

**MOLECULAR CHARACTERIZATION OF GROUP B *STREPTOCOCCUS*
(GBS) AND ITS ASSOCIATION WITH VAGINAL MICROBIOME
AMONG PREGNANT WOMEN FROM LOW-MIDDLE INCOME
COUNTRIES**

Takudzwa Matuvhunye

Student Number: 1933509

**A thesis submitted to the Faculty of Health Sciences, University of the
Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of
Doctor of Philosophy**

Johannesburg 2024

DECLARATION

I, Takudzwa Matuvhunye declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

.....

27th day of August 2024

DEDICATION

I dedicate this work to my beloved wife Yananayi Davu, my unwavering pillar of strength and boundless joy, whose steadfast support has served as the bedrock of my PhD research journey.

To my beloved children, Loiusa Akudzwe (6) and Liam Adiswe (1), whom I dearly cherish and whose innocent smiles and love have been a constant source of inspiration and motivation.

I extend my deepest gratitude to my father Edward Matuvhunye, whose enduring and unswerving support has remained constant throughout both my childhood and in my academic endeavors.

In loving memory of my parents Tsitsi Gwenzi Matuvhunye and Themba Matuvhunye, whose indelible legacy and belief in education and inspiration fueled my passion for a career in life sciences.

PRESENTATIONS ARISING FROM THIS STUDY

1. **Matuvhunye T**, Hinds J, Ali MM, Chaka TE, Berkley JA, Tapia MD, Samba SO, Obara S, Medugu N, Saha SK, Islam M, Izu A, Islam M, Gould K, Kwatra G and Madhi SA. Use of microarray to determine prevalence of Group B *streptococcus* colonization by multiple serotypes in pregnant women from low and middle-income countries. International Symposium on *Streptococcus agalactiae* Disease (ISSAD), Rio de Janeiro, Brazil. October 2023 (Poster Presentation).
2. **Matuvhunye T**, Ali MM, Chaka TE, Berkley JA, Heyderman R, Izu A, Madhi SA and Kwatra G. Association between vaginal Group B *Streptococcus* colonization and the vaginal microbiome in pregnant women from low and middle-income countries of sub-Saharan Africa. University of Witwatersrand Faculty of Health Sciences research day. Johannesburg, South Africa. September 2022 (Poster presentation).
3. **Matuvhunye T**, Hinds J, Ali MM, Chaka TE, Berkley JA, Tapia MD, Samba SO, Obara S, Medugu N, Saha SK, Islam M, Izu A, Islam M, Gould K, Kwatra G and Madhi SA. Vaginal carriage of multiple Group B *Streptococcus* serotypes in pregnant women from low and middle-income countries. University of Witwatersrand Faculty of Health Sciences research day. Johannesburg, South Africa. September 2022 (Poster presentation).
4. **Matuvhunye T**, Ali MM, Chaka TE, Berkley JA, Heyderman R, Izu A, Madhi SA and Kwatra G. Association between Vaginal Group b *streptococcus* colonization and the vaginal microbiome in pregnant women. International Symposium on *Streptococcus agalactiae* Disease (ISSAD), London, United Kingdom. August 2021 (Poster presentation).

PUBLICATIONS ARISING FROM THIS STUDY

Three publications are currently being finalized for submission:

Paper I

Matuvhunye T, Ali MM, Chaka TE, Berkley JA, Heyderman R, Izu A, Kwatra G and Madhi SA. Association between vaginal Group B *Streptococcus* colonization and the vaginal microbiome in pregnant women from low and middle-income countries of sub-Saharan Africa (Manuscript under preparation).

Paper II

Matuvhunye T, Hinds J, Ali MM, Chaka TE, Berkley JA, Tapia MD, Samba SO, Obara S, Medugu N, Saha SK, Islam M, Izu A, Islam M, Gould K, Kwatra G and Madhi SA. Use of microarray to determine prevalence of Group b *streptococcus* colonization by multiple serotypes in pregnant women from low- and middle-income countries (Manuscript under preparation).

Paper III

Matuvhunye T, Dorota J, Ali MM, Chaka TE, Berkley JA, Bhasit U, Tapia MD, Sow O Samba, Obara S, Medugu N, Saha SK, Islam M,, Bently S, Madhi SA, and Kwatra G. Genetic diversity of colonizing Group B *Streptococcus* sequence types and clonal complexes in maternal and newborn isolates from 9 countries in Africa and South-East Asia. (Manuscript under preparation).

ABSTRACT

Introduction: Recto-vaginal colonization by Group B *streptococcus* (GBS) in expectant women is a risk factor for invasive GBS disease in newborns and may predispose to stillbirths and preterm births. The focus of the research included: i. Determine the genetic composition of colonizing GBS sequence types (STs) and clonal complexes (CCs) in pregnant women and their newborns from six sub-Saharan African and two Southeast Asian low and middle-income countries (LMICs); ii. Determine the prevalence of concurrent colonization by more than one GBS serotypes using microarray; iii. Investigate the association between vaginal GBS colonization and the vaginal microbiome in the pregnant women from LMICs.

Methods: Group B *streptococcus* isolates cultured from rectum or vagina of pregnant women (n=816) and skin surface or throat mucosal colonizing isolates from their newborns (n=628) immediately following delivery were whole genome sequenced using next generation sequencing platform (Illumina HiSeq 2000; Illumina, USA) and analyzed using bio-informatic tools, including FastQC for quality control and the short-read sequence typing tool (SRST) to determine the capsular serotypes, sequence types and antimicrobial resistance genes (AMR). Vaginal swabs from pregnant women from Kenya, Ethiopia, Mali, Bangladesh, Nigeria and South Africa were analyzed for vaginal colonization by more than one GBS serotype using microarray, and the vaginal microbiome was investigated using open array quantitative polymerase chain reaction (qPCR) methods.

Results: A total of 53 STs in pregnant women and 51 STs in newborns were identified, including five new described STs. There was high concordance at the serotype (90%; 547/608), ST (88%; 533/608) and CC level (89%; 537/608) between the maternal and newborn isolates. Consequently, further detailed whole genome sequencing analyses focused on the maternal colonizing isolates. The CC23 was the most common lineage in India (47.4%; 65/137), Mozambique (40.0%; 44/110) and Kenya (29.1%; 32/110). The lineage CC17 was common in Mozambique (27.3%; 30/110), Kenya (28.2%; 31/110) and South Africa (20%; 21/105) but absent in Southeast Asian countries. The clonal complex 255 (3.4%; 3/89), CC889 (13.5%; 12/889), and CC1212 (4.5%; 4/89) were

detected exclusively in Bangladesh, albeit at low frequency. There was within serotype geographic variability in relation to ST and CC between countries. Notably, serotype III GBS isolates were variably associated with the hypervirulent CC17, including high prevalence in Mozambique (97%; 30/31), Kenya (89%; 31/35) and South Africa (77%; 17/22), modest prevalence in Mali (45%; 9/20), Nigeria (35%; 6/17) and Ethiopia (33.3%; 2/6); whilst it was not identified in any of eleven serotype III isolates from the Southeast Asian countries. All serotype III isolates (66.9%; 95/142) positive for the *hypervirulent A* gene (*hvgA* gene) were associated with CC17. A rare genotype CC934 (0.2%; 2/816) which also includes the *hvgA* gene, was identified in Ethiopia. The prevalence of any alpha-like family genes (*alp1*, *alp2/3*, *alphaC*, and *rib*) ranged from 80.9% (72/89) in Bangladesh to 100% (110/110) in Kenya. The erythromycin resistance gene was modestly prevalent in Bangladesh (31.5%; 8/89) and Mozambique (20.9%; 23/110), whilst the prevalence thereof was less than 5% in Ethiopia, India, Nigeria, Mali, and Kenya.

Analysis of the vaginal microbiome was done using open array qPCR-based method, which was targeted to identify 52 organisms. Investigation of the vaginal microbiome was done of vaginal swabs in women from Kenya, Ethiopia and South Africa. Four community state types (CSTs) were identified, including CST I dominated by *Lactobacillus crispatus* (3.9%; n=11/282), CST III (*Lactobacillus inners* dominant, 7.1%; n=20/282), CST IV (dominated by diverse anerobic bacteria, 87.2%; n=246/282), and CST V (*Lactobacillus jensenii* dominant, 1.8%; n=5/282). The prevalence of GBS colonization was commonly associated with CST IV-A (56.4%; n=22/39), less common in CST III (30%; n=6/20) but absent in women with CST V (n=5). Colonization with *Veillonella parvula* (adjusted odds ratio, 2.89; 95% confidence interval, 1.53-5.43), *Staphylococcus aureus* (aOR, 12.3; 95% CI, 5.31-28.24) and *Candida albicans* (aOR, 3.98; 95% CI, 2.33-6.79) were positively associated with vaginal GBS colonization, whilst no association was evident for other organisms. Among women with GBS vaginal colonization, more than one serotype was detected using microarray in 10% (n=72/702) of women and ranged from 3.7% (n=2/54) in Bangladesh to 14.4% (n=20/139) in Mali.

Conclusion: Differences in the prevalence of GBS STs and CCs between low and middle-income countries of Southeast Asia and sub-Saharan Africa may account for the regional differences in the burden of invasive GBS disease between countries. In particular, the low prevalence of CC17 associated serotype III isolates among Southeast Asian countries (India and Bangladesh), and the

high prevalence of CC17 serotype III isolates from African countries (including South Africa, Mozambique, Nigeria, and Kenya) may be contributing to the reported differences in the incidence of early-onset GBS disease between the sub-Saharan African and Southeast Asia countries. Vaginal colonization by more than one GBS serotype is low in pregnant women from LMICs. The association between vaginal GBS colonization and other bacteria colonizing the vagina in pregnant women, suggest that the maternal vaginal microbiome may also influence the risk of invasive GBS disease in their newborns.

ACKNOWLEDGEMENTS

I acknowledge:

- My supervisors Professor Shabir A Madhi and Dr Gaurav Kwatra for their support and guidance throughout my PhD studies, including arranging trips abroad for further training at Wellcome Sanger Institute and at Georges University College of London, United Kingdom.
- Dr Jason Hinds and Dr Katherine Gould at St. Georges University College of London for guidance and training on microarray assays and for affording me an opportunity to use the Bioscience Microarray facilities.
- Professor Stephen Bentley, Dr Dorota Jamroz and Dr Bhasit Uzma at the Wellcome Sanger Institute for organizing in-house training on analysis of whole genome sequence data and for permission to use the high-performance computing facilities hosted at the Wellcome Sanger Institute.
- Collaborators of the GBS multicenter study and JUNO project
- Postgraduate students, laboratory staff and the management staff at University of Witwatersrand Vaccines and Infectious Diseases (Wits-VIDA).

The funders:

- South African Medical Research Council Vaccines and Infectious Diseases Analytics Research Unit (Wits-VIDA)
- Department of Science and Technology/ National Research Foundation: Vaccine Preventable Diseases Research Chair.
- Mucosal and Pathogens Research National Institutes of Health (NIH).
- The Wellcome Trust Sanger Institute.

The mothers and their infants who participated in the multicenter GBS study

TABLE OF CONTENTS

DECLARATION	ii
DEDICATION	iii
PRESENTATIONS ARISING FROM THIS STUDY	iv
PUBLICATIONS ARISING FROM THIS STUDY	v
ABSTRACT	vi
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS	x
LIST OF FIGURES.....	xiv
LIST OF TABLES.....	xvi
ABBREVIATIONS.....	xviii
PREFACE	xx
STUDENT CONTRIBUTIONS	xxxii
Chapter 1.0 Introduction	1
1.1 Burden of <i>Streptococcus agalactiae</i>	1
1.2 Microbiology of <i>Streptococcus agalactiae</i>	2
1.3 Genome structure of <i>Streptococcus agalactiae</i>	2
1.4.1 GBS Capsular polysaccharides	3
1.4.2 Surface proteins.....	4
1.5 Pathogenesis of neonatal GBS infection.....	7
1.6 Prevalence of maternal GBS colonization	8
1.7 Density of GBS vaginal colonization	9
1.8 Serotype distribution of colonizing isolates.....	9
1.9 Invasive potential of colonizing serotypes.....	10
1.10 Carriage of multiple GBS serotypes	11

1.11 Detection of multiple serotype carriage.....	14
1.11.1 Serological assays	14
1.11.2 Polymerase chain reaction (PCR) based methods.....	14
1.11.3 Microarray.....	14
1.12 Population Structure of GBS	15
1.13 Whole genome sequencing analysis	19
1.14 Recombination	19
1.15 Antimicrobials in GBS isolates.....	20
1.15.1 Penicillin.....	20
1.15.2 Macrolide and lincosamides antibiotics	21
1.15.3 Aminoglycosides and glycopeptides.....	21
1.15.4 Tetracyclines	22
1.16 Vaginal microbiome.....	22
1.17 The vaginal microbiome community state types	23
1.18 Major vaginal infections	26
1.18.1 Bacterial vaginosis (BV)	26
1.18.2 Sexually transmitted infections (STIs).....	26
1.19 Vaginal microbiome in pregnant women.....	27
1.20 The relationship between the vaginal microbiome and Group B <i>Streptococcus</i> colonization.....	27
1.21 Molecular analysis of the vaginal microbiome.....	32
1.22 Justification of the study	33
1.23 Aims of the study	34
Chapter 2.0 Methods	35
2.1 Study population	35
2.2 Methods for whole genome sequencing analysis	38

2.2.1 Selection of study samples	38
2.2.2 Culturing and Genomic DNA Extraction.....	41
2.2.3 Whole genome sequencing.....	41
2.2.4 Typing of the capsular serotype, surface protein genes and sequence types	41
2.2.5 Recombination detection and phylogenetic analysis of GBS isolates	42
2.2.6 Statistical analysis	42
2.3 Methods for analysis of multiple serotypes	42
2.3.1 Selection of study samples	42
2.3.2 Culturing of vaginal samples and Genomic DNA extraction	44
2.3.3 DNA Fragmentation and labelling	44
2.3.4 Hybridization.....	44
2.3.5 Slide washing, scanning and serotype calls	45
2.3.6 Statistical analysis	45
2.4 Methods for analysis of the vaginal microbiome.....	45
2.4.1 Selection of study samples and sample size	45
2.4.2. Selection of open array target organism and validation assays.....	46
2.4.3 Nucleic acid extraction.....	51
2.4.4 Quantitative Polymerase Chain Reaction.....	51
2.4.5 Statistical analysis	52
2.5 Ethical approval	52
2.6 Funding	53
Chapter 3 Genetic diversity of colonizing Group B <i>Streptococcus</i> isolates in maternal and newborn isolates from 8 countries in Africa and South-East Asia	54
3.1 Abstract.....	54
3.2 Introduction.....	55
3.3 Results.....	56

3.4 Discussion	86
Chapter 4 Use of microarray to determine prevalence of Group b streptococcus colonization by multiple serotypes in pregnant women from low- and middle-income countries	90
4.1 Abstract	90
4.2 Introduction.....	90
4.3 Results.....	91
4.4 Discussion	98
Chapter 5 Association between vaginal Group B <i>Streptococcus</i> colonization and the vaginal microbiome in pregnant women from low and middle-income countries of sub-Saharan Africa	101
5.1 Abstract	101
5.2 Introduction.....	102
5.3 Results.....	103
5.4 Discussion	122
Chapter 6.0 Integrated Discussion and Conclusion	125
Chapter 7.0 References	129
Appendices	162
Appendix 1.....	162
Appendix 2.....	168
Appendix 3.....	168
Appendix 4.....	178
Appendix 5.....	182
Appendix 6.....	189
Appendix 7.....	192

LIST OF FIGURES

Figure 1. 1: The capsular gene organization of nine Group B <i>streptococcus</i> serotypes Ia, Ib, II - IX.	3
Figure 1. 2: Schematic representation of virulence factors in GBS.	5
Figure 1. 3: Stages in the development of neonatal GBS infection.....	8
Figure 1. 4: Global GBS colonization prevalence.....	9
Figure 1. 5: Steps involved in a microarray experiment.....	15
Figure 1.6: The role of <i>Lactobacillus</i> species in preventing colonization by potential pathogens on the vaginal epithelium	23
Figure 1.7: Characteristics of the vaginal microbiome community state types and groups.	24
Figure 1. 8: Vaginal microbiome community state groups within different ethnic groups of women	25
Figure 2. 1: PhD Study design.....	37
Figure 2. 2: Flow diagram of the maternal GBS isolates selected for analyses by whole genome sequencing.	39
Figure 2. 3: Flow diagram of the GBS isolates selected from newborns for analyses by whole genome sequencing.....	40
Figure 2.4: Study cohort diagram.....	43
Figure 3. 3: The phylogenetic tree of GBS serotype III sequence types and CCs in pregnant women from sub-Saharan Africa and Southeast Asia.....	66
Figure 3. 4: The prevalence of CCs in serotype III maternal colonizing isolates from sub-Saharan Africa and Southeast Asian countries.....	667
Figure 3. 5: The STs associated with serotype III in maternal GBS colonizing isolates from sub-Saharan Africa and South Asian countries.	68
Figure 3. 6: The number and phylogenetic tree of <i>hvgA</i> positive isolates in pregnant women from African and South-east Asian countries.	70
Figure 3. 7: The position of the <i>hvgA</i> gene and pathogenic proteins in serotype Ia colonizing isolates.	72
Figure 3. 8: Recombination analysis of <i>hvgA</i> positive maternal colonising isolates.	75

Figure 3. 9: Prevalence of the alpha surface protein genes in colonizing GBS isolates from Southeast Asia and African countries.....	79
Figure 3. 10: Prevalence of the alpha surface protein genes in CCs of colonizing GBS isolates from Southeast Asia and African countries.....	85
Figure 3. 11: The prevalence of antimicrobial resistance GBS isolates colonizing pregnant women from African countries and Southeast Asian countries.	85
Figure 3. 12: The prevalence of antimicrobial resistance GBS CCs colonizing pregnant women from African countries and Southeast Asian countries.	85
Figure 4. 1: Combinations of serotypes in vaginal swabs colonized with more than one serotype.	93
Figure 4. 2: Difference in relative abundance of serotypes among samples colonized with more than one serotype.	94

LIST OF TABLES

Table 1. 1: Studies on carriage of multiple GBS serotypes.....	12
Table 1. 2: The distribution of GBS clonal complexes in infant invasive disease across the globe	18
Table 1. 3: Studies on the vaginal microbiome associated with colonization by Group B <i>Streptococcus</i>	29
Table 2. 1: Vaginal organisms targeted in the 56-targets open array qPCR	47
Table 2. 2: Performance of open array assays	49
Table 3. 1: The number of Group B <i>streptococcus</i> colonizing isolates in mother and newborn dyads sequenced from eight low and middle-income countries of Southeast Asia and Africa.....	57
Table 3. 2: The prevalence of surface proteins in GBS maternal colonizing isolates from eight low and middle-income countries of Southeast Asia and Africa	80
Table 3. 3: Surface proteins targets association with colonizing GBS serotypes in pregnant women from Southeast Asia and Africa.....	83
Table 4. 1: Prevalence of multiple GBS serotypes in countries of low and middle-income countries.	92
Table 4. 2: Serotypes in samples colonized with more than one serotype and their relative abundance.	94
Table 5. 1: Demographic and clinical characteristics of the study population.....	103
Table 5. 2: Site specific analyses of clinical and demographic characteristics of participants in Ethiopia, Kenya, and South Africa	105
Table 5. 3: Association between vaginal organisms and GBS colonization status in pregnant women.	110
Table 5. 4: Site-specific analysis of the association between vaginal organisms and GBS colonization in Ethiopia.....	113
Table 5. 5: Site-specific analysis of the association between vaginal organisms and GBS colonization in Kenya.	115

Table 5. 6: Site-specific analysis of the association between vaginal organisms and GBS colonization in South Africa.....	117
Table 5. 7: Distribution of community state types grouped by GBS status.	121

ABBREVIATIONS

aOR	Adjusted odds ratios.
Alp	Alpha-like protein
BV	Bacterial vaginosis
CAMP	Christie Atkinson Munch-Petersen test
CCs.	Clonal complexes
CDC	Center for Disease Control
CPS	Capsular polysaccharides
DNA	Deoxyribonucleic Acid
EOD	Early onset disease (0-6 days)
GBS	Group B <i>streptococcus</i> (<i>Streptococcus agalactiae</i>)
IQR	Interquartile range
<i>HvgA</i>	<i>Hypervirulent adhesion</i>
HIV	Human Immunodeficiency syndrome
HPV	Human papilloma virus
HSV	Herpes simplex virus
IAP	Intrapartum antibiotic prophylaxis
LMICs	Low- middle income countries
LOD	Late onset disease (7-89 days)
MLST	Multi-locus sequence typing
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
PFGE	Pulse-field gel electrophoresis

PI	Pilus
RIB	Resistance to proteases immunity group B
RDP	Restriction digestion patterns
STGG	Medium containing skim milk, tryptone, glucose, and glycerin
STI	Sexual transmitted Infection
STs	Sequence types
UK	United Kingdom
USA	United States of America
WHO	World Health Organization
Wits-VIDA	University of Witwatersrand Vaccines and Infectious Diseases

PREFACE

This thesis is presented to the reader in the University of the Witwatersrand's recommended "divided block" format. In this format, the thesis consists of two parts. The first part includes the Introduction and Methods chapters that are written in the traditional thesis format, and second part in which the Results and Discussion of the study's objectives are presented as individual chapters. The study aimed to evaluate the genetic diversity sequence types and clonal complexes of GBS colonizing isolates in pregnant women and their newborns from eight low and middle-income countries of sub-Saharan Africa and Southeast Asia. Furthermore, we evaluated the prevalence of colonization by multiple serotypes in African women at the time of delivery. Thirdly, I evaluated the association between vaginal GBS colonization and vaginal microbiome at the time of delivery.

In Chapter 1, there is an introduction on the burden of GBS disease, microbiology of GBS, distribution and invasive potential of colonizing GBS serotypes. Furthermore, there is an overview on prevalence of colonization by more than one GBS serotypes, including methods commonly used in detecting multiple serotypes. Also, the population genetic structure of GBS colonizing isolates and the association of GBS to the rest of the vaginal microbiome is reviewed.

In Chapter Two, there is an overview of the materials and methods used for the studies undertaken. The methods undertaken include culturing of isolates and DNA extraction for whole genome sequencing. Also, the methods for analyzing the vaginal microbiome by Open array methods; and investigation for detection of multiple GBS colonizing serotypes using micro-array methods is described.

The findings of the studies are detailed in Chapters 3 to 5. In chapter 3, the genetic variability of GBS STs and CCs in pregnant women from eight low and middle-income countries of Southeast Asia and sub-Saharan Africa is detailed. Furthermore, the association of capsular serotypes and GBS sequence types and clonal complexes is analyzed, as is the prevalence of genes coding for surface protein antigen targets which are potential vaccine candidates, and antimicrobial resistance genes.

In chapter 4, the use of microarray to investigate the prevalence of colonization by more than one serotype in pregnant women colonized by GBS is described.

Finally, in chapter 5, the association between vaginal GBS colonization and the vaginal microbiome is evaluated for women from three African countries (Kenya, Ethiopia and South Africa).

The thesis concludes with chapter 6, with a summary of the main findings of my research and discusses the significance and highlights research gaps.

STUDENT CONTRIBUTIONS

Protocol and ethics

Takudzwa Matuvhunye (TM) was responsible for obtaining approvals for the study including ethics, protocol and export permits for study samples.

Chapter 3: Genetic diversity of colonizing Group B Streptococcus isolates in maternal and newborn isolates from 8 countries in Africa and South-East Asia.

TM was responsible for sub-culturing the archived GBS isolates, DNA extraction, shipping the extracted DNA for whole genome sequencing at the Wellcome Trust Sanger Institute, bioinformatic analysis of sequenced data and statistical analysis.

Chapter 4: Use of microarray to determine prevalence of Group b streptococcus colonization by multiple serotypes in pregnant women from low- and middle-income countries.

TM was responsible for locating the archived vaginal swab samples, aliquoting and shipping the vaginal swabs to St. Georges College of London in the United Kingdom, performing microarray analysis using the Bioscience microarray facilities and statistical analysis of the results.

Chapter 5 Association between vaginal Group B Streptococcus colonization and the vaginal microbiome in pregnant women from low and middle-income countries of sub-Saharan Africa.

TM conducted a literature search on the vaginal microbiome associated with GBS to select vaginal microbiome targets. TM performed nucleic acid extraction, detection of the vaginal microbiome targets using open array assays and statistical analysis of the results.

Chapter 1.0 Introduction

1.1 Burden of *Streptococcus agalactiae*

Since the 1970s, Group B *Streptococcus* (*Streptococcus agalactiae*; GBS) has been the main cause of invasive disease in newborns (1). GBS normally colonizes the genitourinary and gastrointestinal tracts, with approximately 15-35% of pregnant women being colonized in the rectum or vagina (2, 3). Vertical transmission of GBS from recto-vaginal colonized expectant women to their newborns occurs before or during delivery, which could lead to early-onset invasive disease (EOD; within 0-6 days after birth) in 1-3% of the newborns. The case-fatality risk of EOD is between 5 to 27% (4, 5). Most cases of EOD (80%) occur within the first day of life, suggesting that GBS is acquired in utero (6). The remaining cases of invasive GBS disease in infants usually occurs between 7 to 89 days of life (i.e. late-onset disease; LOD), which could be due to horizontal GBS acquisition from breast milk, community or nosocomial sources (7), and has a case-fatality risk of 2 to 20% (1, 8).

The incidence of GBS disease in young infants (<90 days after birth) range between 0.5 to 2.0 per 1,000 live births globally, and manifests clinically as sepsis, meningitis, or pneumonia (9-11). Sepsis is the predominant clinical syndrome in EOD (80-85%), whereas 25-30% of LOD present as meningitis (12). Approximately 50% of infants who survive invasive meningitis develop chronic developmental abnormalities and neurologic sequelae such as seizures, cognitive impairment, hearing loss, and blindness (13).

The current strategy for preventing EOD involves testing of pregnant women for recto-vaginal colonization by GBS at 34 to 37 weeks of gestation age, followed by use of intrapartum antibiotic prophylaxis (IAP) in women identified to be colonized at least two hours before delivery. Furthermore, IAP is recommended for expectant mothers with other risk factors for EOD in their newborns; i.e. a risk-based approach to IAP. The incidence of EOD decreased by 90%, from 1.8 to 0.23 cases per 1000 live births between 1990 and 2015 in United States of America (USA), after the introduction of routine screening of expectant mothers for GBS colonization and implementation of IAP (14, 15). Despite the decrease in EOD with the use of IAP, GBS remains a major cause of sepsis in newborns from developed countries, as IAP does not lower risk of LOD for which the incidence is 0.31 per 1000 live births in USA (15). In low-income settings with limited preventive strategies, the GBS disease burden is still high, with the African region accounting for 65% of all GBS deaths in infants (5).

The global incidence (per 1000 live births) of invasive GBS disease in young infants (<90 days age) is estimated to be 0.49 with the highest incidence in Southern Africa (2.0) and lowest in Southeast Asia (0.21) (16). In the Caribbean and Latin America, the incidence (per 1000 live births) of invasive GBS disease is 0.49 (4). Incidence rate (cases per 1,000 live births) also varies by country within Africa, with South Africa having higher incidence (2-3) compared with The Gambia (1.33) and Nigeria (0.25) (4, 17, 18). The estimated incidence of GBS LOD was higher in Southern Africa (0.93), compared to the global incidence (0.26).

1.2 Microbiology of *Streptococcus agalactiae*

Streptococcus agalactiae is a lancet-shaped gram-positive cocci that grows in pairs or short chains and divides in a single plane. The diameter of *Streptococci* colonies are 1-4 mm in diameter, surrounded by a distinct mucoid zone and exhibit a narrow zone of beta-hemolysis on blood agar (19). Colonies that are typical of GBS morphology and exhibit a non-beta-hemolytic growth pattern can be further tested by confirmation of Lancefield group B antigen. The GBS is the only *Streptococcus* species that is reactive to the serogroup B antigen immunoprecipitation test. The Christie Atkinson Munch-Petersen test (CAMP) is also employed for the detection of GBS (20), wherein the test colonies are plated vertical to the *Staphylococcus aureus* streak on blood agar. GBS results in an enhanced arrow-head shaped haemolysis when two streaks are in close proximity (21).

1.3 Genome structure of *Streptococcus agalactiae*

The first complete genome of GBS strain NEM316 comprising 2118 protein-coding genes and 2 211 485 base pairs (bp), was published in 2002 (22). Over 7000 draft genomes of GBS have been sequenced to date, with genome sizes range from 1.8 to 2.3 megabase (Mb) and guanine-cytosine GC composition of 35% to 36% (12). Approximately 80% of the core genes are conserved across all GBS species forming a pan-genome which constitute a set of genes common across all species in the same genus or clade. The accessory genes increase with each newly sequenced clade, while the core genes remain constant. Glacier and colleagues reported that 50% of genes in GBS strains are orthologs with biological functions similar to *Streptococcus pyogenes* genes (22).

GBS possesses a high number of tRNAs (80), signal transduction systems, and ABC transporters that are critical for its survival and pathogenicity in the human host (62). The genome contains thirty open reading frames encoding putative proteins with a signaling motif for cell wall sorting

(22). The *cyl* locus comprises genes for cytolysin or hemolysin toxin synthesis, implicated in pulmonary epithelial cell invasion (23). The capsular operon is responsible for capsular polysaccharide synthesis and bacterial encapsulation via the Wzx/Wzy-dependent pathway. Conserved genes within the capsule operon encompass the *neu* genes (*neuA*, *neuB*, *neuC*, and *neuD*), responsible for sialic acid synthesis and the capsule's subsequent sialylation (24). Capsular gene regions spanning *cpsA* to *cpsE* are conserved among all GBS serotypes, while variable capsular genes (*cpsF*, *cpsG*, *cpsH*, *cpsI*, *cpsJ*, *cpsK*, *cpsM*) encode serotype-specific glycosyltransferases and polysaccharides (25). Molecular techniques such as qPCR target the capsular genes to detect GBS serotypes (Figure 1.1).

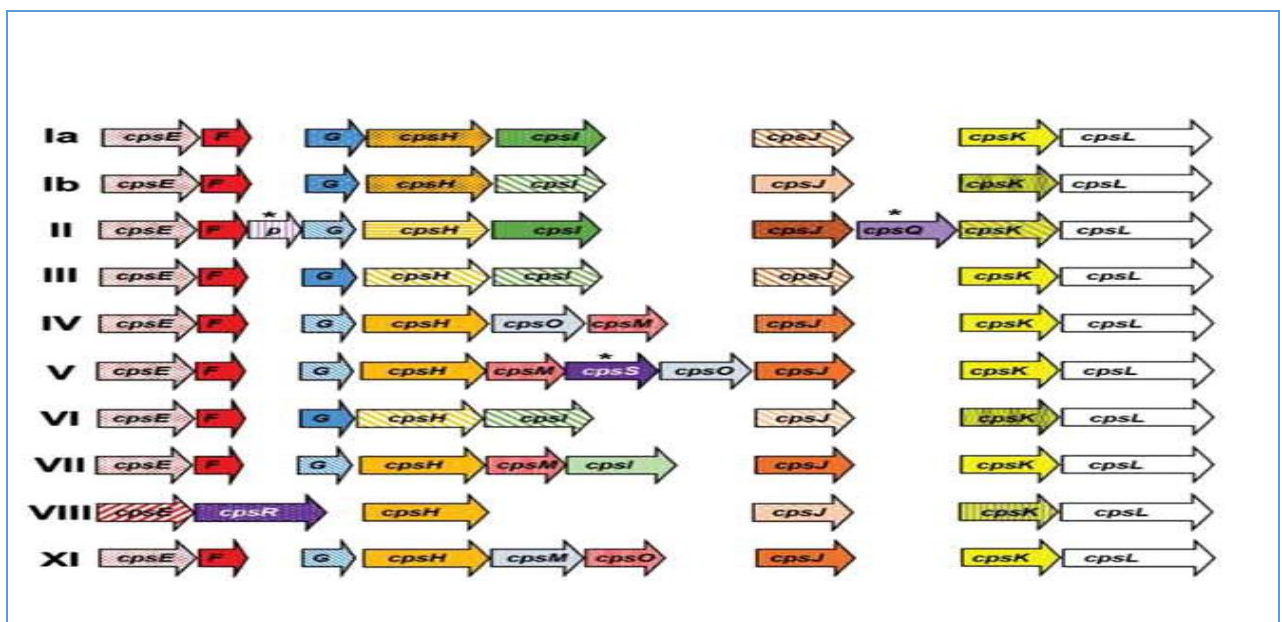


Figure 1. 1: The capsular gene organization of nine Group B *streptococcus* serotypes Ia, Ib, II -IX (12).

1.4 Major GBS virulence factors

1.4.1 GBS Capsular polysaccharides

Capsular polysaccharide (CPS) is a peptidoglycan-anchored antigen, extensively studied as a virulence factor contributing to GBS pathogenesis (13). The CPS increases bacterial invasiveness and facilitates evasion of host defenses through various mechanisms, including mediating biofilm formation and impeding the binding and interaction of antimicrobial peptides and neutrophil extracellular traps (12). Group B *streptococcus* is classified into ten distinct serotypes (Ia, Ib, II, III, IV, V, VI, VII, VIII and IX) based on the immunologic reactivity of CPS (26). The most recent

addition, serotype IX, was identified in 2007 from the blood of a neonate with GBS invasive disease (27).

The CPS consists of tightly-packed repeating units of galactose, glucose, sialic acid and N-acetylglucosamine (22). Isolates of serotype VI are an exception, as it lacks N-acetylglucosamine, whereas in serotype VIII N-acetylglucosamine is replaced by a rhamnose residue. The CPS are terminally sialylated (N-acetylneuraminic acid) with short side chains that are important for evading the host immune system by interfering with opsonic components of the complement system, thus preventing recognition by the host's innate immune system (28). The composition and quantity of CPS determine the virulence and invasiveness of GBS strains. Most GBS isolates causing disease in humans are encapsulated, and the CPS is usually denser in invasive isolates than in colonizing isolates (29). Nevertheless, 5-20% of GBS isolates from humans lack CPS due to loss of a capsular polysaccharide expression system, which is more so in colonizing than invasive isolates (30).

The capsular polysaccharide is a major epitope in GBS vaccine development because of its immunological role in the pathogenesis of invasive GBS disease. Early studies by Rebecca Lancefield revealed that polyclonal rabbit serum specific to capsular polysaccharides was protective against GBS disease in mice (28). Several years later, Baker and Kasper provided evidence that maternal vaccination with the capsular polysaccharide antigens could potentially protect against invasive GBS disease in newborns (31). Coupling of the native GBS polysaccharides, which is a weak immunogen on its own, with a protein carrier such as CRM197 (a nontoxic variant of diphtheria toxin) enhance the immune response against CPS (12). Multivalent polysaccharide conjugate vaccines are currently being developed, aimed at immunizing women during pregnancy to protect their neonates against invasive disease from serotypes included in the vaccine (32, 33). A hexavalent polysaccharide conjugate vaccine candidate, which includes serotypes Ia, Ib, and II-V, is currently under development. (34, 35).

1.4.2 Surface proteins

GBS expresses an array of surface-embedded proteins that play a critical role in its invasiveness, including facilitating adherence to vaginal epithelial cells. Among these surface proteins are the alpha-like protein (Alp) family (alphaC, Alp2, Alp3, Alp4, and Rib) (36). Alp-like proteins are involved in binding to the human glycosaminoglycan, adherence and translocation across epithelial

cells (37). The first surface protein identified as a virulence factor and potential vaccine candidate is the alphaC (38-40). The AlphaC proteins are characterized by long, identical repeats of amino acids expressed as either α - and or β -proteins are present mostly in serotypes Ia, Ib and II (41). Rib protein is also expressed by the majority of GBS serotypes II, III, and a subset of serotype V. The Rib antigen induces protective immunity in mice against lethal doses of rib-expressing strains (42). Additional virulence factors comprise of the immunogenic surface proteins such as serine proteases and C5a peptidases, which impair the host defense system by inactivating the human complement component C5a (28). GBS also secretes a hemolysin (CylE) which induces host tissue damage (Figure 1.2).

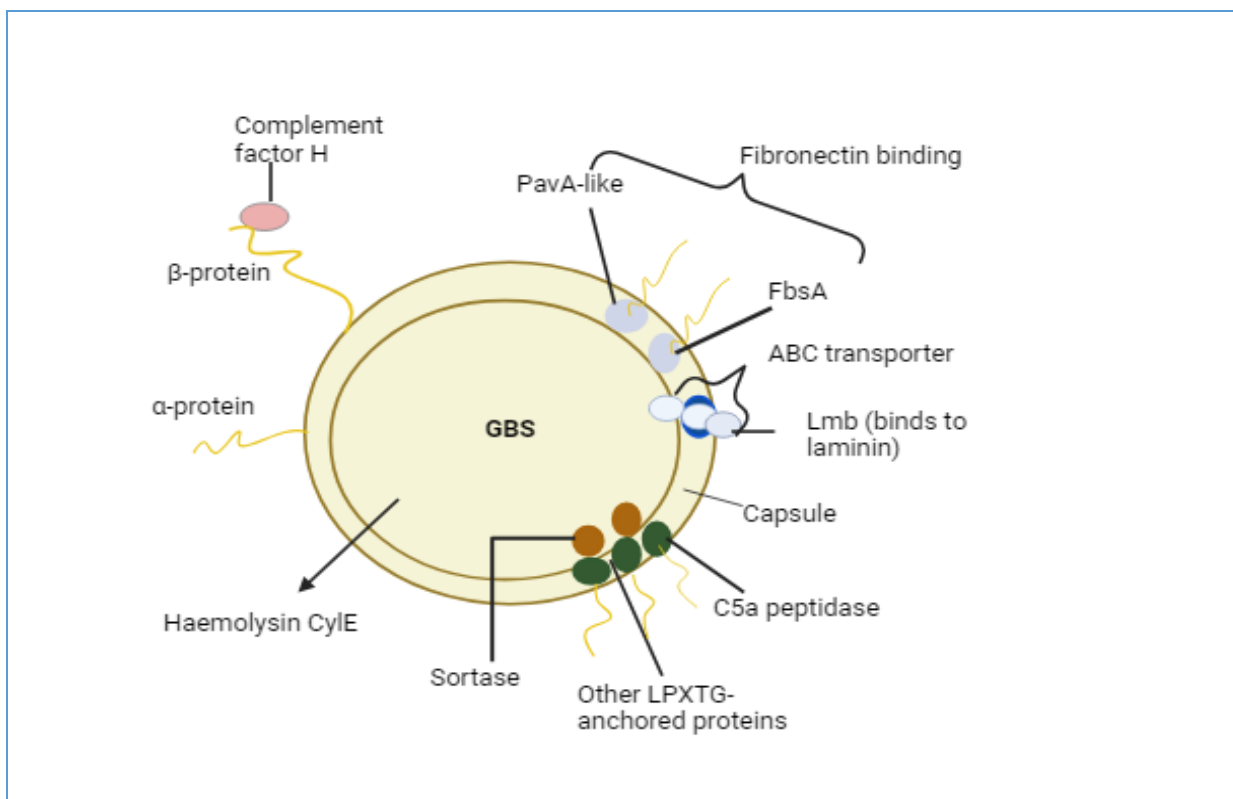


Figure 1. 2: Schematic representation of virulence factors in GBS. Surface and secreted virulence factors in GBS that allow evasion of host immune responses, adapted from Mitchell, 2003 (43).

Another surface-anchored protein is the hypervirulent GBS adhesin (HvgA), which mediates GBS neonatal intestinal colonization and transcellular crossing of the choroid plexus blood–brain barriers, leading to meningitis (13, 44). In mice orally inoculated with GBS, the presence of the

HvgA protein was critical factor in the intestinal colonization and translocation across the intestinal barrier (45). Studies have shown that the majority of LOD cases are caused by GBS isolates that expresses the HvgA protein, being more common in meningitis in EOD (79%) and LOD (83%) (45, 46). Overall, the HvgA and rib proteins are associated with invasiveness of GBS strains (47-50). In a systematics review that analyzed data from 2001 to 2018, at least one of alp 1, alp2/3, alpha C or rib surface protein targets was detected in 87% of GBS isolates in maternal colonization and 93% in infant disease (51). However, the data on surface protein antigen targets was majorly from studies in Europe (45%; 35/78), with a few studies from the sub-Saharan African region (9%; 7/78) and southern Asia (1%; 1/78), which shows paucity of sequence data on vaccine surface protein targets in African and South Asian region (51).

The use of genomics and proteomics has unveiled a range of surface antigens which are potential GBS vaccine candidates. The identification of the pilus in GBS through reverse vaccinology approaches, led to studies on investigating the potential of pilus proteins as vaccine antigens (52, 53). Each of the three pilus variants is encoded by the pilus islands (PI-1, PI-2a, and PI-2b) consisting of five genes (28). Lower levels of naturally occurring antibody to PI-1 and PI-2a, were identified in mothers of newborns with EOD than in controls, indicating PI as potential vaccine antigens (54). A study from South Africa, however, showed no association between maternal antibodies against pilus surface protein and risk of invasive disease in their infants (289). In pre-clinical studies, surface immunogenic protein (Sip) administered to female mice in combination with three other putative surface or secreted proteins (GBS67, GBS80, and GBS104) was associated with 90% (445/495) survival of their pups inoculated with serotypes (Ia, Ib, II, III, V, and VIII), compared with only 18% of pups that survived from unvaccinated dams (55).

The non-capsular based GBS-NN vaccine, consisting of N-terminal domains of alpha-like proteins Rib and alphaC is in clinical trials (phase II), as is the GBS-NN2 prototype vaccine, consisting of fusion proteins of Alp1N and Alp2/3N (56, 57). The development of a protein-based vaccine is being pursued as it offers prospects of overcoming limitations of serotype-specific immunity, absence of protection against non-typeable strains and mitigates the potential of serotype replacement or switching which may transpire with use of polysaccharide capsular based vaccines (58). Protein antigens induce protective T-cell dependent antibody responses, thereby conferring long-lasting immunity, as would also result from conjugation of CPS to protein antigens (28). As

the surface protein-based vaccine are also under development, research is warranted to determine the geographic distribution of surface protein antigens and their association with sequence types.

1.5 Pathogenesis of neonatal GBS infection

The primary risk factor for early-onset GBS disease in newborns is maternal colonization of rectum or vagina by GBS (59), which occurs intermittently or transiently in 10-40% of pregnant women. Approximately 1% to 2% of newborns born to women colonized with GBS at the time of delivery will develop early onset GBS disease (7, 18, 60). Vaginal colonization results from GBS attachment to vaginal epithelial cells (i.e. fibrinogen and laminin in the extracellular matrix). Virulence factors, including capsular polysaccharides, pili, and cell surface-anchored proteins like FbsA, facilitate the attachment of GBS to vaginal epithelial cells (61). GBS vaginal colonization is also coupled with resistance to clearance by mucosal immune defenses (62). Furthermore, GBS is able to adhere to various other human cell types, including the blood-brain endothelium, respiratory tract epithelium and placental membranes (63).

During pregnancy, GBS is able to infiltrate the uterine compartment through ascending infection, which could breach the placental membrane and colonize the fetal skin or mucous membranes (Figure 1.3). Invasion of the amniotic sac by vaginal organisms, including GBS, could lead to placental membrane rupture and induce preterm delivery (61, 63). An alternative route of GBS infection in newborns involves aspiration of into the fetal lungs, either from infected amniotic fluid *in utero* or from vaginal fluid during delivery (61). After infecting the lungs, GBS may cause bacteremia and sepsis in neonates, as well as meningitis following breaching of the blood-brain barrier (64).

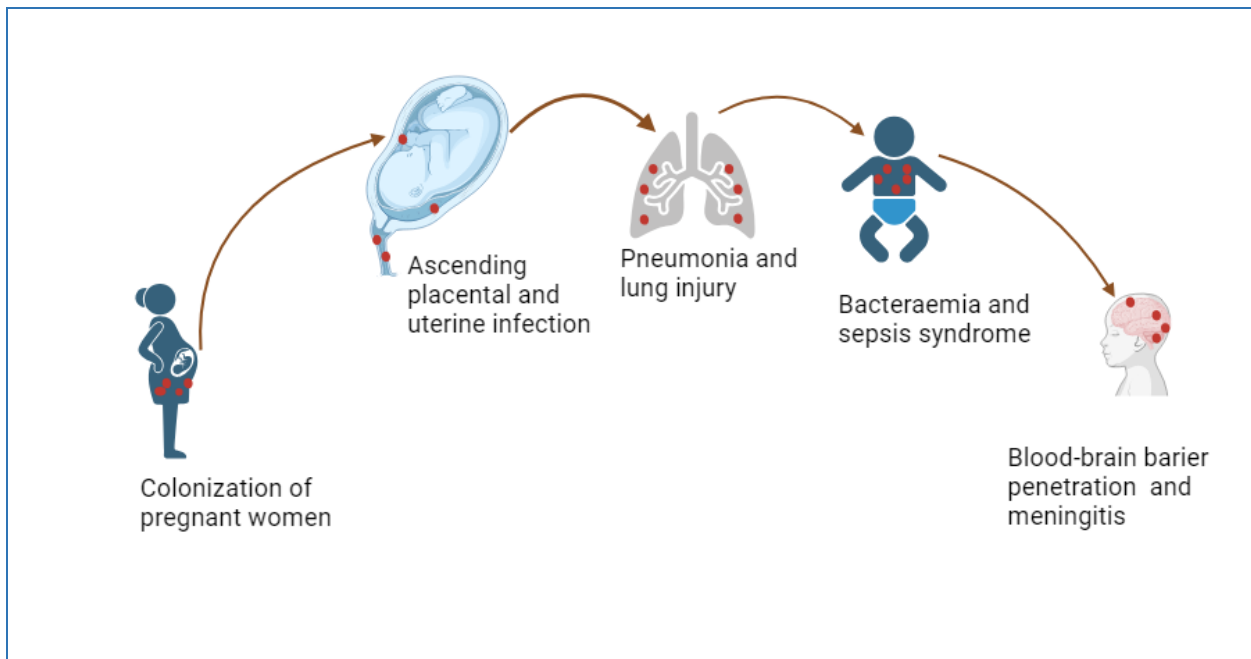


Figure 1. 3: Stages in the development of neonatal GBS infection, adapted from Doran et al, 2004 (61). The dots in red represent GBS.

1.6 Prevalence of maternal GBS colonization

The global prevalence of GBS colonization of the vagina or rectum in pregnant women is 15% to 18% (13, 14), with prevalence being 18.1% to 26.7% in Africa, 6.7% to 22.7% in the Americas (13) and 11.1% in Southeast Asia (Figure 1.4) (13). A recent systematic review reported GBS colonization to being highest in pregnant women from the Caribbean (35%) and lowest in Eastern Asia (11%) (65). The prevalence of recto-vaginal colonization by GBS in pregnant women in the Africa region also demonstrate some heterogeneity, including 25%, 20%, 19% and 14% in southern Africa, northern Africa, central Africa and western Africa, respectively (65). Factors contributing to geographic variability in GBS colonization prevalence among women could include differences in infection acquired capsular antibodies, serotype distribution, genetic predisposition and other environmental factors (65).

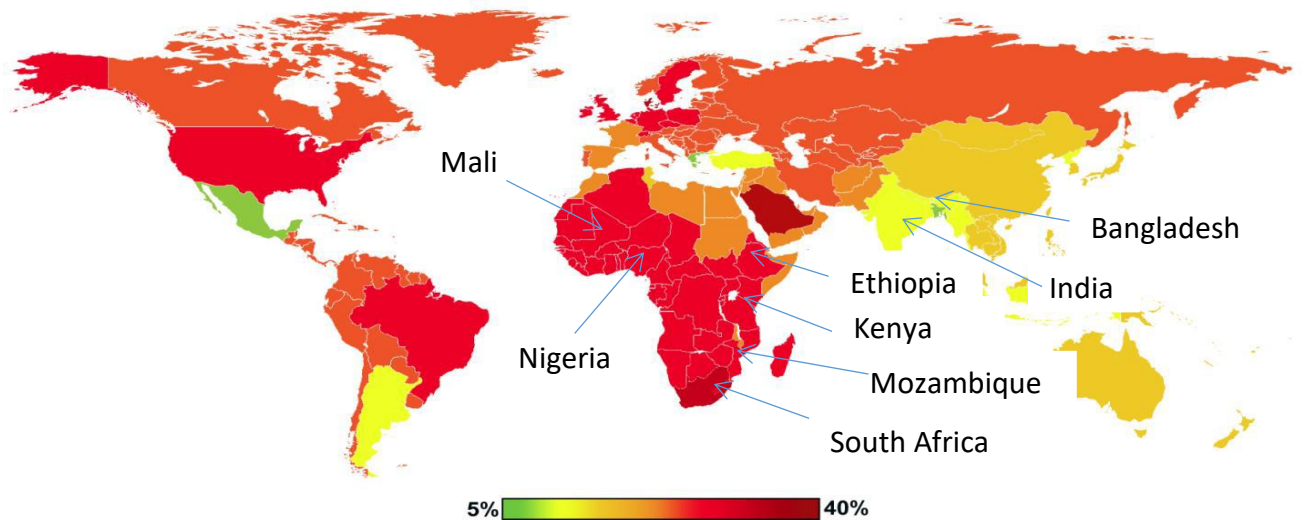


Figure 1. 4: Global GBS colonization prevalence. The heat map shows the prevalence of GBS colonization during pregnancy (prevalence data adapted from Kwatra et al, 2016 (66)). Countries that are shown on the map are the sites for which samples used in the current study were collected.

1.7 Density of GBS vaginal colonization

Rectal and or vaginal GBS colonization in expectant mothers during labor is a significant risk factor in developing invasive disease in the newborn. Pregnant women with higher density of recto-vaginal GBS colonization, variably defined as greater than 10^5 colony forming units per millimeter or growth of GBS colonies in more than two quadrants of semi-quantitative blood agar plates (67), more likely to transmit GBS to their newborns and also at higher risk of delivering prematurely (68). Nevertheless, there is no standardized guidelines for quantifying GBS colonies in vaginal and or rectal samples. The presence of bacteriuria (presence of GBS in urine) is also a surrogate marker suggestive of high density of recto-vaginal colonization in women (69).

1.8 Serotype distribution of colonizing isolates

A hexavalent serotype specific capsular-protein conjugate GBS vaccine is currently being developed. The distribution of GBS serotypes may change over time, hence, ongoing surveillance of circulating serotypes is warranted. The GBS types Ia-III, and V account for the majority of colonizing isolates globally (98%), including in the sub-Saharan region (94%) (70). In South Africa, serotype III (37.3%) is the predominant colonizing serotype, followed by serotypes Ia (30.1%), II (11.3%), V (10.2%), Ib (6.7%), and IV (3.7%) (71). In contrast, serotype distribution

in neighboring Namibia differs, serotype II (60%) being dominant, followed by serotypes III (25%), V (10%), and Ia (3%); illustrating within regional differences in recto-vaginal colonizing serotype distribution (72).

Globally, serotype III is the most common colonizing serotype (25%), albeit varying across regions representing 22% to 33.2% of serotypes in Europe compared with 11%-13% in Asian countries such as Bangladesh and India, as well as in West Africa, the Caribbean and Latin America (51). (51). The other most common maternal colonizing serotypes are Ia (22%) and V (19%) (65, 73). Serotype IV constitute 3% to 4% of colonizing serotypes in Europe, Northern America and Southern Africa (65, 74). Serotypes VI, VII, VIII, and IX are the least prevalent colonizing serotypes overall, albeit constituting 11% to 20% of colonizing serotypes in South Asia and Eastern Asia (Table 1.1) (75). Serotypes VI, VII, VIII, and IX are also commonly detected in Japan (27%-36%) (65), and have also recently been identified as colonizing serotypes in Kenya and South Africa (73, 76). Although serotype distribution has been extensively studied especially in developed countries, there is paucity of data on the prevalence of less common serotypes and of sequencing data from the colonization studies.

1.9 Invasive potential of colonizing serotypes

Serotype III account for 48% to 95% of invasive disease serotypes in newborns, while serotype Ia contributes to 14.2% to 21.8% of cases (77). Serotypes V and II contribute to 5% to 10% of all invasive GBS disease cases in infants globally (5). Collectively, serotypes Ia, Ib, II, III, and V constitute 97% of cases of GBS invasive disease in neonates, and are included in the investigational hexavalent polysaccharide conjugate vaccines under development (5). The detection of capsular type IV in GBS disease (78), contributed to its inclusion in the hexavalent GBS vaccine candidate. The distribution of serotypes causing invasive disease in young infants also vary across geographic regions (4). In the sub-Saharan region, serotypes Ia, III and V are the predominant invasive disease-causing serotypes, with serotype III contributing to 56.0% to 58.3% of cases. In South Africa, the most common disease-causing serotypes in infants are serotypes III (55.4%), Ia (28.2%) and V (7.9%) (79). In Kenya, GBS Ia, Ib, II, III and V cause 98.6% of invasive disease, with serotype III contributing to 60% and 90% of EOD and LOD, respectively (80). Serotypes VI to IX are uncommonly detected serotypes in invasive disease isolates (81).

1.10 Carriage of multiple GBS serotypes

There is limited data on colonization by more than one serotype in pregnant women, particularly in low resourced countries. The reported prevalence of colonization by multiple GBS serotypes vary from 1% to 69.8% between studies, albeit there being variability in methods used to detect multiple colonization (Table 1.1). In a study from The Gambia, the prevalence of colonization by multiple GBS serotypes in pregnant women with GBS carriage was 32.3% (31/96; 95% CI: 24%-42%) (82).

The sensitivity to detect colonization by multiple GBS serotypes can be increased by selecting multiple colonies for serologic testing (83). Nevertheless, serotyping of multiple colonies from the same case is laborious and not commonly undertaken (83). Consequently, molecular-based methods such as quantitative PCR and microarray methods are being investigated to detect multiple colonizing GBS serotypes. Studies evaluating for colonization by multiple serotypes could be informative once GBS vaccines are deployed, to determine whether changes in serotype-colonization are due to replacement by non-vaccine serotypes or unmasking of non-vaccine serotypes which were present at a low density prior to vaccination. Furthermore, colonization by multiple GBS serotypes could promote homologous gene recombination between serotypes, leading to the emergence of non-vaccine serotypes with supra-virulence genes (17). Also, recombination within the capsular polysaccharide locus (vaccine target region) may result in serotype switching and the emergence of capsular polysaccharide locus variants or mutants (84).

Capsular serotype switching involves altering or exchanging the capsular polysaccharide locus, leading to a change in serotype occurs when more than one strain of GBS occupies the same niche concurrently (85). Carriage of multiple serotypes facilitates the dissemination of important genetic determinants, such as antibiotic resistance genes and virulence genes. Although capsular polysaccharide serotype switching is rare among GBS strains, it has been documented (39, 86, 87).

Table 1. 1: Studies on carriage of multiple GBS serotypes

Author / Year	Country	Participants	Specimen type	GBS detection method	Study sample size	GBS Colonization prevalence	More than one serotype
Anthony, 1976 (88)	USA	Pregnant women	Vaginal, Urethra, cervix swabs	Culture and serology	382	15% (57/382)	1% (5/57)
Anthony, 1981 (89)	USA	Pregnant women	Rectovaginal swabs and stool	Serology: micro-immunodiffusion	64	20% (13/64)	1.6% (2/13)
Hoogkamp-Korstanje, 1982 (90)	Netherlands	Pregnant women	Vagina, cervix, and rectum swabs	Culture and serology	762	14% (106/762)	23% (24/106)
Ferrieri, 2004 (91)	USA	Non pregnant women	Vaginal and rectal swabs	Culture and serology	NA	NA	22% (22/102)
Pérez-Ruiz, 2004 (92)	Spain	Pregnant women	Vaginal and rectal swabs	Culture and Pulsed-field gel electrophoresis	NA	NA	3.3% (1/30)
Whitney, 2004 (93)	Thailand, USA, Philippines, Zimbabwe, Myanmar, Ireland,	Pregnant women	Swabs from the cervix and vaginal wall, urine samples	Culture and latex agglutination	1308	11% (148/1308)	1% (1/128)

El Aila, 2009 (94)	Belgium	Pregnant women	Vaginal and rectum swabs	Culture and RAPD PCR	150	24% (36/150)	11% (4/36)
Slotved, 2017 (95)	Ghana	Pregnant women	Recto-vaginal swab samples	Group B latex agglutination	400	27% (107/400)	1% (1/107)
Khatami, 2019 (83)	USA	Non-pregnant, 18-55 years old	Vaginal lavage specimens	The qPCR and latex agglutination	433	21% (91/433)	7% (6/91)
Jisuvei, 2020 (96)	Kenya	Pregnant women	Vagina, rectum swabs	Latex agglutination	292	20.5% (60/292)	70% (37/53)
To, 2021 (82)	The Gambia	Mothers and infants.	Rectovaginal sites, breast milk, infant rectal and	Culture and RAPD PCR	650	39% (253/650)	32% (31/96)

USA: United States of America, RAPD: Randomly amplification of polymorphic DNA analysis, NA: information not available.

1.11 Detection of multiple serotype carriage

1.11.1 Serological assays

Serological assays such as the Lancefield capillary precipitin method, co-agglutination, enzyme immunoassay, double immunodiffusion, and latex agglutination are conventionally used to serotype GBS isolates. The most common method used for GBS serotyping is latex agglutination method, which utilizes anti-rabbit IgG-coated latex particles sensitized and specific to the GBS capsular polysaccharide (97, 98). Serotyping of GBS isolates is typically performed by selecting a single colony per sample and assigning a serotype based on a specific immunologic reaction with the isolate. Nevertheless, the typing accuracy of latex agglutination is suboptimal, with sensitivity below 90% (number of correct results/total number of tests performed) (70) and inability to type 1-5% of serotypes which express low or no CPS (30, 99, 100). Furthermore, testing of only a single colony per sample is sub-optimal to detect simultaneous colonization by multiple serotypes. The detection of multiple GBS serotypes on the sample can be enhanced to detect approximately 90% of all serotypes by serotyping three or more colonies, however, it is labor intensive and costly (91).

1.11.2 Polymerase chain reaction (PCR) based methods.

The PCR-based methods widely used for detecting and serotyping GBS in clinical specimens (101), include the conventional PCR, multiplex PCR and qPCR techniques (102-104). The PCR methods rely on GBS genome sequence information to design primers and probes that target serotype-specific capsular gene regions, allowing the ten GBS capsular types to be detected. The advantages of qPCR-based methods include providing rapid results, high reproducibility and the capacity to directly determine serotypes from the specimen without an initial culturing step, which allows for the detection of non-viable organisms (105). Despite their suitability to detect multiple serotypes, PCR methods have not been widely used (83, 106).

1.11.3 Microarray

Performing a microarray laboratory experiment involves four main steps: DNA extraction from the culture growth, amplification of the target DNA, hybridization of the labelled DNA with oligonucleotide probes immobilized on microarray slides, and scanning of slides and analysis of the results (107) (Figure 1.5). In the initial step, the nucleic acids are labelled with fluorescent dyes, such as the commonly used cyanine dyes Cy5 and Cy3 and hybridized to the array-bound probes. Subsequently, the unbound DNA is removed through washing with buffers (108), to ensure that the signal emitted during image acquisition comes exclusively from the sample bound to the arrays.

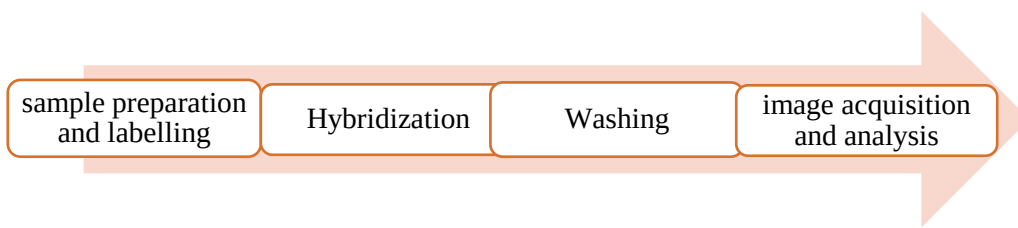


Figure 1. 5: Steps involved in a microarray experiment

The microarray method is a relatively new approach to pathogen typing (109, 110), having been used for detecting *Streptococcus pneumoniae* in colonization studies. In the microarray method, serotypes are identified based on multiple target markers that could span over 200 genes, reducing the possibility of mistyping due to cross-hybridization or single nucleotide polymorphisms. The microarray method is designed to identify colonizing serotypes and their relative abundance by detecting the presence and relative amounts of GBS capsule genes in genomic DNA extracts from clinical samples. Moreover, the microarray method allows for simultaneous analysis, including detection of antimicrobial resistance determinants, surface protein antigens and other co-colonizing organisms.

The advantage of the microarray method is the detection of a higher frequency of identifying colonization with multiple *Streptococcus pneumoniae* serotypes compared with the World Health Organization (WHO) recommended culture-serotyping method (48.8% vs. 11.2%) (111). Among 20 methods (i.e. including culture, real time PCR and reverse line blot) evaluated for identification of multiple serotypes, the molecular serotyping microarray method showed the highest positive predictive value (93.7%) and sensitivity (95.8%) (112). The utility of microarray was also demonstrated by Zurn et al, 2020 in typing of GBS isolates into different clusters based on genetic virulence factors (113). Currently, the microarray method is being evaluated for molecular serotyping of GBS serotypes, aiming to deploy the assay for assessing carriage, including colonization with multiple serotypes as an endpoint in GBS vaccine studies.

1.12 Population Structure of GBS

Studying the sequence type (ST) diversity of colonizing GBS isolates is crucial as colonizing STs are reservoirs of virulence factors and serotype diversity. Furthermore, studies on capsular gene sequences diversity are essential as genetic alterations within the capsular genes could impact the effectiveness of GBS capsular polysaccharide-protein conjugate vaccines (114). The capsular

polysaccharides-based typing methods, such as by PCR, have limited resolution due to the minor representation of capsular polysaccharides (0.5% to 1.5%) within bacterial chromosomal DNA (110). Enhanced resolution is, therefore, achieved by typing of multiple loci using different methods such as multi-locus sequence typing (MLST), pulse-field gel electrophoresis (PFGE), and restriction digestion patterns (RDP).

The MLST technique is widely used in GBS epidemiological studies to compare the genetic relatedness of GBS strains based on standardized protocols. The MLST method is based on analysis of sequences of the 7 highly conserved DNA fragments (459 to 519 bp), which includes the phenylalanyl tRNA synthetase (PheS), alcohol dehydrogenase (adhP), serine dehydratase (sdhA), glucose kinase (glcK), transketolase (tkt), amino acid transporter (atr), and glutamine synthetase (glnA) (115). Sequence types (STs) are defined by the allele combination of the house keeping genes in each isolate (116, 117). There are currently 2099 STs in the MLST database (117). Globally, the most common STs in maternal colonizing GBS isolates are ST1, ST19 and ST23, which constitute over 80% of colonizing isolate STs (115, 118). Studies from North America, New Zealand, and the United Kingdom reported that ST23, ST19 and ST1 have poor invasive disease potential, are more adapt to colonizing the vaginal mucosa and frequently identified in asymptomatic pregnant women (13, 114, 118, 119). In contrast to colonizing STs, ST17 is a hyper-virulent clone that consists exclusively of serotype III and commonly causes meningitis in infant invasive GBS disease (120).

Capsular serotypes can be associated with more than a single STs such as serotype Ia being associated with ST23 and ST24 (118). Furthermore, different capsular serotypes may share a common sequence type. The distribution of STs of GBS colonizing isolates in women varies geographically, with ST17 and ST23 being the most prevalent in Spain, Portugal; and Italy, whereas ST19 predominates in Canada and the USA (121). Furthermore, the frequency of STs can change over time, including in the same person, reflecting acquisition of a new ST or capsular switching due to the acquisition of genetic sequences from a different sequence type isolate (87, 122).

Sequence types are further divided into clonal complexes (CCs), with isolates that share at least five or six identical alleles of house-keeping genes clustering within the same clonal complex (CC). The major GBS clonal complexes associated with human colonization are CC1, CC10, CC17,

CC19, and CC23 (123). Overall, the most common CCs in maternal colonizing isolates are CC19 (22%, 95% CI: 21-24%), CC23 (19%, 95% CI: 17-21%), CC1 (17%, 95% CI: 16-19%) and CC17 (15%, 95% CI: 14-17%) (123). In a recent systematic review, which included MLST data from Europe (45%; 35/78), sub-Saharan Africa (9%; 7/78) and one study each from northern Africa and southern Asia, the most common invasive disease causing CCs in infants were CC17 (42%, 95% CI: 40-44%), CC19 (18%, 95% CI: 16-20%), and CC23 (15%, 95% CI: 13-16%) (51). The lineage CC17 is a hypervirulent clone, that accounts for EOD or LOD cases ranging from 3.7% to 29% in Asia and up to 94% in Sub-Saharan region (Table 1.2) (13, 124, 125). The increased invasiveness of the CC17 is attributable to the presence of virulence genes, including the hypervirulent (*hvgA*) gene, which codes for a cell wall surface protein for invasion of human intestinal and blood brain barriers (46).

The MLST has the limitation of providing limited genotypic information and inability to detect recombination events due it being based on targeting selected loci (126). In contrast, whole genome sequencing (WGS) could provide additional genetic information to STs, as well as delineate the prevalence antibiotic resistance genes and surface protein antigens. The use of WGS also enables analyses of genome level changes at genome level that could drive lineage expansion, such as recombination events in bacterial populations (127, 128). Single nucleotide polymorphisms (SNPs) identified through WGS analysis may provide insight into transmission patterns within geographic regions and phylogenetic relationship between different STs or CCs (129).

Table 1. 2: The distribution of GBS clonal complexes in infant invasive disease across the globe

Author/Year Reference	Region	Country	Number of isolates sequenced	CC1	CC10	CC12	CC17	CC19	CC23	Other CC
Manning, 2009 (48)	Europe	Canada	183	17%	0	10%	31%	23%	18%	1.0%
Martins, 2017 (130)		Portugal	218	10%	4.0%	0	50%	11%	23%	2.0%
Brzychczy-Wloch 2014 (131)		Poland	42	2.0%	10%	0	10%	17%	49%	12%
Meehan, 2014 (132)		Ireland	72	15%	0	8%	46%	7%	13%	11%
Andersen, 2022 (133)		Denmark	212	9.0%	8.0%	0	42%	17%	19%	5.0%
van der Mee-Marquet, 2018 (134)		France	120	9.0%	5.0%	0	47%	11%	26%	2.0%
Morozumi, 2015 (135)	Eastern Asia	Japan	150	10%	11%	0	29%	30%	20%	0
Wu, 2019 (136)		China	29	0	14%	66%	0	17%	3.0%	0
Ryu, 2014 (137)		South Korea	41	44.0%	32%	0	0	15.0%	9.0%	0
Sigaugue, 2018 (138)	Sub-Saharan	Mozambique	35	3.0%	0	0	94%	0	3%	0
Seale, 2016 (139)	Africa	Kenya	73	3.0%	7.0%	0	70%	5.0%	15%	0

1.13 Whole genome sequencing analysis

The WGS platforms, the Illumina GA, Roche 454, and ABI Solid technology (140), generate raw reads that are cleaned and subsequently processed by mapping against a closely related reference genome or through de novo assembly. The short read mapping tools, including Bowtie2 and BWA, utilize Burrows-Wheeler Transformation (BWT) compression techniques to find optimal matches within an acceptable computation time. Portions of the genome may not be reliably called due to regions with ambiguous consensus nucleotides, often present in the repetitive regions (141, 142).

De novo genome sequence assembly involves reconstructing the genome from short-read sequences without relying on the reference genomes. The process utilizes the Bruijn graph, which employs read-depth information to construct scaffolds and contigs, albeit with longer computational duration (143). De novo assembly enables the detection of insertions and deletions and the identification of novel genetic variants within genome sequences. In most studies, mapping and de novo assembly are both applied in the whole genome sequence analysis workflow to capture the diversity of organisms within a population. The identification of variants between reference and sequenced genome, known as variant calling, involves the detection of single nucleotide variants (SNVs), structural variations, and indels (short insertions and deletions of less than 50 bases) (144).

The identified variants are annotated to understand the biological and functional implications, and to predict the potential effect of mutations such as disease causation. Variants predominantly fall into two categories: synonymous (silent) or non-synonymous that leads to amino acid changes. Nonsense variants can also arise, changing a protein codon into a stop codon and forming a truncated gene product. If the nonsense variants are located within the stop codon region, it can lead to the elongation of a protein (145). Many studies focus on indels and non-synonymous single nucleotide polymorphisms within protein coding-regions, which drives bacterial evolution and lineage expansion (146).

1.14 Recombination

Recombination within bacterial genomes shapes bacterial population evolution by promoting rapid dissemination of resistance elements and virulence determinants among clonal lineages (127). The DNA segments from a donor cell of the same species replace homologous segments in the recipient organism (120). While recombination primarily occurs among bacterial strains that are closely

related, it also occurs between distantly related bacterial species. The process is facilitated by unidirectional horizontal gene transfer mechanisms, such as transformation, transduction, and conjugation (147). Homologous recombination is a common phenomenon in GBS genomes, involving large portions of the genome within clonal lineages. Brochet *et al.* reported that *Streptococcus agalactiae* clones exchange chromosomal regions through “cis and trans-mobilization by conjugative elements”, with chromosomal deletions in GBS genomes ranging from 181 base pairs to 49 kilobase pairs (148, 149). Notably, recombination is common within the capsular gene locus, within the *cpsH* gene responsible for conferring serotype specificity for GBS Ia and III (148). Extensive recombination drives clonal diversification and the emergence of less common GBS serotypes VI, VII, and VIII (87).

1.15 Antimicrobials in GBS isolates

Antibiotic resistance is a significant public health threat, exacerbated by the excessive and uncontrolled use of antibiotics, resulting in diminished effectiveness of the first or second line of antibiotics (150). The current recommended prevention strategy for GBS in pregnant women colonized or presenting with GBS risk factors is the IAP. The first choice antibiotic for IAP therapy is penicillin, while lincosamides or macrolides (such as clindamycin and erythromycin) are recommended in cases of penicillin allergies (151). Although penicillin remains effective against GBS, resistance to second-line antibiotics (clindamycin and erythromycin) is increasing among GBS isolates (152).

1.15.1 Penicillin

Penicillins are grouped as either natural or synthetic antibacterial agents, derived primarily from fungi. The three common structural features shared by the penicillin antibiotics include: a side chain, and a thiazolidine ring coupled to a beta-lactam ring (153). The monobactams, carbapenems and cephalosporins (beta-lactam antibiotics), also harbor the typical beta-lactam ring element (154). Penicillins' mode of action involves preventing the formation of bacterial peptidoglycans in penicillin-binding proteins (PBPs), subsequently resulting in cell death due to an imbalance in cell wall maintenance (155). Alterations within PBPs amino acid, also result in reduced sensitivity in penicillin antibiotics (156). Group B *streptococcus* isolates exhibiting reduced penicillin susceptibility were reported in Japan, ranging from 2.3% to 14.7% (157). In contrast, GBS isolates from infant invasive disease in the United States of America, United Kingdom and Ireland were all susceptible to penicillin (15).

1.15.2 Macrolide and lincosamides antibiotics

Macrolides and lincosamides are utilized to treat infections caused by gram-positive bacteria, including *Streptococci*. The antimicrobial mode of action is based on binding to the 50S subunit of the bacterial ribosome, which leads to dissociation of peptidyl-tRNA and prevents protein synthesis (158). Reduced susceptibility to macrolides arises due to several mechanisms, including the acquisition of *erm* (erythromycin ribosome methylation) genes. The *erm* gene also confers resistance to macrolides, lincosamides and streptogramin B (MLSB) by altering the antibiotic binding site on the 23S rRNA, which can be either inducible (iMLSB) or constitutive (cMLSB) (159).

Active efflux of antibiotics through an ATP-dependent efflux pump, which is encoded by the *mef/msr* (*msrA*, *msrB* and *msrC*) genes constitutes the second mechanism of resistance to macrolides and lincosamides (160). Resistance is also attributed to the *Inu* (A, B, C D, E and F) genes, encoding a lincosamide nucleotidyl transferases enzyme that directly modify the antibiotic (161). Ribosomal protection is another form resistance, which involves the ABC-F proteins that are coded by the *Isa* and *mre* genes (162). The ABC-F proteins confer resistance by binding to the bacteria ribosome, thereby triggering the antibiotics to be released and resumption of protein synthesis.

The prevalence of erythromycin and clindamycin in GBS varies geographically (163), with China (64.9%) (164) and USA (44.8%) having highest rates of erythromycin resistance among invasive disease isolates (15). Higher prevalence of resistance to clindamycin was detected in GBS strains from Taiwan (65.9%), China (52.4%) (164) Algeria (43.2%) and Portugal (34%) (153, 165, 166), whereas lower prevalences were reported in Ghana (3.1%), Brazil (2%) and Argentina (1.8%) (167, 168).

1.15.3 Aminoglycosides and glycopeptides

Glycopeptides and aminoglycosides (i.e. vancomycin and gentamicin) are valuable for treating GBS infections, with vancomycin being the last option of antibiotics for IAP. Vancomycin inhibits the cross-linking transpeptidation and trans-glycosylation by binding to the D-Ala-D-Ala termini of peptidoglycan precursors (169). The aminoglycoside inactivating enzyme 6'-acetyltransferase-2'' phosphotransferase *aac* (6')-aph (2''), coded by the *aacA-aphD* gene confers resistance to gentamycin antibiotic (170). Vancomycin resistance arises through the synthesis of peptidoglycan

precursors with altered residues or modifications in the glycopeptide target site (171). Based on the altered amino acid, vancomycin resistance phenotypes include the Van A, Van B, Van D, and Van G (172). Overall, gentamycin resistance is low GBS, with prevalence being 8.8% in France (153) and 4.1% in GBS isolates in the United Kingdom (160).

1.15.4 Tetracyclines

Tetracyclines were extensively used against bacterial infections, and exert antibacterial effects by halting the elongation phase of bacterial protein synthesis (173). Mechanisms of resistance to tetracyclines in bacteria involves enzymatic inactivators, ribosomal protection proteins, and efflux proteins (174). In GBS, tetracycline resistance is predominantly caused by ribosomal protection proteins (*TetM* and *TetO*), and efflux proteins (*TetK* and *TetL*) encoded by the *tetM*, *tetO*, *tetK*, and *tetL*, respectively. The acquisition of *tetM* and *tetO* tetracyclines resistance genes by hypervirulent CC17 isolates, has led to the expansion of the GBS clones (123). Due to the widespread use of tetracycline antibiotics in the 1940s, prevalence of tetracycline resistance in GBS isolates is high, ranging from 80% to 100% in invasive isolates (118, 156).

There is limited information on the prevalence of antibiotic resistance genes in GBS colonizing isolates from low and middle-income countries.

1.16 Vaginal microbiome

The vagina hosts different organisms, including bacteria, viruses, and fungi, which generally maintain a synergistic relationship in their human hosts (175). Vaginal organisms provide protection against opportunistic infections through several mechanisms, including producing lactic acid that helps to maintain the acidic vaginal pH below 4.5 and inhibit the growth of potential pathogens. Metabolites secreted by *Lactobacillus* species inhibit bacterial and viral infections, such as herpes simplex virus (HSV) type 2 (176, 177). Vaginal organisms such as *Lactobacillus* species, that exist as commensals in the vaginal microbiome, prevent potential pathogens from attaching to the vaginal epithelium through competitive exclusion (Fig 1.6) (178). When the vaginal microbiome is disrupted, potential pathogenic microorganisms replace *Lactobacillus* species, the integrity of the vaginal epithelium is weakened and vaginal dysbiosis may ensue. Vaginal colonization by potential pathogenic organisms could also lead to in utero infection of the fetus or premature delivery as the pathogenic vaginal organisms invade into the amniotic cavity.

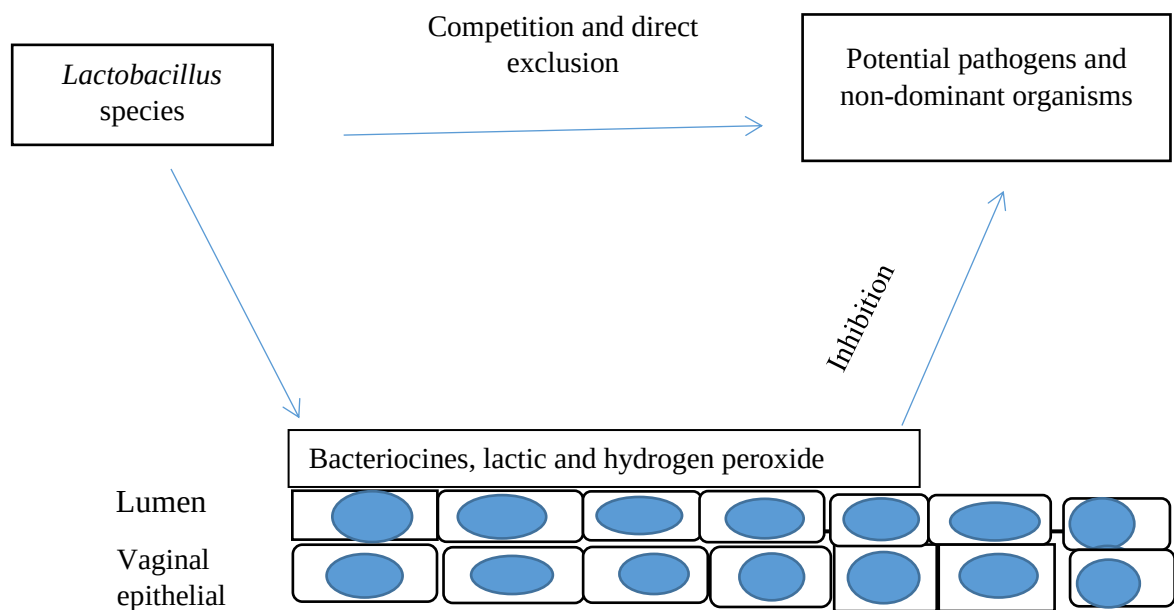


Figure 1.6: The role of *Lactobacillus* species in preventing colonization by potential pathogens on the vaginal epithelium

1.17 The vaginal microbiome community state types

A study by Ravel and colleagues, grouped the vaginal microbiome into five major community state types (CST I-V) based on the relative abundance of the most common vaginal bacteria (179). Four CSTs, i.e. CST I, CST II, CST III and CST V, are predominantly colonized by *Lactobacillus crispatus*, *L. gasseri*, *L. inners* and *L. jensenii*, respectively. In contrast, CST IV has high proportions of anaerobic bacteria (including *Atopobium*, *Corynebacterium*, and *Prevotella* species) and the absence or low proportion of *Lactobacillus* species (180). The CST IV is further stratified into either CST IVA (a mixed group of *Lactobacillus* species and anaerobic bacteria) or CST IVB which is composed of anaerobic bacteria with absence of *Lactobacillus* species.

A vaginal microbiome with *Lactobacillus* predominance (CSTI, CST II, CST III, and CST V) is healthy and protective. *Lactobacillus* species maintain vaginal health by secreting hydrogen peroxide, antibacterial bacteriocins, and lactic acids, which create an inhospitable environment for

potential pathogens (181, 182). The *Lactobacillus* species vary in their stability, ability to resist hormonal changes and menstruation, and to protect against vaginal colonization by anaerobic bacteria linked to bacterial vaginosis (BV) (183). *Lactobacillus crispatus* dominated vaginal microbiome has greater stability and associated with a lower risk of vaginal poly-microbial overgrowth compared to other *Lactobacillus* species (184). *Lactobacillus inners* produces less D-lactic acid compared with *L. crispatus*, which maintains a low pH level in the vaginal environment (185). Consequently, *Lactobacillus inners* is less stable and less protective against vaginal infections among *Lactobacillus* species (182), enabling colonization by other vaginal organisms.

A vaginal microbiome rich in anaerobic bacteria (CST-IV), but lacking *Lactobacillus* species is deemed unhealthy, and has been linked with poly-microbial overgrowth and a high vaginal pH (above 4.5) which favor colonization by potential pathogenic organisms (Figure 1.7). Overgrowth of anaerobic bacteria disrupts vaginal immune homeostasis and promotes rupturing of the epithelial barrier, leading to increased colonization by pathogenic organisms, including sexually transmitted pathogens such as human immunodeficiency virus (HIV) and human papilloma virus (HPV) (186-188).

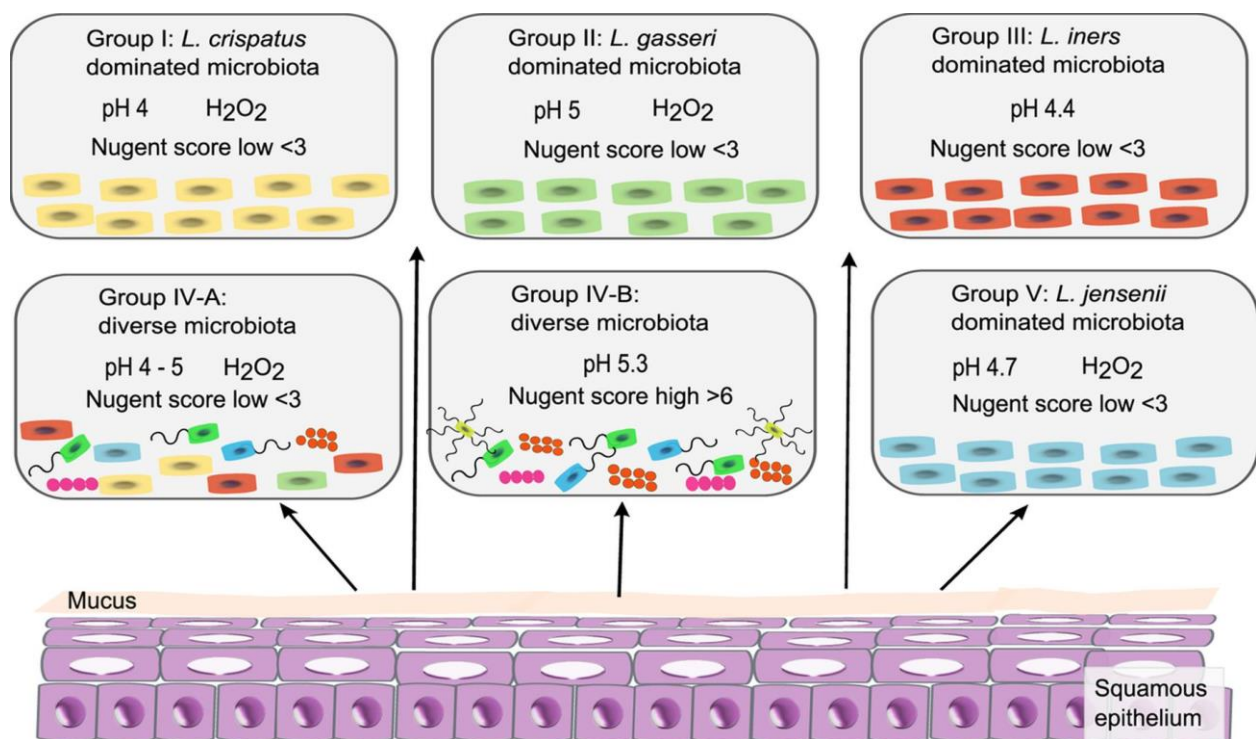


Figure 1.7: Characteristics of the vaginal microbiome community state types and groups (from Petrova and colleagues, 2013) (189).

The vaginal microbiome dominated by *Lactobacillus* species (Group I, II, III and V) are less diverse whilst the vaginal microbiome dominated by non-*Lactobacillus* species are more diverse. Nugent score is a gram stain diagnostic test that defines the vaginal microbiome as normal vaginal flora (≤ 3), intermediate vaginal flora (4-6) or bacterial vaginosis vaginal flora (> 6).

There is variability in the vaginal microbiome composition by geographical region and among ethnic groups. The vaginal microbiota of Asians and Caucasians is dominated by *Lactobacillus* (80% to 90%), compared with Hispanic and African women (45% to 60%) (179, 190). Conversely, the CST-IV is highly prevalent in African women (40%) compared to Asians and Caucasians (10% to 20%) (Figure 1.8) (179). The reasons for variability in vaginal microbiome composition are unclear, but could be due to differences in diet, environmental factors, sexual practices, immune status and genetic predisposition (182).

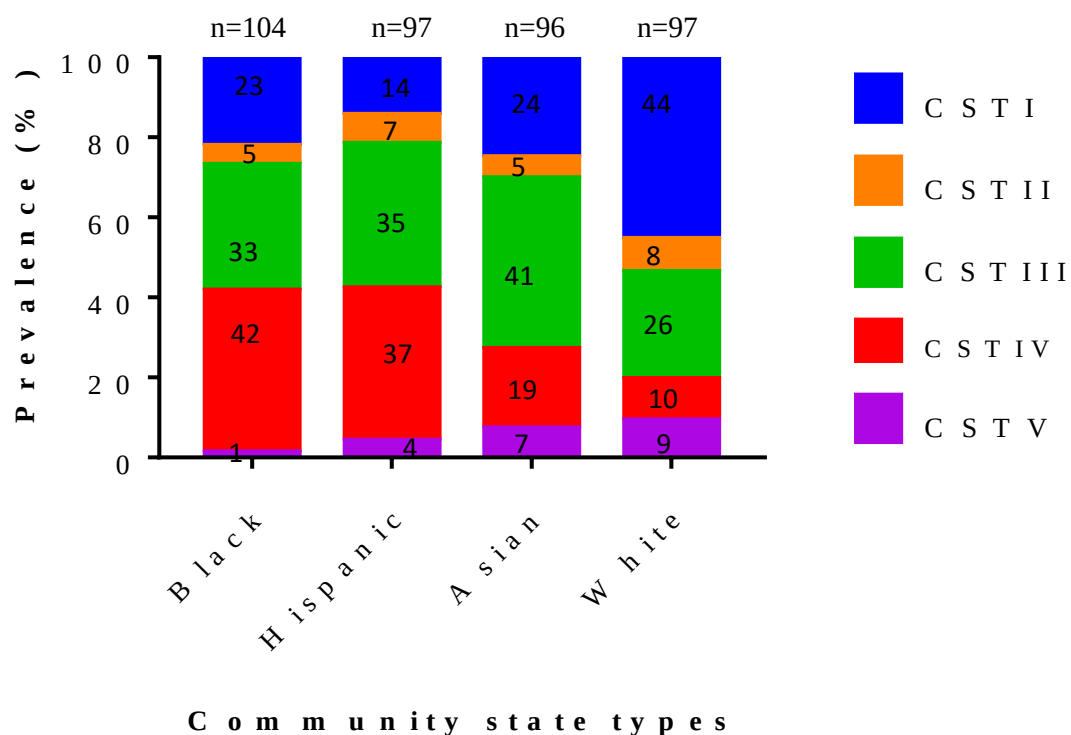


Figure 1. 8: Vaginal microbiome community state groups within different ethnic groups of women, data adapted from Ravell *et al*, 2011 (179). The CST IV (diverse vaginal microbiome) was more prevalent in women of Black and African ancestry. Women of European ancestry (white) had *Lactobacillus* predominate in the vaginal microbiome.

In a recent study analyzing the vaginal microbiome of 76 European and 246 pregnant African American women (182), African American women had a relatively large diversity of vaginal organisms that consisted of *Mycoplasma hominis*, *Lactobacillus inners*, Bacterial vaginosis associated bacterium (BVAB) 1, BVAB2, *Sneathia* spp, *Megasphaera* spp, *Prevotella amnii*, and *Gardnerella vaginalis*. In contrast, European women had low diverse vaginal microbiome. Additionally, a longitudinal revealed that in African women, the vaginal microbiome is less likely to transition from a dysbiosis-associated microbiome to *Lactobacillus* dominated microbiome, but prone to transitioning from *Lactobacillus* to dysbiosis associated microbiome (191).

1.18 Major vaginal infections

1.18.1 Bacterial vaginosis (BV)

Bacterial vaginosis (BV) is portrayed by an overgrowth of anaerobic bacteria and diagnosed by high Nugent scores (>6) (192, 193). The BV is associated with poor pregnancy outcomes and an increased susceptibility to acquisition of sexually transmitted infections. Bacterial vaginosis is negatively associated with GBS colonization, while being positively associated with *Candida albicans* and *Escherichia coli* carriage (194). Nevertheless, a study from South Africa did not identify any association between BV and vaginal colonization by GBS (195). The prevalence of BV associated bacteria is reportedly high in African women (182), including 34% to 58% in South Africa (196).

1.18.2 Sexually transmitted infections (STIs)

Sexually transmitted infections (STIs) are prevalent globally and negatively impact the health outcomes of pregnant women and their newborns. According to 2020 estimates by the World Health Organization (WHO), there were approximately 374 million new infections of four curable STIs (*syphilis*, *gonorrhea*, *trichomoniasis* and *chlamydia*), while 417 million people were infected with herpes simplex virus type 2 (HSV-2) globally (197). The relationship between GBS colonization and STIs is not well researched. A vaginal microbiome rich in *Lactobacillus* species has, however, been shown to reduce HPV and *Chlamydia trachomatis* infections (186). Conversely, the presence of BV associated bacteria with paucity of *Lactobacillus* species in the vaginal microbiome is correlated with increased occurrence of genital tract shedding of HSV-2 (198).

1.19 Vaginal microbiome in pregnant women

During pregnancy, the vaginal microbiome narrows to become low diverse, with *Lactobacillus* species dominating and associated with low proportions of anaerobic bacteria (182, 199). A low diverse vaginal is accompanied by high stability during pregnancy, and even when changes occur, the vaginal microbiome transitions from and to another *Lactobacillus* dominated CST (182). In African women, *Lactobacillus* species are more abundant during pregnancy than in non-pregnant women (191).

The presence of polymicrobial communities such as *Atopobium vaginae*, *Mycoplasma* species, *Gardnerella* species, *Prevotella*, and *Sneathia* species during pregnancy, may promote rupture of the membranes and premature delivery (182). Other vaginal organisms such as GBS, *Staphylococcus* species, and *Escherichia coli*, frequently present in the vaginal flora of pregnant women, can also infect the fetus through ascending infection route (194, 200). Collectively, GBS, and *Staphylococcus* species are the most common cause of infantile and neonatal sepsis (201, 202).

1.20 The relationship between the vaginal microbiome and Group B *Streptococcus* colonization

The prevalence of GBS colonization varies in different geographic regions and among different ethnic groups (65). Regional heterogeneity in the prevalence of GBS may also be attributed to biological factors such as the vaginal microbiome. Mechanisms by which the vaginal microbiome influences colonization by potential pathogenic organisms i.e. GBS, include reduced epithelial barrier function in the lower female genital tract, amplification of specific inflammatory cytokines, and lowered levels of antimicrobial factors (203).

There are, however, conflicting data on the relationship between GBS colonization and the vaginal microbiome. Some studies reported a positive association between vaginal GBS colonization and reduced *Lactobacillus* species (194, 204, 205), whereas no association between GBS colonization and the vaginal microbiome has been identified in other studies (206, 207). Most previous studies have been conducted in high-income countries (Table 1.3). Consequently, the generalizability of the results to low and middle-income countries remains uncertain.

In vitro studies have shown that *Lactobacillus* species have an antimicrobial effect on GBS growth and inhibit GBS adherence to vaginal epithelial cells (208, 209). Studies in women from China and USA suggested that GBS exists synergistically with vaginal bacteria such as *Gardnerella vaginalis*

and *Prevotella* species, possibly due to metabolic dependencies between species (200, 210). In non-pregnant women, colonization with GBS was positively associated with *Prevotella bivia* colonization, suggestive of synergistic relationship between the two organisms (200). Previous studies from Japan and South Africa found a positive association between GBS and *Candida* species colonization in women (194, 204), possibly due to GBS tropism for *Candida* hyphae filaments as indicated by in vitro studies (211). Similarly, there is an inverse association between GBS and other beta-haemolytic *Streptococcus* strains and *Streptococcus pneumoniae* due to having similar metabolic pathways or competition for nutrition (212).

As vaccines against GBS are currently under development, their introduction could lead to a niche in the vaginal microbiome that could be replaced by other vaginal organisms. Also, interventions aimed at changing the vaginal microbiome, could inadvertently influence the risk of GBS vaginal colonization acquisition.

Table 1. 3: Studies on the vaginal microbiome associated with colonization by Group B *Streptococcus*.

Author/Year (Reference)	Study population	Country	Study Technique	Key Findings
Antonio, 1999 (213)	301 sexually active women	USA	Culture and Staining	No significant differences in detection of GBS between women with <i>Lactobacillus</i> species isolated (27/154; 18%) and women not colonized by <i>Lactobacillus</i> species (36/147; 24%, p=0.15).
Kubota, 2002 (204)	4025 pregnant women at 22-36 weeks gestation age. 408 GBS positive and 3617 GBS negative women	Japan	Culture and gram staining	The prevalence of coagulase-negative <i>Staphylococcus</i> (11.5% vs 18.3%; p=0.0003), <i>Prevotella</i> species (1% vs 2.6%; p=0.0003), <i>Lactobacillus</i> (34.4% vs 53.5%; p<0.0001), <i>Enterococcus faecalis</i> (2.5% vs 3.8%; p=0.03), microaerophilic <i>Streptococcus</i> species (1.7% vs 5.7%; p<0.0001), <i>Candida</i> species (6% vs 8.8%; p=0.001) and <i>Peptostreptococcus</i> species (0.4% vs 1.2%; p=0.02) was significantly lower in the GBS colonized compared to women not colonized by GBS.
Altöparlak, 2004 (205)	150 pregnant women, 17- 45 years. 41 women were GBS positive and 109 were negative for GBS	Turkey	Culture and gram staining	The prevalence of BV was 12.2% (5/41) in GBS colonized and 9.2% (10/109) in GBS negative group and GBS colonization was not associated with risk of BV. The percentage of <i>Lactobacillus</i> species was significantly lower in GBS colonized (21.3%) compared with uncolonized (47.6%) women (p<0.01). Presence of <i>Candida</i> species was positively

				associated with GBS colonization in pregnant women (14.9% vs 2.9%; p <0.001).
Ronnqvist, 2006 (214)	191 non-pregnant women, over 18 years. 79 GBS positive and 103 GBS negative	Sweden	Culture and sequencing	The density of vaginal <i>Lactobacilli</i> was lower in GBS colonized women (<3 colony forming units per millimeter) compared with those not colonized (>6 colony forming units per millimeter; p=0.036).
Cools, 2016 (194)	424 women, including 60 pregnant women. 69 GBS positive (27 Kenya, 6 Rwanda and 36 South Africa) 355 GBS negative	South Africa (199 women) Kenya (165 women) and Rwanda (60 women)	Culture, Nugent scoring and qPCR	The presence of BV (nugent score 7-10) was negatively associated with GBS carriage (19% in GBS positive and 35% in GBS non-colonized women; OR 0.43, 95% CI: 0.21-0.88; p=0.022). The prevalence of <i>Candida albicans</i> was significantly higher in GBS positive (25%) compared with GBS negative (8%) women (OR 3.25, 95% CI: 1.5-7.06; p=0.003). GBS carriage was positively associated with the presence of <i>Escherichia coli</i> (42% in GBS colonized vs. 25% in GBS non-colonized women; OR 2.01 95% CI: 1.10-3.80).
Rick, 2017 (207)	94 pregnant women (≥35 weeks gestation). 8 GBS positive women and 84	Guatemala	Culture, Latex agglutination, 16S Sequencing	The overall vaginal microbiome composition was not associated with GBS. A lower of <i>Gardnerella</i> was detected among GBS colonized (2.2%) compared with non-colonized (11%) women (p<0.05). In contrast, the presence of <i>lactobacilli</i> was higher among GBS colonized women

	GBS negative women			compared with GBS uncolonized women (89.9% vs 51%; $p<0.05$). The relative abundance of <i>Corynebacterium</i> was higher in GBS colonized (3.6%) than in GBS noncolonized (0.2%) women ($p<0.05$).
Rosen, 2017 (200)	428 non-pregnant women aged 18–55 years. 92 women positive for GBS. 336 women negative for GBS.	USA	Amplicon Sequencing	Taxa with positive correlation with GBS colonization include <i>Prevotella bivia</i> ($r=0.4$; $p=0.02$), <i>Veillonella</i> species ($r=0.7$; $p=0.04$). Taxa negatively correlated with GBS colonization include BVAB1 ($r=0.7$; $p=0.04$), <i>Prevotella</i> genogroup 3 ($r=0.1$; $p=0.02$), <i>Megasphaera</i> species type 1 ($r=, 0.5$; $p<0.05$).

USA: United States of America, vs: versus

1.21 Molecular analysis of the vaginal microbiome

Before the advent of sequencing methods, the traditional methods for studying the vaginal microbiome were the culture-based methods. Vaginal swabs are examined under the microscope following gram-staining and heat treatment (215). Culture methods, however, have several limitations, including being labour intensive and the lack of standard culture media for propagating vaginal organisms. In contrast, the deployment of molecular methods is less laborious and enables evaluation of a larger spectrum of organisms, including having led to the identification of previously unrecognized vaginal organisms such as *Atopobium vaginae*, *Sneathia* species and *Lactobacillus inners* (216). The most widely employed molecular method for studying vaginal organisms are the next-generation sequencing methods, which allows for detecting different organisms.

Next-generation sequencing methods identify vaginal organisms through sequencing the 16S rRNA gene, which is highly conserved within the same species and can distinguish between different taxa (217). The limitations of sequencing methods include being a semi-quantitative method and less sensitive in detecting low density organisms that could influence the vaginal microbiome. Furthermore, resolution of the next generation sequencing method is determined by the choice of primers used for sequencing analysis, as the V1-V9 regions provide different resolution of the vaginal organism profile (218).

Other molecular methods include PCR based methods that allow quantification of vaginal organisms. Several studies mainly from high income countries i.e. USA, UK and Italy used broad-spectrum PCR assays for characterizing the vaginal microbiome community (219). The broad-range PCR, however, is less sensitive in detecting species that have low cell count within the microbial community. The limitations of the conventional PCR method include inability to detect multiple organisms in a single reaction which warranted the development of multiplex qPCR methods such as the open array assays, which can detect multiple organisms simultaneously in a single experimental run.

The open array technology is based on a metal plate that is photo-lithographically patterned to form a rectilinear array of 3072 through-holes, arranged in 48 sub-arrays of 18 to 64 through-holes each (220). Each through-hole utilizes 33 nl of the PCR reaction mixture and loaded with an individual TaqMan assay. The assays are labelled with probes to profile multiple target pathogens from the

same sample DNA or RNA template and for simultaneous analysis of several samples in a single open array reaction.

The combined positive control with genetic inserts of multiple targets is included in each reaction to monitor reagent quality, amplification efficiency, and homogeneity over time in different reaction plates (221). The advantage of the open array is that gene targets from non-culturable organisms are included in the design and detected by the assay. Open array technique has low limit of detection as it allows the determination of the quantities of gene targets, which corresponds to the absolute concentrations of the target organisms in study samples.

1.22 Justification of the study

The burden of GBS related disease is highest in low and middle-income countries, particularly in African countries, where GBS is a leading cause of mortality in neonates. The strongest predictor of early-onset GBS invasive disease in newborns' is recto-vaginal GBS colonization in the mothers. Currently, there is no effective method to prevent maternal GBS colonization. The most promising approach to prevent maternal colonization and invasive GBS disease in neonates is through maternal GBS vaccination. GBS polysaccharide-protein conjugate vaccines that are under development induce serotype-specific immunity, targeting serotype-specific epitopes, whilst protein-based vaccines under development target specific surface protein antigens expressed by a subset of GBS serotypes.

Sequencing of GBS genomes to explore the genetic diversity of GBS sequence types and clonal complexes in different countries, could assist in understanding the heterogeneity in disease across settings, as well as evaluate the future potential impact of vaccines thereon. Genetic alterations in the gene region targeted by vaccine could result in phenotypically distinct isolates and vaccine escape recombinants that reduce vaccine effectiveness. In addition, sequencing of GBS genomes is essential to determine the prevalence of GBS surface proteins, virulence factors, and antimicrobial resistance within GBS serotypes and clonal complexes.

There is scarcity of information on the prevalence of concurrent colonization by multiple GBS serotypes, which would be important for predicting the potential impact of multivalent GBS vaccines. Detection of multiple GBS serotypes is not routinely performed, so colonization with more than one serotype is largely undefined in many settings.

There is a paucity of studies on the vaginal microbiome from African countries, and whether GBS colonization is associated with specific vaginal organisms. The vaginal microbiome in the peripartum period was investigated in women from three African countries, including investigation of association of the microbiome and specific organisms in relation to GBS colonization.

1.23 Aims of the study

The focus of the research included: i. Determine the genetic composition of colonizing GBS sequence types (STs) and clonal complexes (CCs) in pregnant women and their newborns from six sub-Saharan African and two Southeast Asian low and middle-income countries (LMICs); ii. Determine the prevalence of concurrent colonization by more than one GBS serotypes using microarray; iii. Investigate the association between vaginal GBS colonization and the vaginal microbiome in the pregnant women from LMICs.

Study objectives

- i. To determine the diversity of sequence types and clonal complexes of GBS vaginal colonizing isolates in pregnant women and infants from across eight low and middle-income countries (Bangladesh, South Africa, Ethiopia, Kenya, Nigeria, Mali, Mozambique, and India).
- ii. To determine the frequency of multiple GBS serotypes in women at the time of delivery from six low and middle-income countries (Bangladesh, South Africa, Ethiopia, Kenya, Nigeria, and Mali).
- iii. To determine the association between vaginal GBS colonization and the vaginal microbiome at the time of delivery in pregnant women from South Africa, Kenya, and Ethiopia.

Chapter 2.0 Methods

2.1 Study population

The study was a retrospective cross-sectionally designed study in which vaginal swabs and GBS colonizing isolates from expectant mothers and their respective newborns were analyzed. Samples were collected as part of a multicenter GBS prevalence study, which aimed to investigate the prevalence of recto-vaginal GBS colonization in women at the time of delivery and vertical transmission thereof to newborns (222). The participants were pregnant women aged 18 to 45 years who were HIV-1 seronegative and their newborns from across eight low and middle-income countries. The countries included Bangladesh, India, Ethiopia, Kenya, Mozambique, Nigeria, Mali, and South Africa. Samples were collected following informed consent from the women, including on behalf of their newborns, at the time of labour or delivery. The study enrolment took place between January 2016 to December 2018 (222). Standardized procedures for collecting study samples and demographic or clinical data were used at all the sites. The demographic and clinical data of participants were captured in a study database by trained research staff, with personal identifiers being anonymised.

The inclusion criteria for enrollment of participants in the parent multicenter GBS prevalence study were:

- Pregnant women aged 18 to 45 years.
- Gestational age at delivery ≥ 34 weeks
- Documented to be HIV-uninfected prior to enrolment in the study.
- Able to understand and agree with the study procedure.
- Provide informed consent and for their newborns prior to enrollment in the study.

Exclusion criteria:

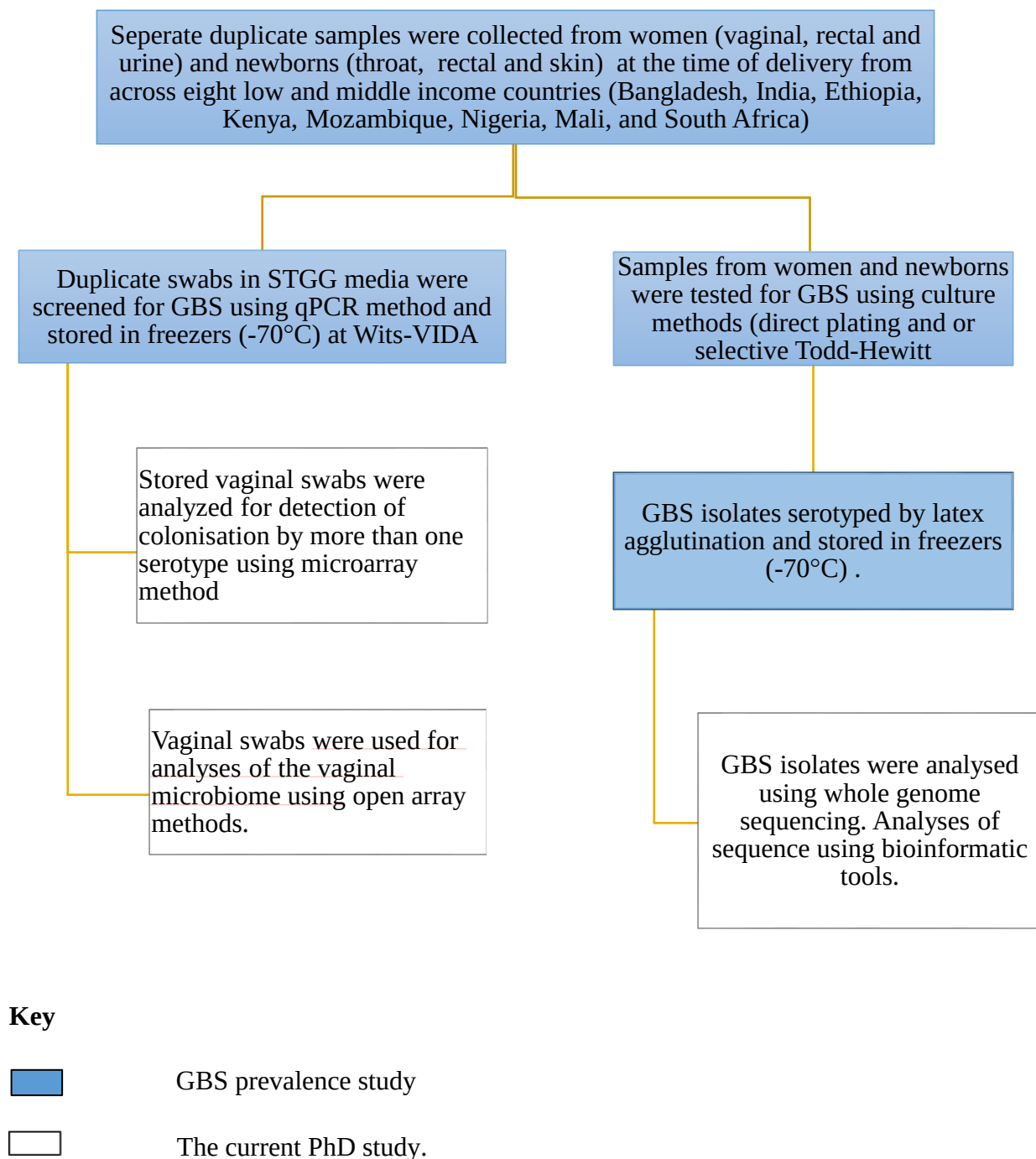
- Antibiotic treatment in the 2 weeks prior to delivery.
- Unable to obtain samples.
- Not consenting to study procedures.

- Any acute illness, symptomatic vaginal discharge, and any suspected condition in which vaginal swab collection was contradicted.

In brief, duplicate lower vaginal swabs, rectal swabs and urine samples were collected separately from expectant women at the time of delivery, prior to rupture of amniotic sac membranes. Separate throat, rectal, and skin surface (umbilicus, outer ear and axillary fold) swabs were also collected from newborns at birth, prior to cleansing or bathing. The samples were collected with rayon-tipped swabs (MW170, Medical wire, UK) and stored separately in Amies transport medium without charcoal, before being transported on dry ice to study site laboratories for GBS detection or isolation. At the site laboratory, one of the duplicate swabs was tested for GBS using two culture methods: direct plating on chromogenic media (CHROM agar StrepB; CA) and inoculation into selective Todd-Hewitt broth supplemented with colistin and nalidixic acid (Lim broth), then followed by subculturing onto the selective media (Selective Broth method, SB) (223).

The GBS-like colonies exhibiting purple-pinkish color were further tested for the presence of Christie Atkins Munch-Petersen factor (CAMP), catalase negative, absence of esculin hydrolyses and detection of antigen for group B *Streptococci* (106) (222). Confirmed GBS colonies were kept in skim milk-tryptone-glucose-glycerol (STGG) media at -70 °C and shipped on dry ice for serotyping to the Vaccines and Infectious Diseases Analytics Research Unit (Wits-VIDA) at the University of the Witwatersrand in South Africa. Serotyping was performing using latex agglutination assay with specific antisera against GBS types Ia, Ib and II to IX CPS antigens (Statens Serum Institute, SSI, Denmark) (224). Duplicate swabs in 1ml STGG media were also shipped on dry ice to the Wits-VIDA for further GBS testing using the PCR method and stored at -70 °C in freezers until further analysis (Figure 2.1).

Figure 2. 1: PhD Study design



2.2 Methods for whole genome sequencing analysis

2.2.1 Selection of study samples

Group B *streptococcus* isolates used in the whole genome sequencing study were collected from the parent study evaluating the prevalence of GBS colonization (222). The GBS isolates were selected based on the serotypes detected by latex agglutination or PCR. (Figure 2.2 and Figure 2.3). For each country, at least five pairs of mother-newborn dyad isolates of the same serotype were randomly selected using Stata software (v13.1). Among the mothers, vaginal isolates were prioritized over rectal isolates if both sites were colonized by GBS. Similarly, among newborns with colonization at multiple sites, the order of prioritization for the isolate selected for sequencing in order of hierarchy was from the throat, rectum and skin surface. In instances, at the country level, where the number of paired isolates of uncommon serotypes was fewer than five paired isolates, the number of isolates was supplemented by including unpaired isolates of that serotype or available paired isolates from the same country. For India, all the available mother-newborn dyad samples were sequenced.

In addition, at least three pairs of maternal urine samples and newborn isolates per each serotype were randomly selected from each of the study countries. In instances where the serotypes of paired maternal (rectal and vaginal) and newborns (throat, rectal or skin surface) were different, all the maternal and newborn isolates were selected for sequencing. The serotype distribution is based on whole genome capsular serotypes. Multiple isolates from the same person of differing serotypes, ST and CCs were included in the analysis.

Figure 2. 2: Flow diagram of the maternal GBS isolates selected for analyses by whole genome sequencing.

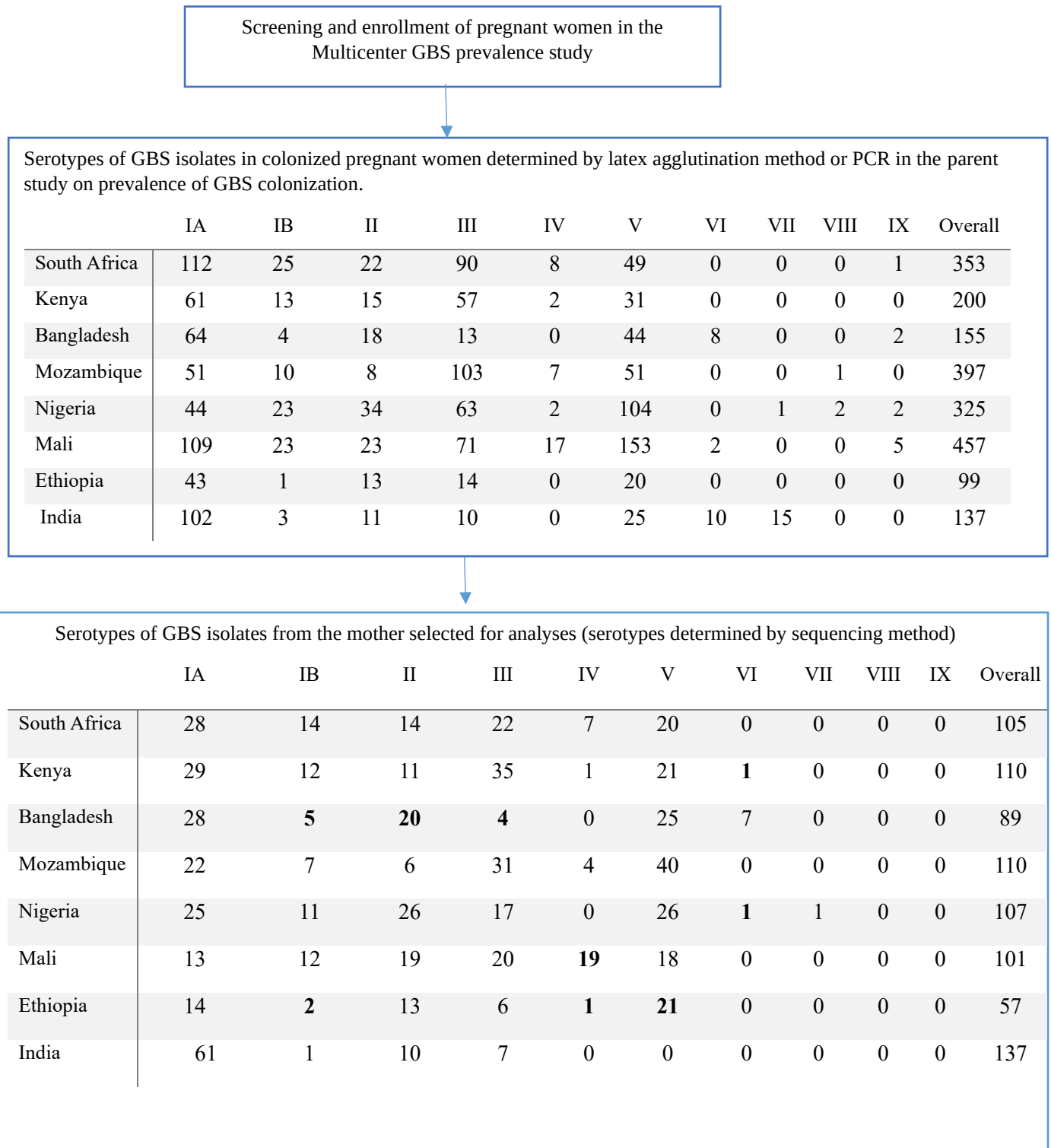
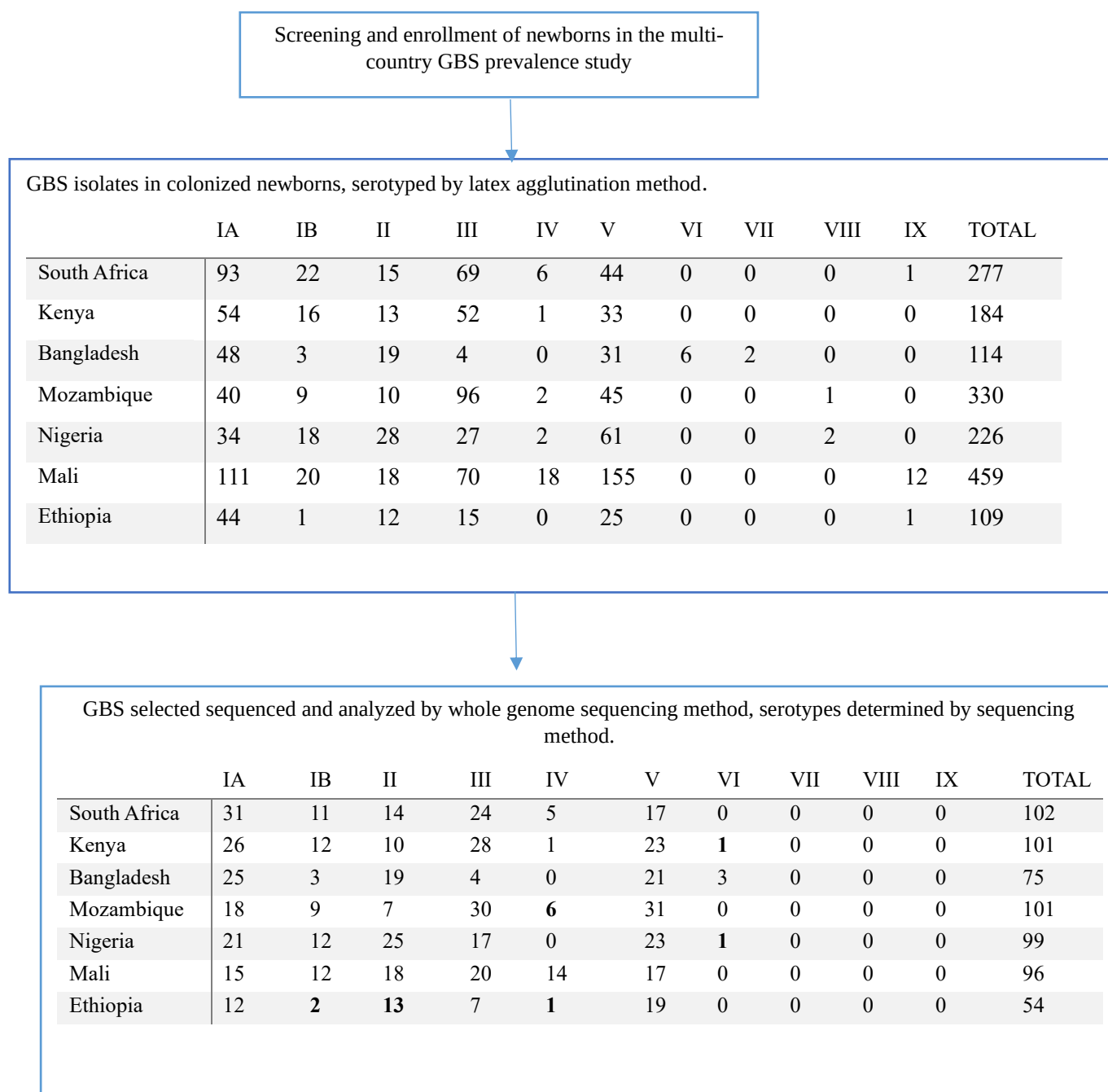


Figure 2. 3: Flow diagram of the GBS isolates selected from newborns for analyses by whole genome sequencing.



The numbers that in bold are due to mismatch between latex agglutination results and serotyping by whole genome sequencing.

2.2.2 Culturing and Genomic DNA Extraction

Stored GBS isolates were cultured on chromogenic media (Colorex StrepB agar), and individual colonies were sub-cultured on sheep blood agar overnight at 37°C and 5% carbon dioxide (CO₂). The harvested GBS colonies were subjected to lyses for 60 min at 37 °C using an enzymatic lysis buffer (200 µL) of lysozyme (40 mg/mL; Sigma-Aldrich, Dorset, UK) and mutanolysin (79 units/mL; Sigma-Aldrich). Genomic DNA extraction was performed using Qiagen manual kits (Qiagen, UK), followed by DNA quantification using a QubitTM fluorometer (Life Technologies) using the dsDNA broad-range kit.

2.2.3 Whole genome sequencing

Whole genome sequencing of genomic DNA was performed at the Wellcome Sanger Institute (Hinxton, United Kingdom) using the next-generation sequencing platform (Illumina HiSeq 2000; Illumina, USA). In brief, the genomic library was generated using Nextera XT V2 (Illumina, USA), with bead-based normalization was performed according to the manufacturer's protocol. Paired end short reads of 125 bp were generated using a HiSeq platform (Illumina, USA). Raw sequence reads were subjected to quality control procedures (225), including trimming low-quality reads with Trimmomatic v0.35 and discarding contigs with per base read quality <28. Reads of good quality were chosen based on criteria such as base quality per sequence (33-40), per base N content (NIL), phred scores (36-39), per base GC content (approximately 37%) and duplication level (<10%), Denovo genome assemblies were generated using Spades v3.9.0 with kmers of 43 to 127 bases (226). Statistics for genome assemblies, including total genome length ranging from 1.3 to 2.8 Mbp, number of contigs between 13 to 55, a 33-37% GC content and N50 ranging from 8.0 Kbp to 1.17 Mbp, were summarized using Quast v2.1. The PROKKA tool (v1.11) (227) and the *Streptococcus agalactiae* database were used to annotate the GBS genomes (228).

2.2.4 Typing of the capsular serotype, surface protein genes and sequence types

The GBS capsular serotypes Ia, Ib and II-IX were determined from the raw reads by employing previously published primers for silico PCR (101, 229). Sequence types were assigned based on the MLST scheme using the short-read sequence typing tool (SRST) (230) and the *Streptococcus agalactiae* database (117). Isolates were grouped to CCs using the multi-locus sequence data (117), core genome phylogeny from a reference-free approach (231), and cluster assignment from Fastbaps clustering (232, 233). Surface protein genes that are potential vaccine antigen targets, such as the *srr1* or *srr2* genes encoding serine-rich repeat proteins, pilus island family genes (PI1, PI2A, and PI2B), alpha protein family genes (*alpha*, *rib*, *alp2/3*, *alp1*), and the *hvgA* gene were

identified via in silico methods (234-236). The AMR genes were checked on sequenced isolates using SRST v.0.2 (230) and the Resfinder databases (237), complemented by sequences of GBS resistance determinants described previously (156). For the construction of a core genome phylogeny, a core genome alignment was generated from genome annotations using Panaroo v1.3 (231). The core SNPs identified with SNP sites from the alignments were used to construct a phylogenetic tree (maximum likelihood; ML) using IQ-TREE and viewed using the Interactive Tree of life (iTOL) (238, 239).

2.2.5 Recombination detection and phylogenetic analysis of GBS isolates

Recombination hotspots were identified using Gubbins and visualized on Phandango (240). Manual visualization of the *hvgA* gene sequences was performed using Artemis genome viewer and blast features on the NCBI website (241). A phylogenetic tree was built for each CC, serotype III and *hvgA* positive isolates based on mapping the GBS isolates to CC specific references: CC1, CC10, CC17, CC19 and CC23. Recombination regions were identified and removed using Gubbins (240). The phylogenetic trees were visualized using the Interactive Tree of life (iTOL) (242).

2.2.6 Statistical analysis

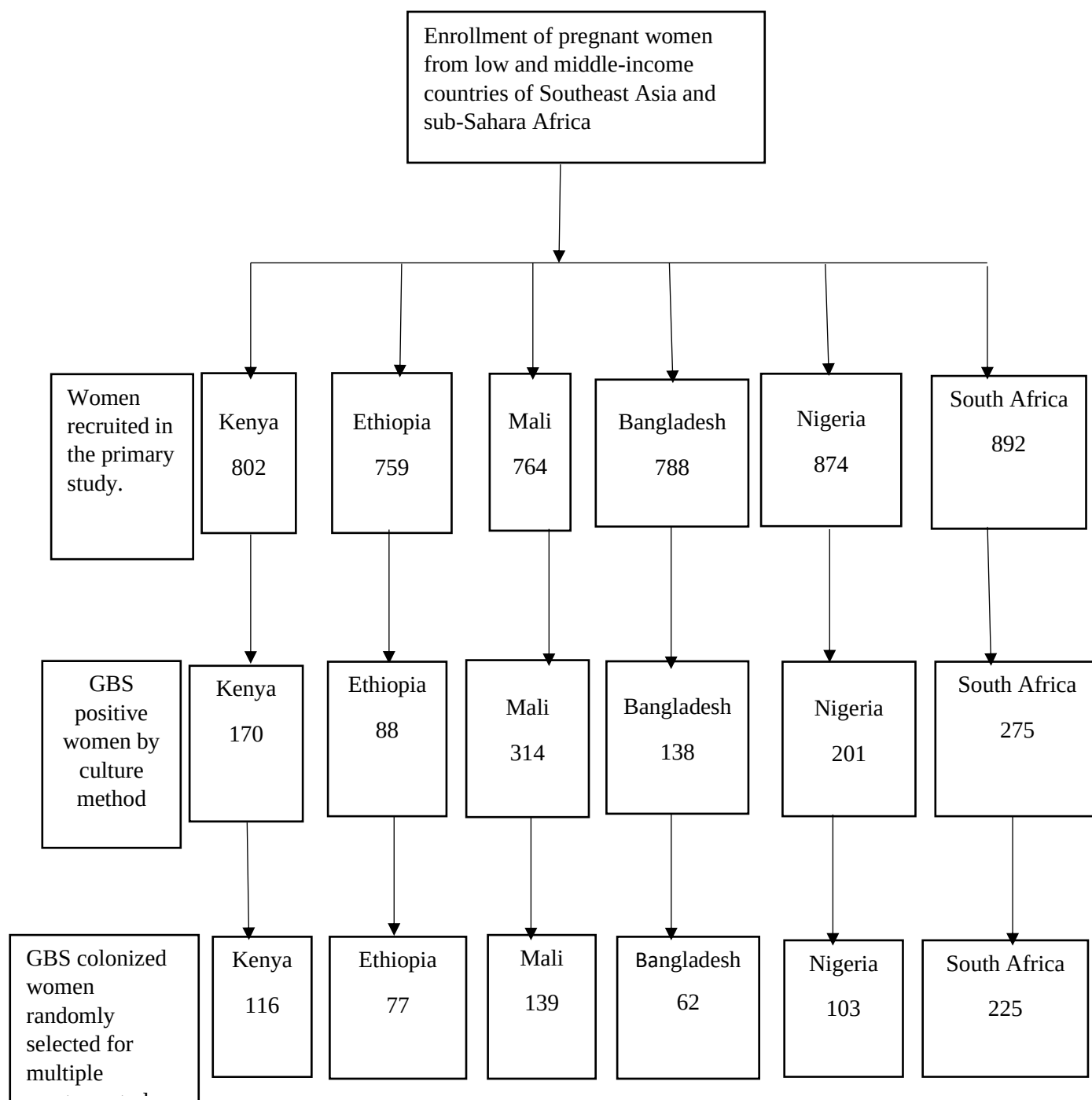
The prevalence of STs and CCs were reported as proportions (percentages) and binomial proportional confidence intervals (CIs) were generated using the Wilson score (243, 244). Concordance in ST or CC between mother and newborn dyads isolates was determined by dividing the number of pairs with the same ST or CCs by the total number of paired isolates with identical serotypes. Data summaries were computed using R (v4.2).

2.3 Methods for analysis of multiple serotypes

2.3.1 Selection of study samples

For the detection of colonization by more than one GBS serotypes, a total of 733 vaginal swabs in which GBS was identified by culture method were randomly selected from six study sites in South Africa, Kenya, Ethiopia, Bangladesh, Mali, and Nigeria (Figure 2.4). Assuming 12% prevalence of multiple serotype colonization in with GBS positive women, a sample size of at least 451 GBS-positive women allow the detection of at least 54 women with colonization by more than one serotype, with a precision of $\pm 3\%$ and 95% confidence interval. 12% is the average expected prevalence of multiple GBS serotypes in GBS colonized women from previous studies (6.6-16.7%) (83, 106)

Figure 2.4: Study cohort diagram. Consort of pregnant women selected for detection of colonization with more than one GBS serotype.



There were 6514 pregnant women enrolled in the GBS prevalence study from across nine low and middle-income countries, of which 1572 tested positive by culture method. A total of 733 GBS positive women from across 6 countries were selected for detection of colonization by more than serotype using by microarray method. A few sites were selected for analysis as the study was exploratory.

2.3.2 Culturing of vaginal samples and Genomic DNA extraction

Vaginal swabs stored in STGG were thawed and vortexed for two minutes to release the bacteria into the transport media. Aliquots of 110 µl were shipped on dry ice to BUGS Bioscience Ltd. (London, UK) for analysis by microarray methods. Thawed vaginal swabs were cultured on Brilliance™ agar plates and incubated overnight at 37 °C and 5% CO₂. The bacterial growth on the entire plate was collected and suspended in 1 mL of phosphate-buffered saline (PBS) (Sigma-Aldrich, Germany) and subsequently pelleted in a 1.5 mL microcentrifuge tube by centrifugation at 13,000 rpm for 5 minutes. The bacterial pellet was re-suspended in 180 µl lysis buffer, which consisted of 20 mg/mL lysozyme dissolved in 20 mM Tris-HCl: pH 8, 2 mM ethylenediaminetetraacetic acid (EDTA), and 1.2% Triton. This was followed by incubation at 37 °C for 1 hour. After adding 20 µl proteinase K (0.5 mg/mL) and 200 µl buffer AL, incubation was performed at 56 °C for 30 minutes. Subsequently, 4µL RNase A (100 mg/mL) was added, with incubation at room temperature for 5 minutes. The bacterial DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, MA, USA), according to manufacturer's protocol. Finally, the DNA concentration was determined using a NanoDrop (Labtech, UK).

2.3.3 DNA Fragmentation and labelling

DNA extracts were analyzed using microarray methods according to previously described protocols (245). Briefly, DNA samples were normalized, and heat fragmented in a thermal cycler (Eppendorf, Germany) at 95 °C for 60 minutes. The DNA labelling was performed by adding 1.2 µl of the cyanine-green (Cy3-ULS) and cyanine-red (Cy5-ULS) labelling master mix to the DNA samples in tubes 1-48 and 49-96, respectively. The sample labelling reactions were incubated in a thermal cycler at 85 °C for 30 minutes, with a final holding step at 4 °C for 3 minutes. The ULS-labelled DNA samples (10 µl) were combined by adding one Cy3- and one Cy5- labelled sample per well in the purification plate and centrifuged at 10 000 rpm to remove any unreacted fluorescent labelling dye.

2.3.4 Hybridization

A volume of 12 µl purified ULS-labelled DNA was combined with 10 ul blocking agent (Agilent-CGH) and 22 ul Agilent hybridization buffer (2x), followed by incubation at 70 °C for 15 minutes. The remaining ULS-labelled DNA was used to measure dye incorporation using NanoDrop (ND 1000, Labtech, UK). The hybridization assembly was prepared by placing clean gasket slides (8x)

in the hybridization chamber base and the gasket wells were loaded with 40 µl of hybridization solution in a “drag and dispense manner”. The microarray slides were placed with the numerical barcode facing up, and the chamber covers were clamped tightly onto the assembly. The assembled chambers were rotated vertically to assess bubble mobility by tapping the hard surface to displace any stationary bubbles. The hybridization chambers were then incubated in a hybridization oven at 65 °C overnight with a rotation of 20 rpm.

2.3.5 Slide washing, scanning and serotype calls

The array-gasket sandwich was transferred to the Oligo ACGH wash buffer 1 in trough 1, from the hybridization chamber. The slides were transferred from the array-gasket sandwich to trough 2 containing wash buffer 1 and stirred for 5 minutes. The slides were transferred to trough 3 with preheated (37 °C) wash buffer 2 for 1 minute. The slides were then transferred from wash buffer 2 and scanned using a Sure-scan microarray scanner, coupled with feature extraction software (Agilent Technologies). Microarray data were analyzed as previously described using a Bayesian hierarchical model (245) to determine serotypes and their relative proportions or abundances in samples with multiple serotypes. Results were retrieved in the form of a Microsoft Excel spreadsheet from the Bugs Bioscience SentiNet website (246).

2.3.6 Statistical analysis

In samples with more than one GBS serotype, the serotype with a relative abundance greater than 50% was considered dominant, while serotypes with a relative abundance less than 50% were considered minor serotype(s). The ratio of serotypes in samples with multiple serotype colonization was calculated as follows: relative abundance of the most abundant serotype/ the relative abundance of the less abundant serotype. Analyses of data were carried out using Stata version 13 (StataCorp, College Station, TX).

2.4 Methods for analysis of the vaginal microbiome

2.4.1 Selection of study samples and sample size

The study included 282 pregnant women from three countries (South Africa, Ethiopia and Kenya) selected for an exploratory analysis of the vaginal microbiome. The study subjects were stratified by GBS colonization status, 47 pregnant women with vaginal GBS and 47 without GBS

colonization were randomly selected from each of the three study countries. The GBS colonization was identified by standard microbiological methods in the multicenter GBS prevalence study which aimed to evaluate the prevalence of GBS recto-vaginal colonization in women at the time of labor and vertical transmission thereof to their newborns (222). The study sample size of 282 pregnant women in overall, provided at least 90% power at $\alpha=0.05$ to detect proportional differences in the prevalence of vaginal organisms between women with and women without GBS colonization, assuming a prevalence of 53% of *Lactobacillus* species in women without vaginal GBS colonization and 34% in pregnant women with vaginal GBS colonization (247).

2.4.2. Selection of open array target organism and validation assays

The target organisms were selected by performing an extensive literature search on vaginal organisms that are commonly found in the vaginal microbiome or previously associated with vaginal GBS colonization and clinical outcomes such as preterm delivery or BV (Appendix 1). The primer sets of target organisms were evaluated for specificity by comparing the nucleotide sequences of the primers to the genomic sequences of all other species in the National Center for Biotechnology Information (NCBI) database (241). The selected primer gene targets were subsequently spotted on a custom-designed 56-target panel by Thermo Fisher Scientific (Waltham, Massachusetts, USA) (248). The open array panel included 56 targets consisting of 42 bacteria (4 commensal *Lactobacilli* species and other pathobionts), 6 viruses, 4 fungi (*Candida* species), and 4 internal controls (Table 2.1).

The open array assays were optimised and evaluated for linearity and qPCR efficiency within our laboratory at Wits-VIDA. Standard curves for target organisms were generated using a tenfold serial dilution of the plasmid DNA template (10^8 copies/ μ l), which contained the concatenated sequences of each target organism. The plasmid DNA (combined positive control) dilutions were represented as standards (std) as follows: std1- 5×10^1 , std2- 5×10^2 , std3- 5×10^3 , std4- 5×10^4 , std5- 5×10^5 , std6- 5×10^6 , std7- 5×10^7 . The amplification efficiency of the target assays was calculated using the following formulae: $e = (10^{-1/m}) - 1$, where m is the slope of the standard curve. The correlation coefficient (r^2) of each standard curve was also calculated. Positive detection was defined as a cycle threshold (Ct) <24 for all target organisms. The efficiency of all target assays varied from 90% to 103%, while correlation coefficients (r^2) ranged from 0.9 to 1, as presented in Table 2.2.

Table 2. 1: Vaginal organisms targeted in the 56-targets open array qPCR

The target organisms consisted of 41 bacteria, 4 fungi, 1 protozoa, 6 viruses and 4 internal controls.

Bacteria		Viruses	Fungi	Internal Controls
<i>Acinetobacter baumannii</i>	<i>Megasphaera</i> 1	Cytomegalovirus	<i>Candida albicans</i>	RNaseP
<i>Atopobium vaginae</i>	<i>Megasphaera</i> 2	Hepatitis B virus	<i>Candida glabrata</i>	Broad range 16S
Bacterial vaginosis associated bacterium 2	<i>Mobiluncus curtisii</i>	Herpes simplex virus 1	<i>Candida dubliniensis</i>	18S
<i>Bacteroides fragilis</i>	<i>Mobiluncus mulieris</i>	Herpes simplex virus 2	<i>Candida tropicalis</i>	Human globin
<i>Bacteroides vulgatus</i>	<i>Mycoplasma hominis</i>	Human herpesvirus 6		
<i>Chlamydia trachomatis</i>	<i>Mycoplasma genitalium</i>	Parvovirus B19		
<i>Clostridium septicum</i>	<i>Neisseria gonorrhoeae</i>			
<i>Corynebacterium diphtheriae</i>	<i>Peptostreptococcus anaerobius</i>			
<i>Eggerthella lenta</i>	<i>Prevotella bivia</i>			
<i>Enterobacter aerogenes</i>	<i>Prevotella intermedia</i>	Protozoa		
<i>Enterococcus faecalis</i>	<i>Proteus vulgaris</i>			
<i>Enterococcus faecium</i>	<i>Sneathia amnii</i>	<i>Trichomonas vaginalis</i>		
<i>Escherichia coli</i>	<i>Staphylococcus aureus</i>			
<i>Gardnerella vaginalis</i>	<i>Streptococcus agalactiae</i>			
<i>Haemophilus influenzae</i>	<i>Streptococcus pneumoniae</i>			
<i>Klebsiella pneumoniae</i>	<i>Streptococcus pyogenes</i>			
<i>Lactobacillus crispatus</i>	<i>Treponema pallidum</i>			

<i>Lactobacillus gasseri</i>	<i>Ureaplasma parvum</i>			
<i>Lactobacillus inners</i>	<i>Ureaplasma urealyticum</i>			
<i>Lactobacillus jensenii</i>	<i>Veillonella parvula</i>			
<i>Listeria monocytogenes</i>				

Table 2. 2: Performance of open array assays

Target organism	Efficiency (%)	Coefficient (r2)
<i>Lactobacillus crispatus</i>	101	0.99
<i>Lactobacillus gasseri</i>	101	0.99
<i>Lactobacillus inners</i>	101	0.99
<i>Lactobacillus jensenii</i>	101	0.99
<i>Atopobium vaginae</i>	101	0.99
<i>Bacteroides fragilis</i>	101	0.99
<i>Bacteroides vulgatus</i>	93	1
<i>Bacterial vaginosis associated bacterium 2</i>	103	0.99
<i>Eggerthella lenta</i>	95	0.99
<i>Gardnerella vaginalis</i>	101	0.99
<i>Ureaplasma parvum</i>	90	1
<i>Ureaplasma urealyticum</i>	103	0.99
<i>Mobiluncus curtisii</i>	98	0.99
<i>Mycoplasma genitalium</i>	91	1
<i>Mycoplasma hominis</i>	93	1
<i>Mobiluncus mulieris</i>	91	1
<i>Megasphaera 1</i>	90	1
<i>Megasphaera 2</i>	100	0.99
<i>Neisseria gonorrhoea</i>	92	1
<i>Peptostreptococcus anaerobius</i>	94	1
<i>Prevotella bivia</i>	101	0.99

<i>Prevotella intermedia</i>	93	0.99
<i>Leptotrichia amnionii</i>	93	1
<i>Acinetobacter baumannii</i>	98	0.99
<i>Corynebacterium diphtheria</i>	93	1
<i>Clostridium septicum</i>	90	1
<i>Enterobacter aerogenes</i>	102	0.99
<i>Chlamydia trachomatis</i>	101	0.99
<i>Escherichia coli</i>	90	1
Group B <i>Streptococcus</i>	91	1
<i>Enterobacter aerogenes</i>	102	0.99
<i>Enterococcus faecalis</i>	93	1
<i>Enterococcus faecium</i>	97	0.99
<i>Haemophilus influenza</i>	101	0.99
<i>Klebsiella pneumoniae</i>	94	0.99
<i>Listeria monocytogenes</i>	90	0.99
<i>Proteus vulgaris</i>	101	0.99
<i>Staphylococcus aureus</i>	92	1
<i>Streptococcus pneumoniae</i>	93	1
<i>Streptococcus pyogenes</i>	90	1
<i>Treponema pallidum</i>	91	1
<i>Trichomonas vaginalis</i>	97	0.99
<i>Veillonella parvula</i>	104	0.99
<i>Candida albicans</i>	90	1

<i>Candida dubliniensis</i>	101	0.99
<i>Candida glabrata</i>	90	1
<i>Candida tropicalis</i>	90	1
<i>Hepatitis B virus</i>	90	0.99
<i>Human cytomegalovirus</i>	93	0.99
<i>Human herpesvirus 6</i>	90	1
<i>Human simplex virus 1</i>	100	0.99
<i>Human simplex virus 2</i>	102	0.99
Parvovirus B19	94	1

2.4.3 Nucleic acid extraction

Vaginal swabs suspended in STGG were thawed, followed by a 2 minutes vortexing step. A volume of 250 µl of STGG medium was aliquoted and extracted for nucleic acids using the Nuclisens easyMAG system (bioMérieux, Marcy l'Etoile, France). Since GBS is encapsulated, vaginal swab extracts were lysed for 60 min at 37 °C in a lysis buffer that contained mutanolysin (79 units/ml; Sigma-Aldrich) and lysozyme (40 mg/ml; Sigma-Aldrich, Dorset, UK), before nucleic acid extraction. The nucleic acids were eluted using 100 µl of elution buffer. The negative controls (STGG media) were also subjected to extraction with the vaginal swab samples.

2.4.4 Quantitative Polymerase Chain Reaction

The extracted nucleic acids were analyzed to detect vaginal organisms using open array qPCR assay (TaqMan Open Array platform; Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer's instructions (248). In brief, a PCR pre-amplification step was performed using 5 µl of sample combined with 2.5 µl of 2X Taman Preamp Master Mix and 2.5 µl of the TaqMan pre-amplification pool. Reactions were amplified on a PCR thermal cycler (Bio-Rad, CA, USA) under the following conditions: 25 °C for 2 min, 50 °C for 30 min, 95 °C for 20 s, 14 cycles of 95 °C for 15 s and 60 °C for 2 min, then a final step of 99.9 °C for 10 min and a hold at 4 °C. The amplification products (2 µl per sample) from the PCR reaction were diluted with 18 µl of nuclease-free water. Finally, 5 µl of the assays were loaded from the 384-well sample plate into the open

array plate using the AccuFill (Thermo-Fisher Sc, Waltham, MA, USA), an automated pipetting instrument.

The open-array plates were cycled on the Quant-Studio 12K Flex real-time PCR system (Thermo-Fisher Sc, Waltham, MA, USA), and amplification data were analyzed using QuantStudio software (Thermo-Fisher Sc, Waltham, MA, USA), according to the manufacturer's specifications. The abundance of vaginal organisms (measured as genome copies per swab) was determined using standard curves generated from a 10-fold dilution of the plasmid DNA template. In a single run, a total of 48 samples were simultaneously analysed, including internal controls that were monitored to ensure amplification in positive controls and the absence of amplification in negative control. Samples that failed were re-extracted and analysed using the open array systems. For quantitative analysis, the Ct values obtained for each organism were compared with those from a standard curve to extrapolate the absolute densities of the target organisms.

2.4.5 Statistical analysis

Samples were assigned into community state types (CSTs) based on the organism with the highest abundance (more than 50% of the total genome copies in a swab). A diverse vaginal microbiome had a mixture of vaginal organisms with less than 50% of the total genome copies in a swab (249). Proportions of CSTs were analyzed using Fisher's exact test.

To examine whether GBS colonization was associated with other vaginal organisms, a multivariable logistic regression analysis was performed adjusting for potential confounding factors including age and country of sample collection. Adjusted odds ratios (aOR) were presented with a 95% confidence interval (CI), and Bonferroni adjusted P values ≤ 0.0013 taking multiple comparisons into consideration, were regarded statistically significant. The differences in the abundance of vaginal organisms between women with and without vaginal GBS colonization were assessed using the Wilcoxon rank-sum test.

All statistical analyses were performed using Stata version 13 (StataCorp, College Station, TX).

2.5 Ethical approval

Ethical clearance for this research study was obtained from the Medical Human Research Ethics Committee (HREC) of the University of Witwatersrand (Protocol no: M191171; Appendix 2). The parent GBS prevalence study was also approved by the Medical Human Research Ethics Committee (HREC) of the University of Witwatersrand (Ethics reference number 150214) and by

the local ethics committees of the countries in which the study samples were collected, for recruitment, sample collection and analyses of samples and storage of samples. Participants provided written informed consent at the time of initial enrollment, agreeing to participate in the study and to use the samples for future research studies.

2.6 Funding

The PhD research study, including all laboratory work was funded by the South African Medical Research Council: Vaccines and Infectious Diseases Analytics Research Unit. The whole genome sequencing study was also funded by the Bill & Melinda Gates Foundation. The studies on the vaginal microbiome and colonization with multiple GBS serotypes were funded by the NIHR Mucosal Pathogens Research Unit (grant 5750133).

Chapter 3 Genetic diversity of colonizing Group B *Streptococcus* isolates in maternal and newborn isolates from 8 countries in Africa and South-East Asia

3.1 Abstract

Background Recto-vaginal colonization by Group B *streptococcus* (GBS) in expectant women is a risk factor for invasive GBS disease in newborns and may predispose to stillbirths and preterm delivery. Five major GBS clonal complexes (CCs) (CC1, CC10, CC17, CC19 and CC23) cause the majority of invasive GBS disease, with CC17 being the most virulent. We undertook GBS genotype characterization of colonizing isolates from mother-newborn dyads across six sub-Saharan African and two Southeast Asian low and middle-income countries (LMICs).

Methods: Whole genome sequencing was done on GBS isolates cultured from the rectum or vagina of pregnant women (n=816) and mucosal colonizing isolates from their newborns (n=628) at the time of delivery. The GBS sequencing data were analyzed for capsular serotypes, sequence types (STs), CCs, select surface protein profiles and antimicrobial resistance (AMR) genes.

Findings: The multi-locus sequence typing (MLST) scheme revealed 53 and 51 STs in maternal and newborn colonizing isolates, respectively; including five newly identified rare STs. There was high concordance at the serotype (90%; 549/608), ST (88%; 533/608) and CC level (89%; 543/608) between the maternal and newborn isolates. Consequently, further detailed whole genome sequencing analyses focused on the maternal colonizing isolates.

The most common lineage in India (47.4%; 65/137), Mozambique (40.0%; 44/110) and Kenya (29.1%; 32/110) was CC23. The lineage CC17 was common in Kenya (28.2%; 31/110), Mozambique (27.3%; 30/110), and South Africa (20%; 21/105) but absent in Southeast Asian countries. Clonal complex 889 (13.5%; 12/889), and CC1212 (4.5%; 4/89) and CC255 (3.4%; 3/89), were detected exclusively in Bangladesh, albeit at low frequency. There was variation in serotypes associated with CC between countries. In particular, although serotype III strains from Mozambique (97%, 30/31), Kenya (89%; 31/35) and South Africa (77%; 17/22) were mainly associated with CC17, none of the eleven serotype III isolates from the Southeast Asian countries clustered to CC17. The *hvgA* gene was detected in 12.4% (101/816) of maternal isolates, mostly associated with serotype III (94%; 95/101), with the remaining isolates being serotype IV (4%; 4/101) and serotype Ia (1.98%; 2/101). Whereas 66.9% (95/142) of serotype III isolates included the hypervirulent A gene (*hvgA* gene), all of which were CC17; the *hvgA* gene was not present in

any of the serotype III isolates from Southeast Asia. Furthermore, the rare genotype CC934 (0.2%; 2/816) which also includes the *hvgA* gene, was identified in Ethiopia.

The prevalence of any alpha-like family genes (*alp1*, *alp2/3*, *alphaC*, and *rib*) ranged from 79.2% (80/101) in Mali to 100% (110/110) in Kenya. Tetracycline resistance genes were the most common antimicrobial resistant gene, ranging from 59.1% (81/137) in India to 100% (101/101) in Mali. The erythromycin resistance gene was modestly prevalent in Bangladesh (31.5%; 28/89) and Mozambique (20.9%; 23/110), whilst the prevalence thereof was less than 5% in Ethiopia, India, Nigeria, Mali, and Kenya.

Interpretation: The geographical differences in GBS genotype, and prevalence of virulence surface protein such as *hvgA* gene, possibly contributes to the reported regional heterogeneity of higher incidence of early onset GBS disease in sub-Saharan Africa compared with Southeast Asian countries. Phylogenetic analysis identified an emerging serotype Ia/CC934 lineage which possesses the pathogenic surface protein *hvgA* gene.

3.2 Introduction

Group B *streptococcus* (GBS) colonization of the rectum and vagina in expectant women is the major risk factor for invasive GBS disease in newborns, predispose to maternal GBS disease and is associated with stillbirth and preterm births (75, 250). Globally, approximately 18% of pregnant women are colonized by GBS, with regional heterogeneity ranging from 12.8% in Asia, 21.3% in Africa to 34.7% in the Caribbean (251, 252). The estimated incidence (per 1 000 live births) of GBS invasive disease in neonates (i.e. early-onset disease; aged 0-6 days) ranges from 0.21 in Southeast Asia to 2.0 in Southern Africa (5). Although likely multi-factorial, differences in the prevalence of GBS colonization and colonizing strains may contribute to regional differences in EOD (77, 252).

There are ten known GBS serotypes (GBS Ia, II, III, IV, V, VI, VII, VIII, and IX) according to the immunologic reactivity of capsular polysaccharides (26). The GBS strains are classified into sequence types (ST) using the multi-locus sequence typing method (MLST) based on nucleotide sequences of seven housekeeping genes (117) and are further stratified into clonal complexes (CC). The major CC associated with invasive GBS disease in young infants in Europe, America, Africa and China are CC1, CC10, CC17, CC19 and CC23 (139, 253-256). The invasive disease

potential of GBS may differ by CC, with CC17 being identified as a hypervirulent clone that commonly causes EOD as well as late-onset disease (7 to 89 days of age; LOD), including meningitis (257, 258). The increased invasiveness of the CC17 clones is attributable to the presence of the *hypervirulent (HvgA)* gene, which codes for a cell surface protein for invasion of human intestinal and blood brain barriers (259). Investigating the genetic structure of GBS strains associated with recto-vaginal colonization in pregnant women could provide insight into the epidemiology of EOD and assist in vaccine development.

Vaccination of pregnant women with a hexavalent (GBS Ia, Ib, II to V) polysaccharide-protein conjugate vaccine (35), as well as multi-antigen GBS vaccine targeting alp-like proteins (*rib*, *alp1*, *alp2/3* and *alphaC*) are under investigation to prevent invasive GBS disease in their young offspring (260). The GBS-NN is a fusion protein comprising the N-terminal domains of the alpha-like proteins (*Alp*) *alphaC* and *rib*, members of the alp family that includes *alp1-alp4* (56). The distribution of the alp-like protein vaccine antigen targets varies between serotypes, with *rib* primarily associated with serotype III (261). Furthermore, other GBS surface proteins, including pilus island proteins are under investigation as potential vaccine candidates (262). Computational algorithms identified the surface protein *gbs2106* as potential vaccine antigens, and their immunological and virulence profile are being evaluated (263).

The primary objective of this study was to determine the genetic composition of colonising GBS sequence types and clonal complexes in pregnant women and their newborns from across eight low and middle-income countries. Furthermore, we specifically analysed the capsular serotypes and GBS clonal complexes in relation to prevalence of genes coding for potential surface protein vaccine antigen targets and antimicrobial resistance (AMR) genes.

3.3 Results

A total of 816 isolates from pregnant women and 628 isolates from newborns were sequenced. Two isolates of differing serotypes were included for 29 pregnant women and 5 newborns. The majority of isolates from pregnant women (83.8%; 684/816) were from vaginal swabs, followed by urine samples (9.1%; 74/816) and rectal swabs (7.1%; 58/816). Among colonizing isolates from the newborn, 67.8% (426/628) were from the throat, 19.7% (124/628) from the skin surface and 12.4% (78/628) from rectal swabs. The number of isolates sequenced ranged from 137 in India to

57 in Ethiopia for maternal isolates, and from 101 in Kenya and Mozambique to 54 in Ethiopia for newborn isolates; Table 3.1.

Table 3. 1: The number of Group B *streptococcus* colonizing isolates in mother and newborn dyads sequenced from eight low and middle-income countries of Southeast Asia and Africa

Country	Maternal isolates (N=816); n (%) ^a	Newborn isolates (N=628); n (%) ^b	Paired isolates (N=608); n (%) ^c
South Africa	105 (12.9%)	102 (16.2%)	97 (16%)
Kenya	110 (13.5%)	101 (16.1%)	98 (16.1%)
Mozambique	110 (13.5%)	101 (16.1%)	97 (16%)
Nigeria	107 (13.1%)	99 (15.8%)	96 (15.8%)
Mali	101 (12.4%)	96 (15.3%)	96 (15.8%)
Bangladesh	89 (10.9%)	75 (11.9%)	72 (11.8%)
India	137 (16.8%)	None	None
Ethiopia	57 (7.0%)	54 (8.6%)	52 (8.6%)

*None – the analysis of the newborn isolates from India was not done as isolates were not shipped to South Africa. ^{a,b,c}Data are numbers (column percentage).

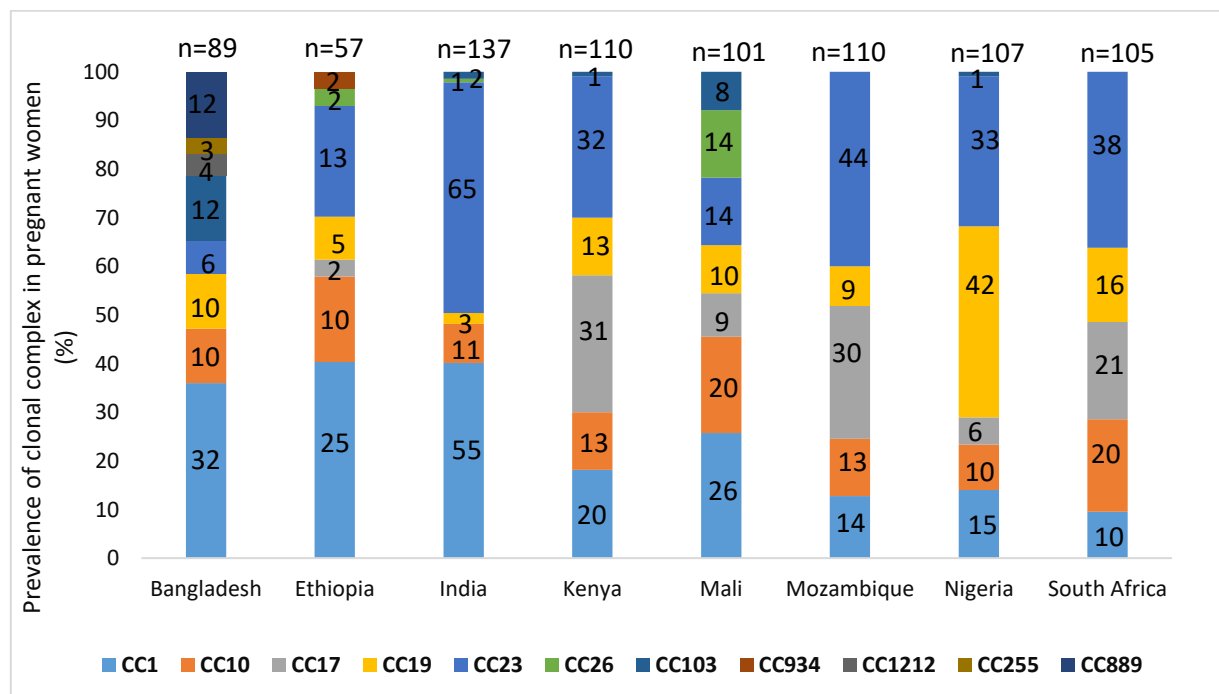
Distribution of sequence types and clonal complexes in colonizing isolates from low and middle income countries of Southeast Asia and Africa

There were 53 distinct STs in maternal and 51 STs in newborn colonizing isolates identified by MLST. Among the 608 paired mother and newborns isolates included in the study, 549 (90%; 95% CI: 88%-92%) were of the same serotype and ST, whilst 59 paired isolates were of differing serotype and ST. The proportion of similarity at ST or CC level between mother-newborn pairs were 88% (533/608; 95% CI: 85%-90%) and 89% (543/608; 95% CI: 87%-92%), respectively. Consequently, further whole genome sequencing analyses only focused on the maternal colonizing isolates, with the results for the newborn isolates expected to be similar to that of maternal isolates.

There were five new STs identified in maternal isolates, including ST2090 (serotype Ib, Bangladesh), ST2091 (serotype Ib, Bangladesh), ST2092 (eight serotype II isolates, Bangladesh), ST2093 (three serotype II isolates, Ethiopia) and ST2095 (serotype II, Mozambique). The STs of GBS isolates further clustered into eleven CCs which varied across geographic localities (Figure 3.1). The top five major CCs which accounted for 93% (756/816) of the maternal isolates were CC1 (24%; 197/816), CC10 (13%; 107/816), CC17 (12%; 99/816), CC19 (13%; 108/816) and CC23 (30%; 245/816).

Country-specific analysis showed some similarities in the distribution of CCs across sites, with CC23 being the most common lineage in India (47.4%; 65/137), Mozambique (40%; 44/110) and Kenya (29.1%; 32/110). Furthermore, CC1 was the most common lineage in Ethiopia (43.9%; 25/57), Bangladesh (36%; 32/89) and Mali (25.7%; 26/101). The lineage CC17, composed mainly of ST17 (72.7%; 72/99), was common in Mozambique (27.3%; 30/110), Kenya (28.2%; 31/110) and South Africa (20%; 21/105) but absent in Southeast Asian countries (Figure 3.1 and Appendix 3). Clonal complex 889 (13.5%; 12/89), and CC1212 (4.5%; 4/89) and CC255 (3.4%; 3/89), were detected exclusively in Bangladesh, albeit at low frequency.

Figure 3.1: Prevalence of clonal complexes in maternal colonizing isolates from sub-Saharan Africa and Southeast Asian countries.

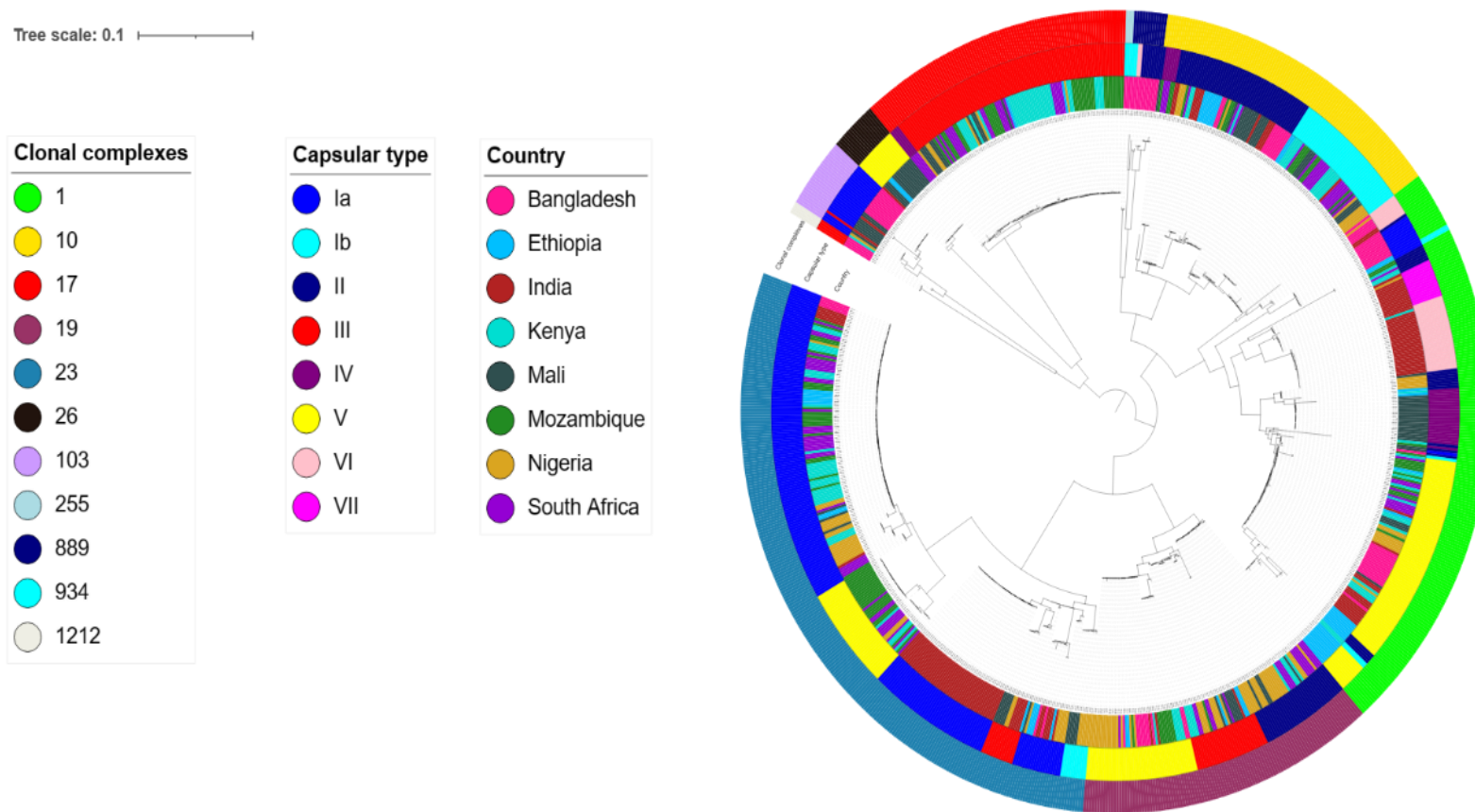


Further analysis showed intra serotype-CC or serotype-STs association which was varied by country (Figure 3.2, Appendix 4 and Appendix 5).

Figure 3.2: Mid-point rooted phylogenetic trees of GBS colonizing isolates among pregnant women from sub-Saharan Africa and Southeast Asian countries. (Fig 3.2a) The population structure of CCs in GBS colonizing isolates in pregnant women from sub-Saharan Africa and South Asian countries. The inner ring represents the country of isolate collection, the middle ring shows one of the eight capsular serotypes identified among the 816 maternal colonizing GBS isolates, and the outer ring indicates 11 CCs identified in the study. The phylogenetic tree of each of the major CCs built using SNPs after correcting for recombination for CC1 (Fig 3.2b), CC10 (Fig 3.2c), CC17 (Fig 3.2d), CC19 (Fig 3.2e) and CC23 (Fig 3.2f). The rings of the major CCs phylogenetic trees are colored according to country of isolate collection (inner ring), capsular serotype (middle ring) and ST (outer ring).

In overall, 11 clonal complexes represented by 8 serotypes were identified in all the maternal colonizing isolates (Figure 3.2a). The most common CC23 (30%; 245/816) consisted mainly of serotypes Ia (75%; 184/245), V (16%; 39/245) and III (5%; 13/245). Conversely, the least common clonal complexes were 889 (2%; 12/816), CC1212 (0.5%; 4/816) and CC255 (0.4%; 3/816).

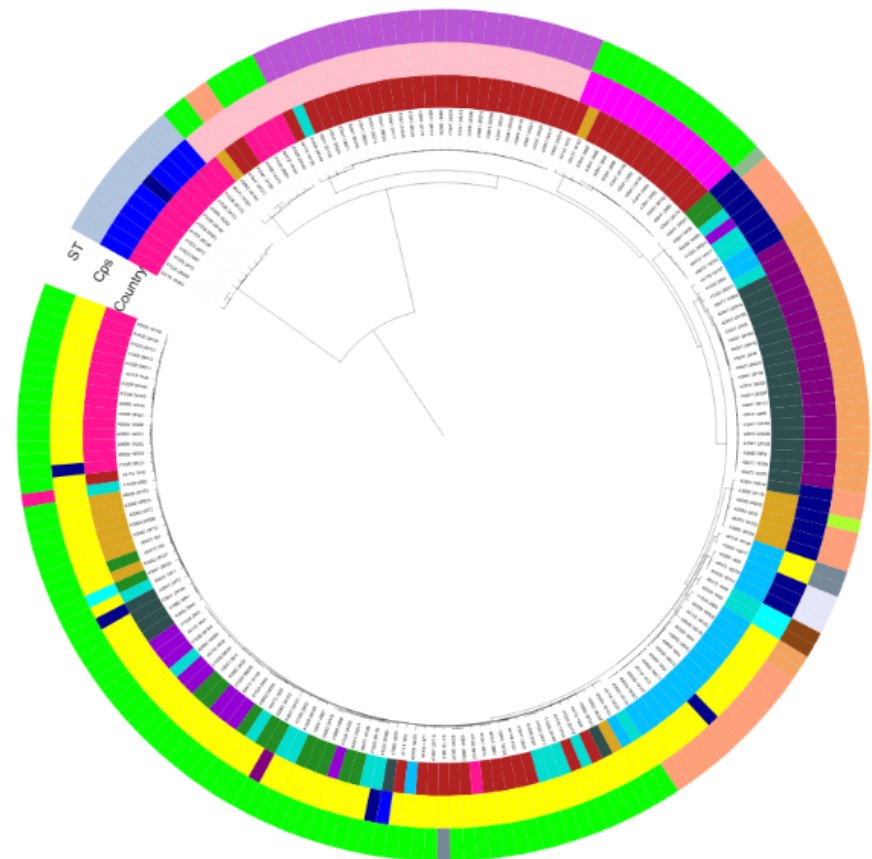
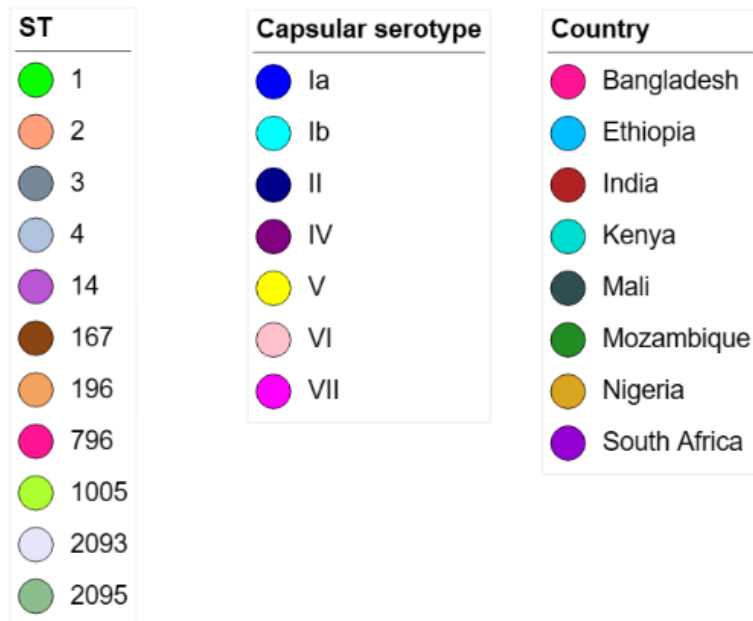
Fig 3.2a: Group B *streptococcus* colonizing isolates clonal complex distribution by serotype and country in pregnant women at the time of delivery.



The clonal complex 1 was composed mainly of serotype V (45%; 89/197), VI (18%; 35/197) and IV (11%; 22/197). Most of the CC1 isolates from Mali were serotype IV (35%; 9/26), while Indian isolates were serotype VI (51%; 28/55) (Figure 3.2b).

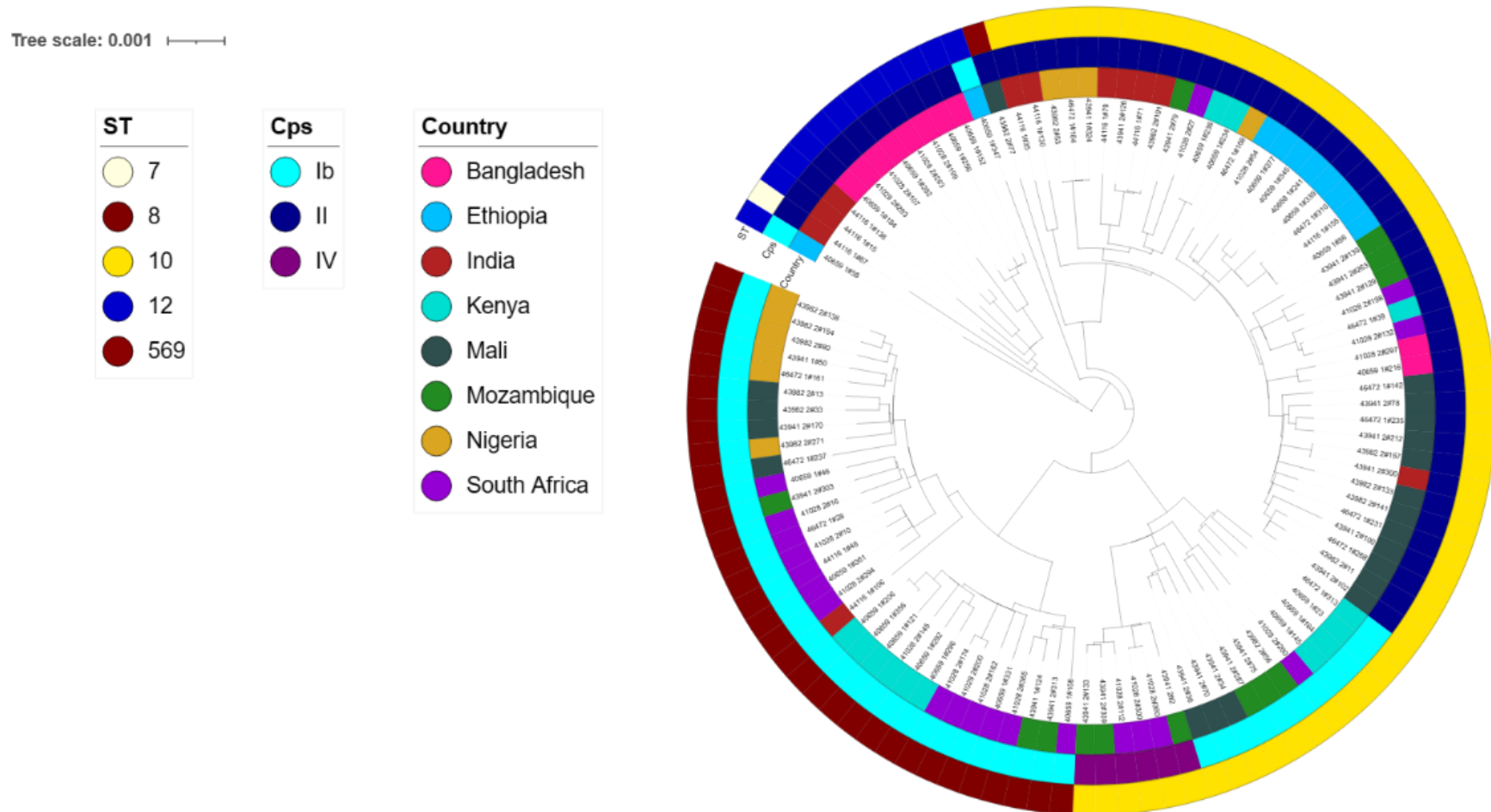
Fig 3.2b: Group B *streptococcus* colonizing isolates of clonal complex 1 distribution by serotype, sequence type and country in pregnant women at the time of delivery.

Tree scale: 0.01



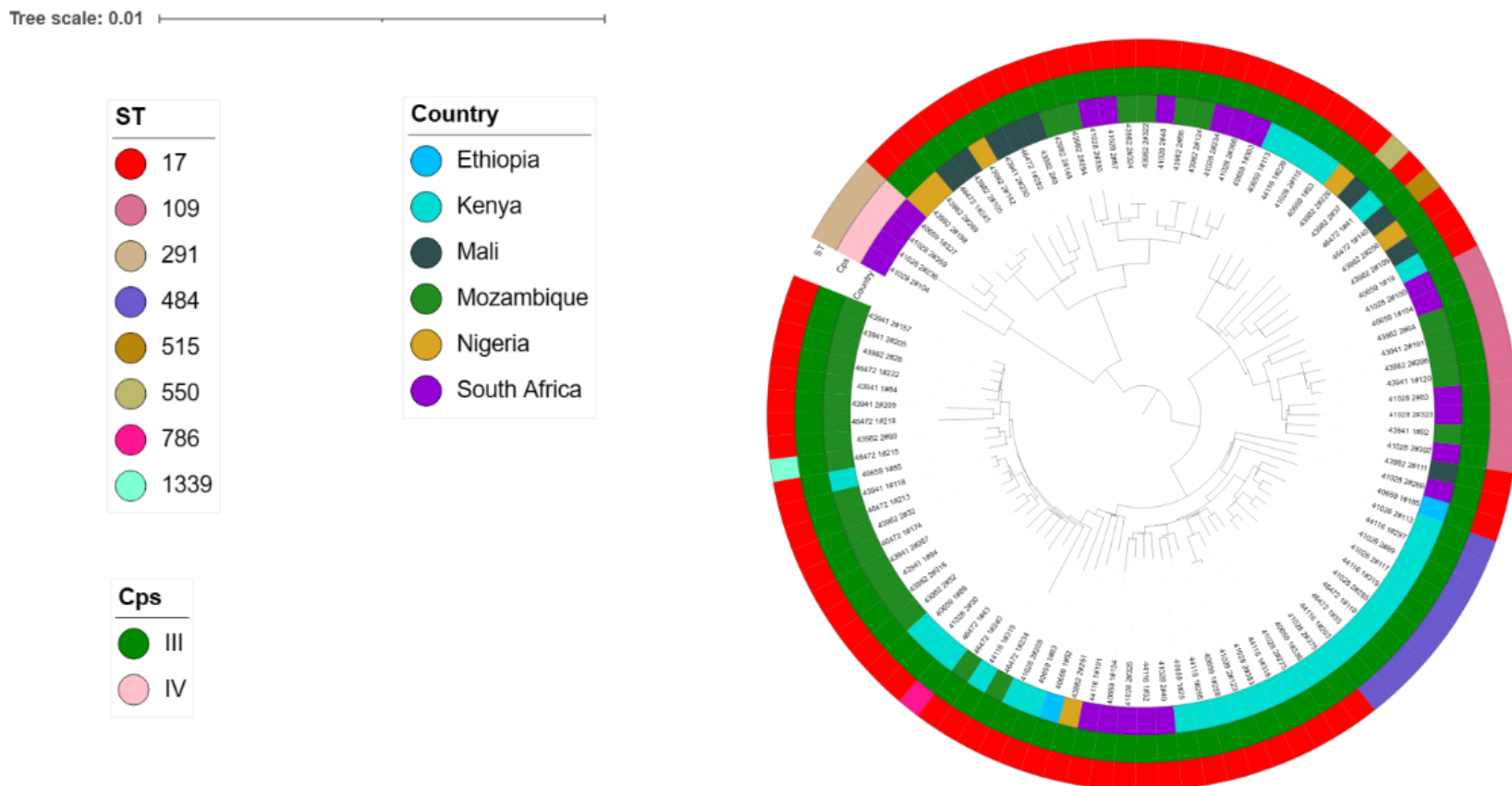
The majority of CC10 isolates were ST 10 (56%; 60/107) and also serotype II (51%: 55/107) (Figure 3.2c).

Fig 3.2c: Group B *streptococcus* colonizing isolates of clonal complex 10 distribution by serotype, sequence type and country in pregnant women at the time of delivery.



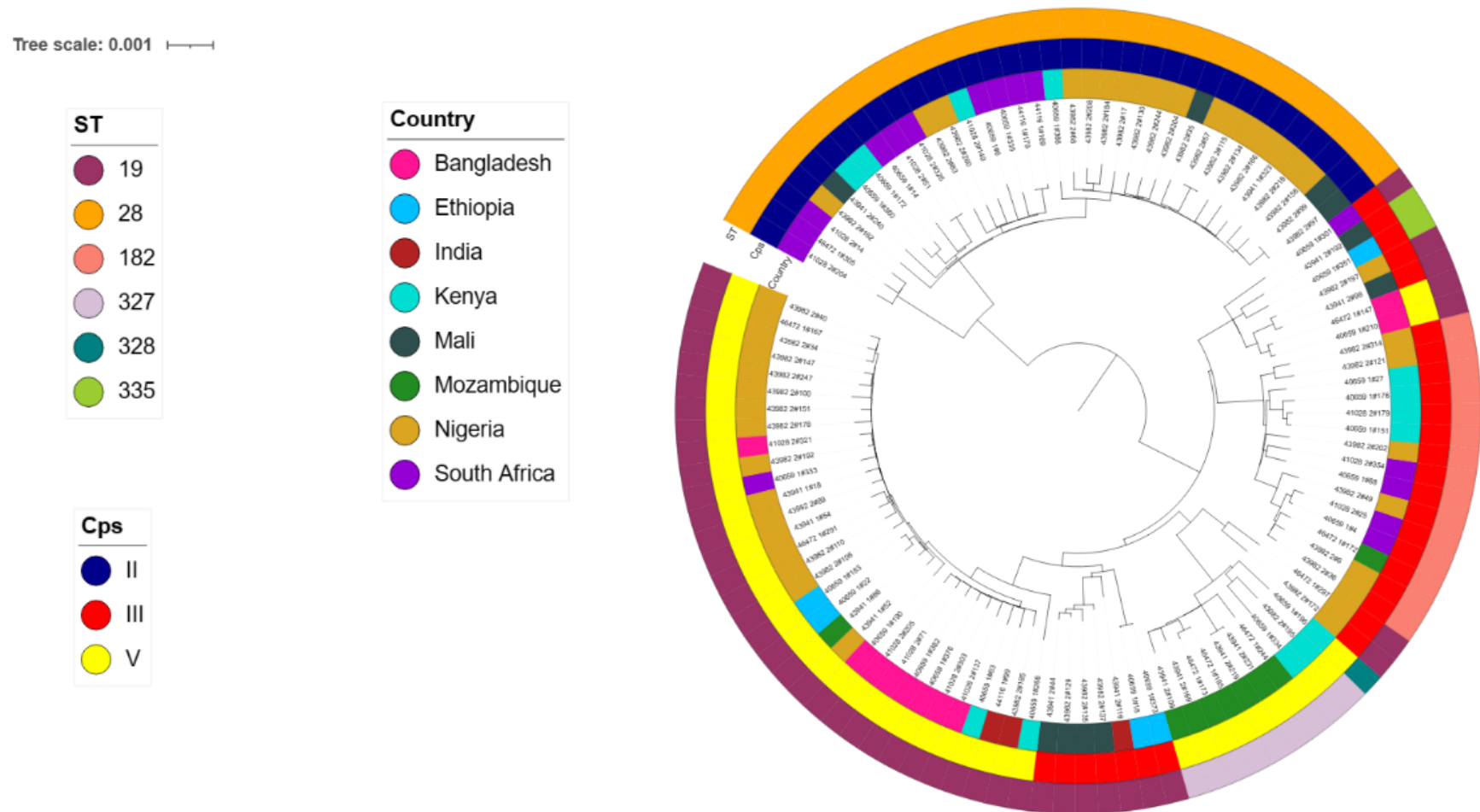
The CC17 consisted of eight sequence types including the most being ST17 (63%; 62/99), ST109 (10%; 10/99) and ST484 (7%; 7/99). Sequence type 17 was associated exclusively of serotype III, while four South African isolates of CC17/ST109 were associated with serotype IV (Figure 3.2d).

Fig 3.2d: Group B *streptococcus* colonizing isolates of clonal complex 17 distribution by serotype, sequence type and country in pregnant women at the time of delivery.



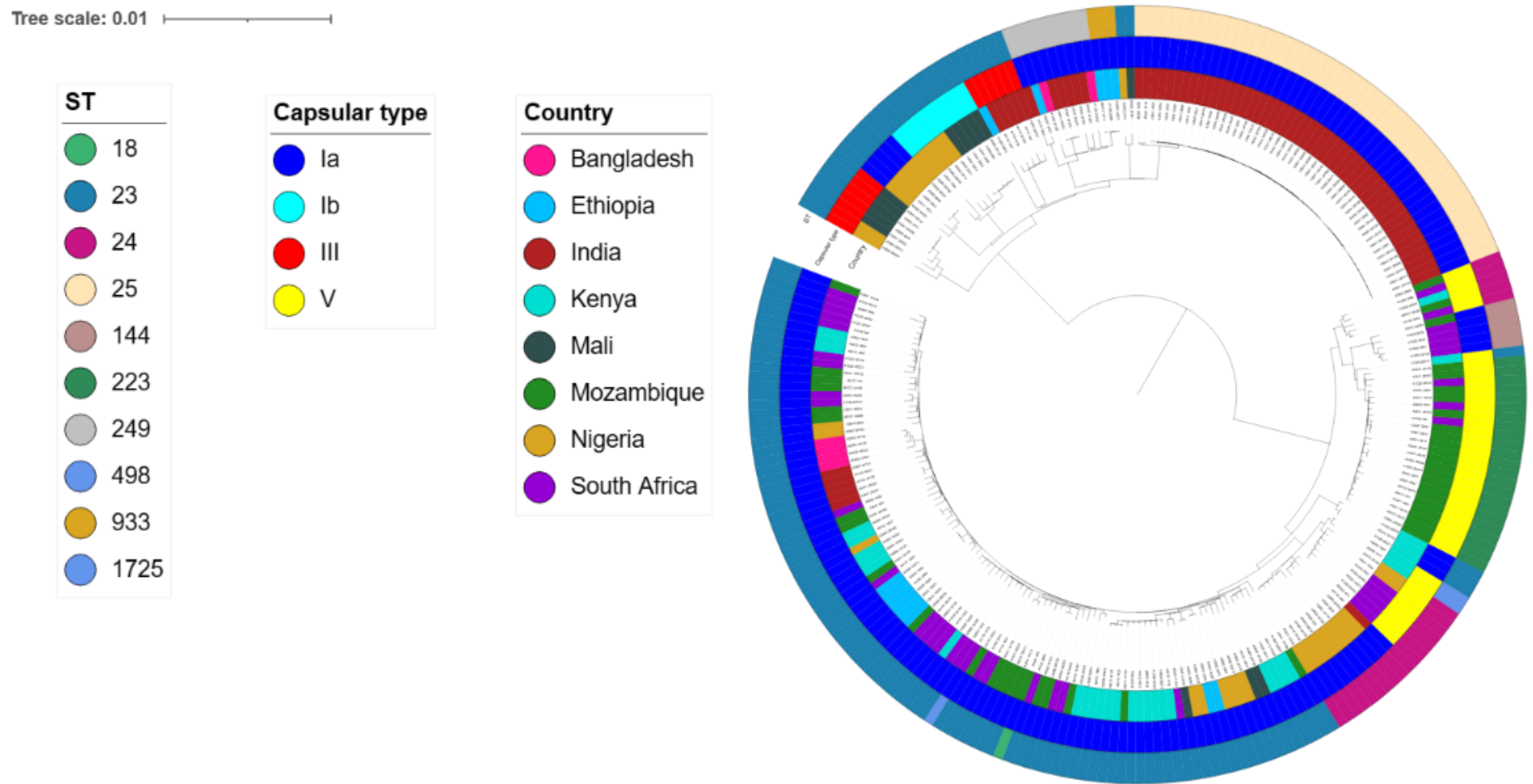
Clonal complex 19 was composed of serotypes II (32%; 35/108), III (27%; 29/108) and V (41%; 44/108) (Figure 3.2e). The CC19 isolates from African countries were associated with serotype III, and non-serotype III in isolates from Bangladesh.

Fig 3.2e: Group B *streptococcus* colonizing isolates of clonal complex 19 distribution by serotype, sequence type and country in pregnant women at the time of delivery.



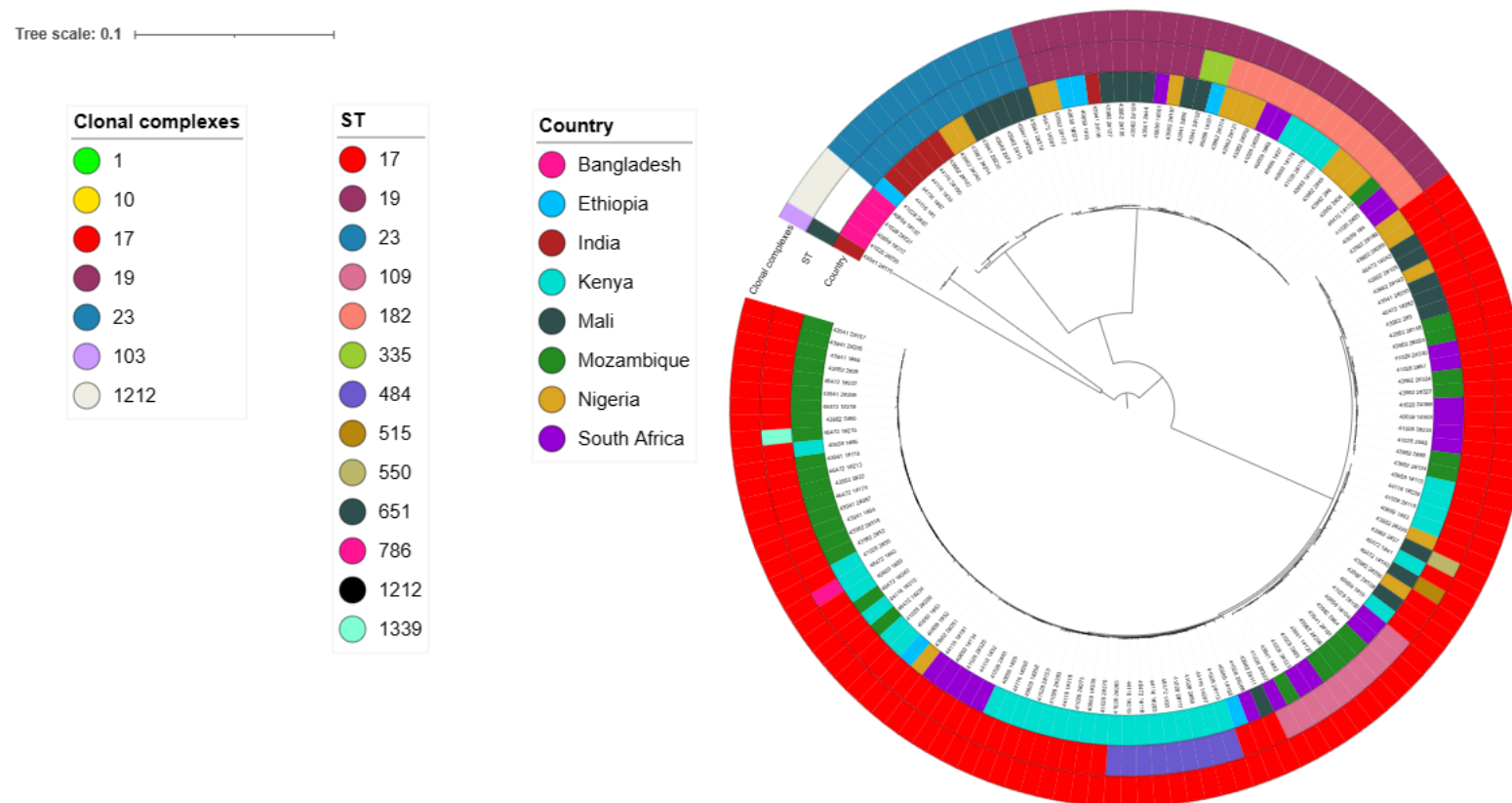
The most common serotypes in CC23 isolates were serotype Ia (184/245), V (38/245) and Ib (10/184). Indian and Malian CC23 isolates were also associated with serotype III (figure 3.2f).

Fig 3.2f: Group B *streptococcus* colonizing isolates of clonal complex 23 distribution by serotype, sequence type and country in pregnant women at the time of delivery.



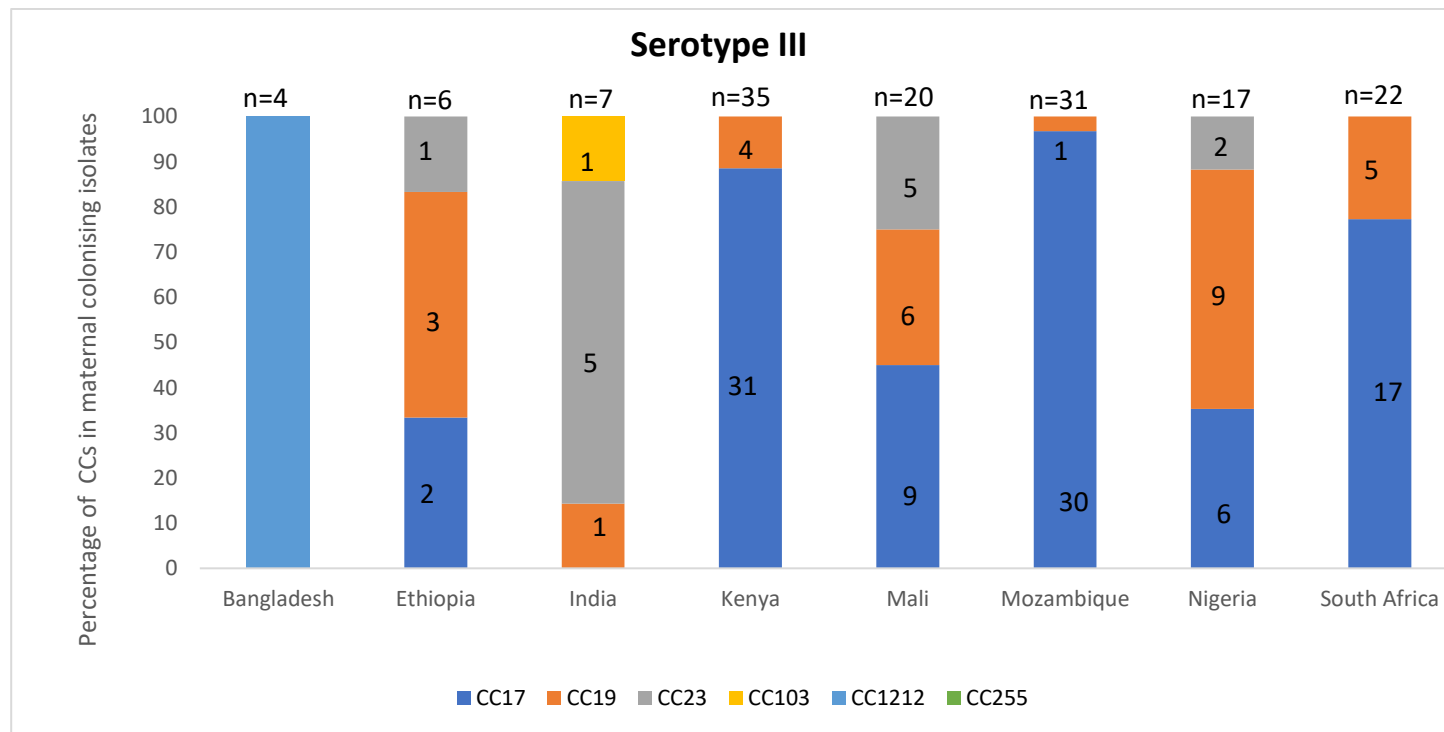
Clonal complex-17 lineage isolates were dominantly serotype III (67%; 95/142), followed by CC19 (20%; 29/142), CC23 (10%; 13/142), CC1212 (2.8%; 4/142) and CC103 (0.7%; 1/142) (Figure 3.3).

Figure 3. 1: The phylogenetic tree of GBS serotype III sequence types and CCs in pregnant women from sub-Saharan Africa and Southeast Asia. The rings of the phylogenetic tree are colored according to country of isolate collection (inner ring), ST (middle ring) and CC (outer ring).



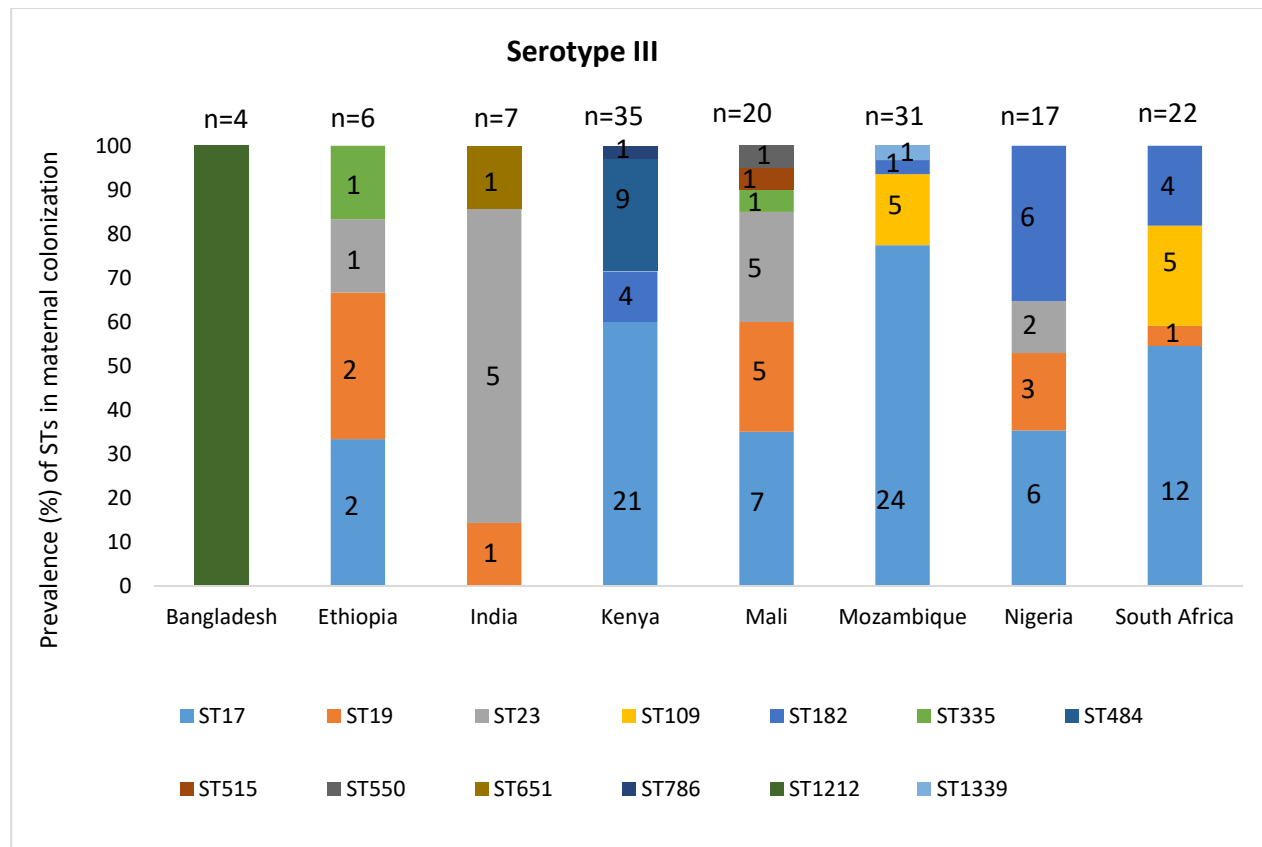
Serotype III isolates were mainly associated with CC17 in African countries; 88.6% (31/35) in Kenya, 96.8% (30/31) in Mozambique and 77% (17/22) in South Africa. In Nigeria, serotype III isolates were associated with CC17 in 35.3% (6/17), CC19 in 52.9% (9/17) and CC23 in 11.8% (2/17). Serotype III were associated with CC17 (45%; 9/20), CC19 (30%; 6/20) and CC23 (25%; 5/20) in Mali, whilst being associated with CC17 (33.3%; 2/6), CC19 (50%; 3/6) and CC23 (16.7%; 1/6) in Ethiopia (Figure 3.4).

Figure 3. 2: The prevalence of CCs in serotype III maternal colonizing isolates from sub-Saharan Africa and Southeast Asian countries



Fifty-five percent (72/131) of the serotype III isolates in sub-Saharan countries carried a ST17 backbone (Figure 3.5). Serotype III only constituted 4.9% (11/226) of GBS isolates in Southeast Asian countries and was associated with CC23 (71.4%; 5/7), CC19 (14.3%; 1/7) and CC103 (14.3%; 1/7) in India and CC1212 (100%; 4/4) in Bangladesh. Furthermore, serotype III isolates from India belonged to ST19 (14.3%; 1/7), ST23 (71.4%; 5/7), ST651 (14.3%; 1/7), and to ST1212 in Bangladesh (100%; 4/4) (Figure 3.5).

Figure 3. 3: The STs associated with serotype III in maternal GBS colonizing isolates from sub-Saharan Africa and South Asian countries.



The *hvgA* surface protein antigen targets in maternal isolates from Sub-Saharan countries

The *hvgA* gene was detected in 12.4% (101/816) of the maternal colonizing isolates, and only in isolates from Sub-Saharan countries, including Kenya (28.2%; 31/110), Mozambique (27.3%; 30/110), Ethiopia (7%; 4/57) and South Africa (20%; 21/105). Most *hvgA* positive isolates were serotype III (94.1%; 95/101) and belonged to CC17 (98%; 99/101). In addition, there were four *hvgA* gene positive isolates associated with serotype IV/CC17 (4.0%; 4/101) in South Africa and two serotype Ia/CC934 (2.0%; 2/101) isolates from Ethiopia (Figure 3.6a). The *hvgA* gene was also identified in ST291 (serotype IV in South Africa, n=4) and ST934 (serotype Ia in Ethiopia, n=2). The *hvgA* gene was not identified from any of the serotype III isolates from India (n=7) or Bangladesh (n=4); and also not identified in relation to any other GBS isolates from these countries.

Figure 3. 4: The number and phylogenetic tree of *hvgA* positive isolates in pregnant women from African and South-east Asian countries. Fig 3.6a The prevalence of *hvgA* gene in maternal colonizing GBS isolates from African and South-east Asian countries. Fig 3.6b The phylogenetic relatedness of *hvgA* gene positive GBS isolates. Note: The phylogenetic tree of GBS isolates is annotated according to the CCs and serotype. The presence of the alpha family proteins (blue), *srr* gene targets (brown), pilus proteins (purple) and the absence (yellow) of genes are plotted with respect to each isolate’s phylogenetic placement.

Fig 3.6a: The percentage of *hvgA* gene positive colonizing isolates in pregnant women from Southeast Asia and Africa.

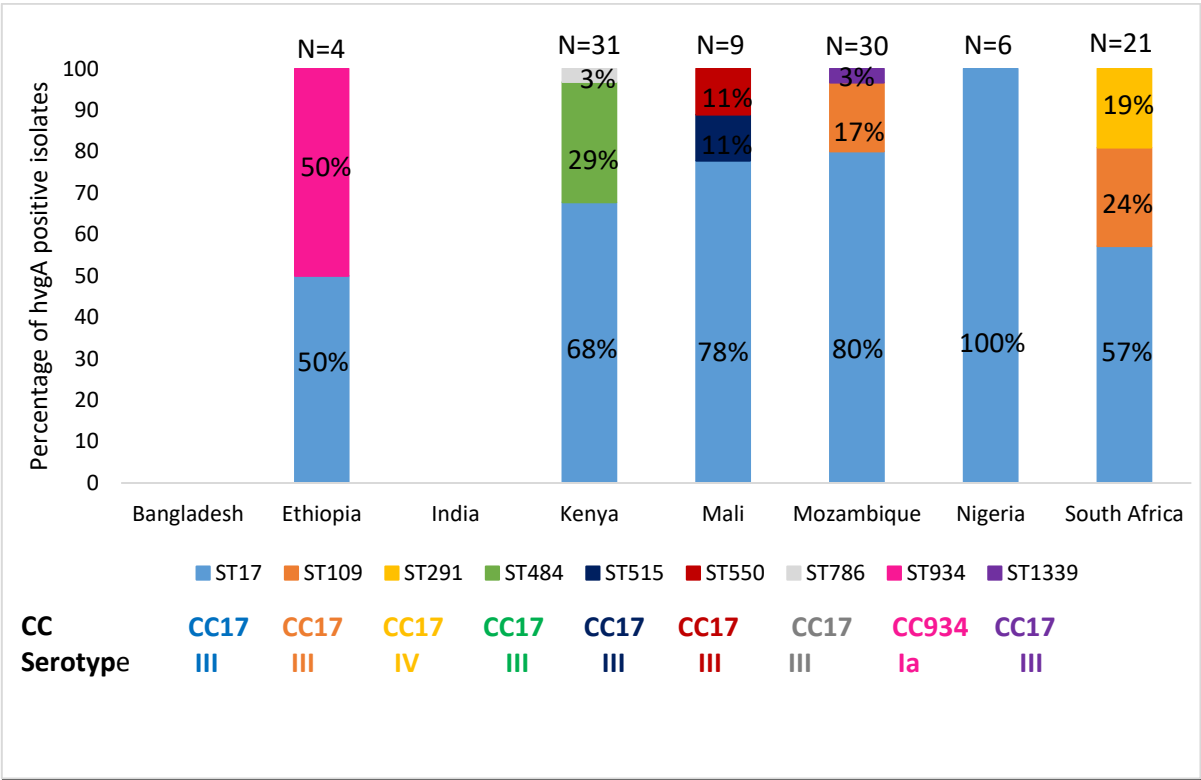
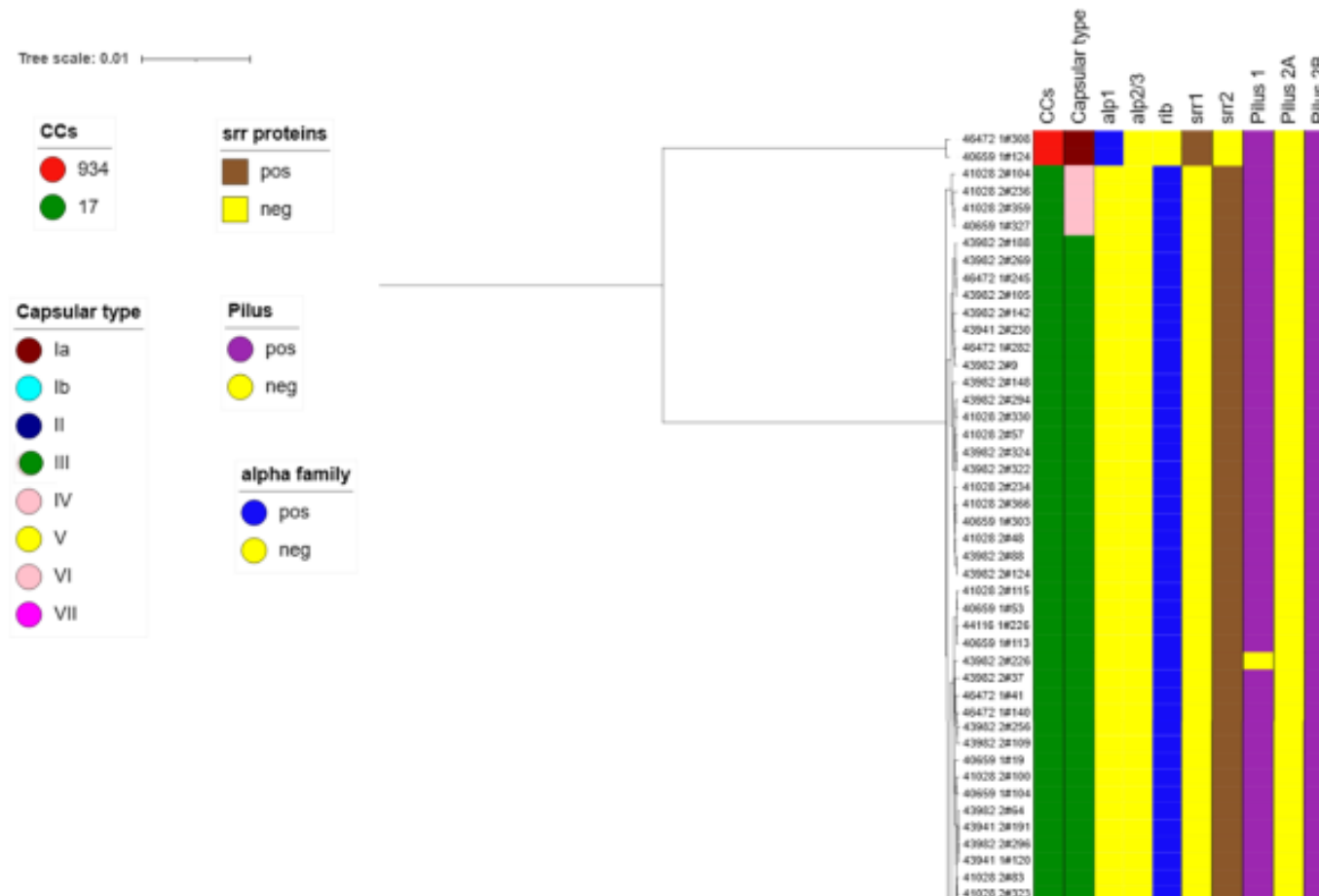


Figure 3.6b shows the phylogenetic relatedness of the *hvgA* positive isolates, with serotype IV isolates (n=4) having a similar genetic backbone as serotype III/CC17 isolates (n=95) in relation to surface protein antigens, while the serotype Ia (n=2) isolates formed a separate clade.

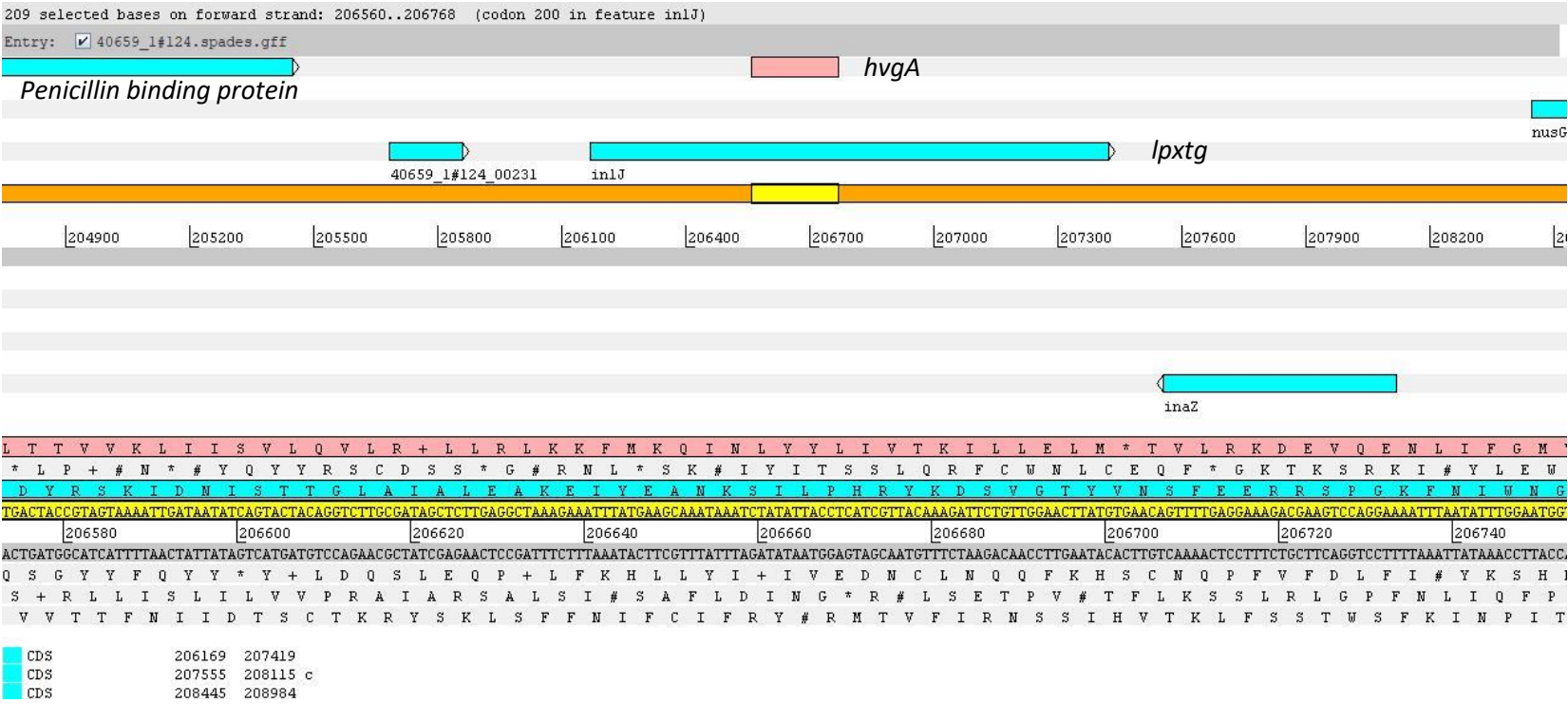
Fig 3.6b Phylogenetic relationship of *hvgA* positive isolates and the genomic backbone of the *hvgA* positive isolates



Further analysis of the two non CC17 *hvgA* gene positive isolates, sampled in 2017 and 2018 showed that the 209 kb *hvgA* gene was located in the genome of isolate 40659_1#124 at coordinates 206560-206768 (264)(Fig 3.7a). The *hvgA* gene was located within the *IniJ* feature (locus tag 40659_1#124_00232), a putative peptidoglycan linked protein (*LPXTG* motif) containing a YSIRK signal peptide which had 99.6% identity with *bibA* gene (GenBank accession number: AM183383.1).

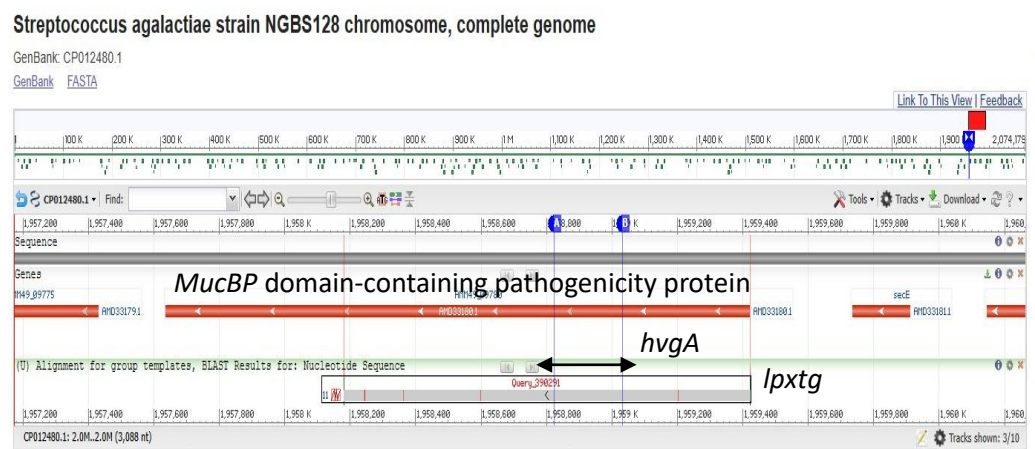
Figure 3. 5: The position of the *hvgA* gene and pathogenic proteins in serotype Ia colonizing isolates. The *hvgA* is indicated as a yellow block within the genome, and genomic features are indicated in blue. Fig 3.7a The genome view of the *hvgA* gene, *lpxtg* within the serotype Ia isolate 40659_1#124. Fig 3.7b The position of the pathogenicity protein, *lpxtg* and *hvgA* in the reference CC17 genome NGB S128. Fig 3.7c The genome view of the *hvgA* gene, *LPXTG* motif gene in the genome of serotype Ia isolate 46472_1#308.

Fig 3.7a The position of the *hvgA* gene in the genome of isolate 40659_1#124.



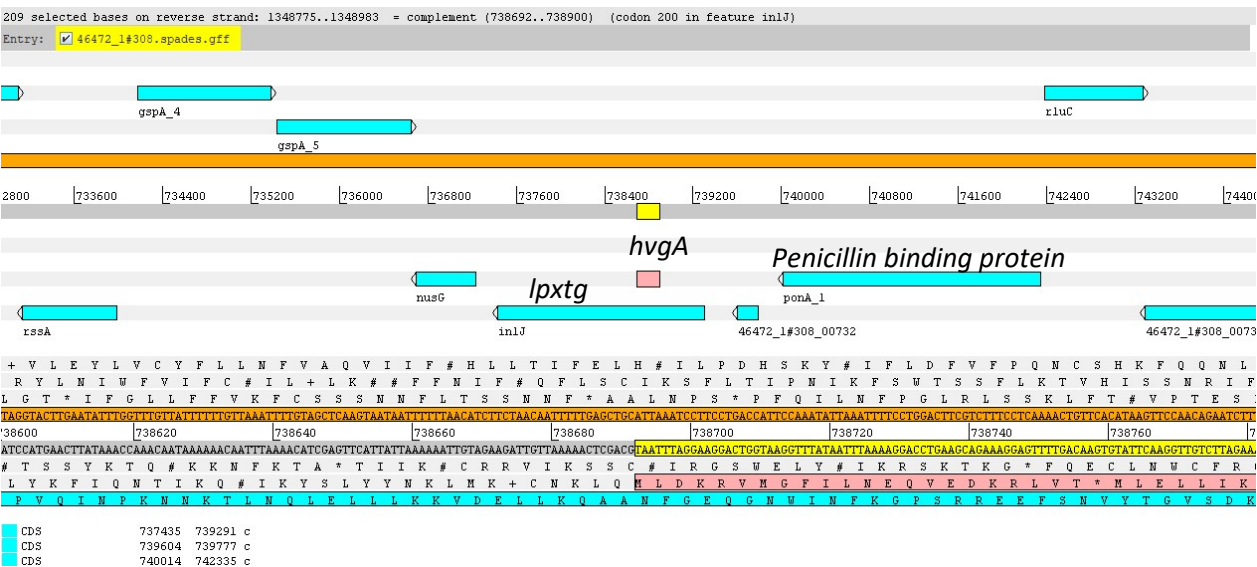
The analysis showed that the *hvgA* and *LPXTG* in isolate 40659_1#124 was within a region which had sequence homology with the AMD49, a *MucBP* domain-containing pathogenicity protein (NGBS locus tag: AMM49_RS09790), located close to a penicillin binding protein (*pbp2a*) (Fig 3.7b).

Fig 3.7b The position of the *hvgA* and *lpvg* position within the reference genome NGBS128.



Similarly, in serotype Ia isolate 46472_1#308, the *hvgA* was located on the *LPXTG* gene (coordinates: 738692-738899), a procycline repetitive protein, also located next to preprotein translocus and *penicillin binding protein* gene (Fig 3.7c).

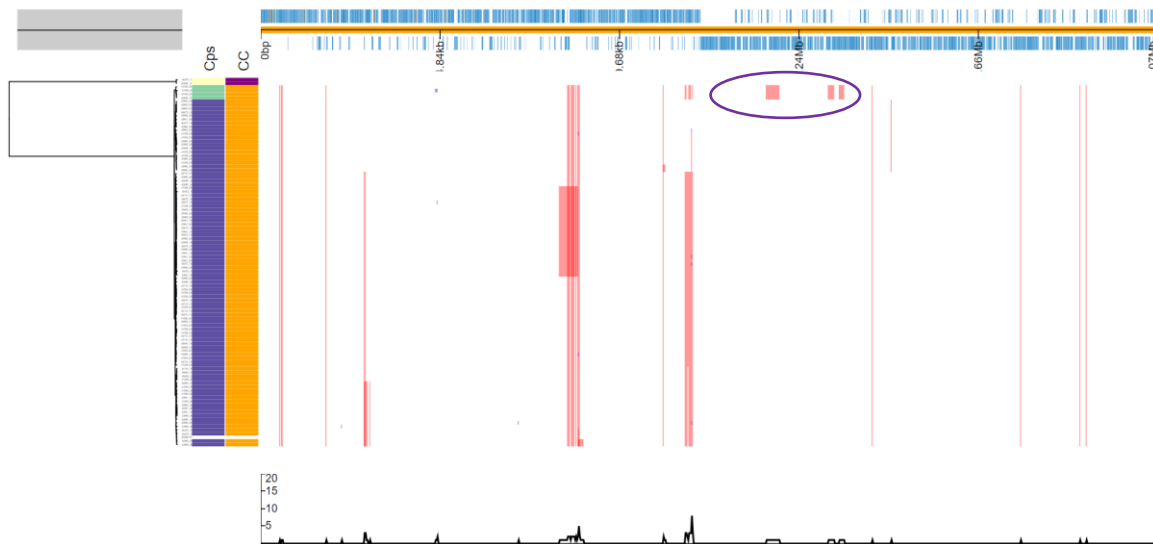
Fig 3.7c The genome view of the *hvgA* in serotype Ia isolate 46472_1#308.



Further analysis was undertaken to determine whether the serotype Ia isolates acquired the *hvgA* gene by recombination, however, there was no evidence of recombination within the *hvgA* genomic region and capsular gene region in the two serotype Ia isolates (Fig 3.8a, b and c). Analysis of the serotype IV/CC17 strains showed unique recombination blocks within the genomes of the four isolates, relative to the reference serotype III/CC17 genome (Fig 3.8a). The recombination region was identified in a 31.3kb region (coordinates within the reference NGBS128 genome: 1169247-1200575), which included genes for the glycosyltransferase family proteins, capsule biosynthesis protein *CapK* (locus tag:AMM49_RS06100) and capsular polysaccharide biosynthesis protein *Cps4B* (locus tag:AMM49_RS06130) (Fig 3.8d). The recombination block covering the capsular gene region was also observed in only the serotype IV isolates among the CC17 Isolates (Fig 3.8e and Appendix 6)

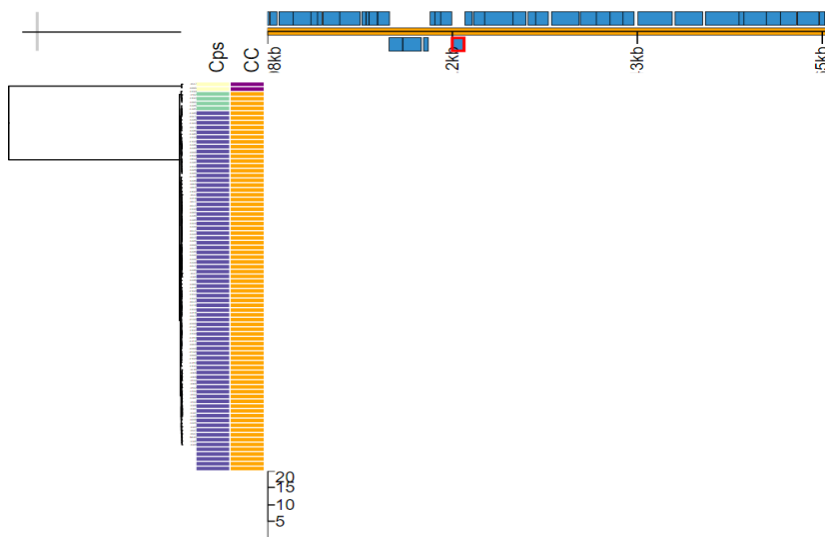
Figure 3. 6: Recombination analysis of *hvgA* positive maternal colonising isolates. The phylogenetic tree (far left) is annotated with metadata, including the capsular serotype III (purple colour), IV (green) and Ia (yellow). The CC17 is indicated by orange colour and CC934 by purple colour. The blocks in red represent recombination in more than one isolate, whilst the blue colour represents recombination within a single isolate. The top blue blocks represent the position of genes in the reference genome (serotype III CC117 NGBS128 isolate). (Fig 3.8a) Recombination blocks in all *hvgA* positive colonizing isolates. The recombination blocks unique to four serotype IV isolates are highlighted in a purple circle (Fig 3.8a). A zoomed part of the *hvgA* genomic region showed no evidence of recombination in the isolate (Figure 3.8b) 40659_1#124, and (Fig 3.8c) isolate 46472_1#308. (Figure 3.8d) A recombination block common across the serotype IV isolates among *hvgA* positive isolates. (Fig 3.8e) A recombination block observed across the four serotype IV isolates but absent in other CC17 maternal colonizing isolates.

Fig 3.8a: Recombination regions within the genome of *hvgA* positive isolates



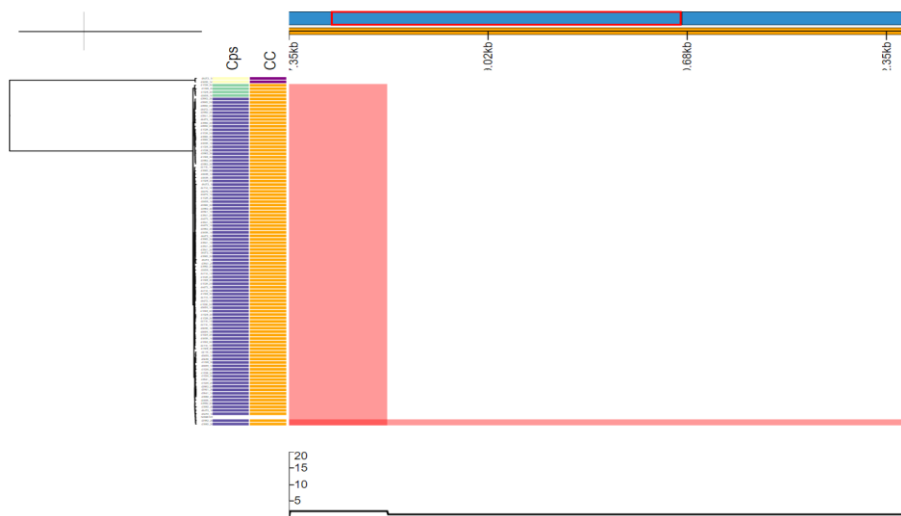
The *hvgA* genomic region showed no evidence of recombination in the isolate 40659_1#124 (Figure 3.8b).

Fig 3.8b The genomic region of the *hvgA* gene in serotype Ia isolate 40659_1#124



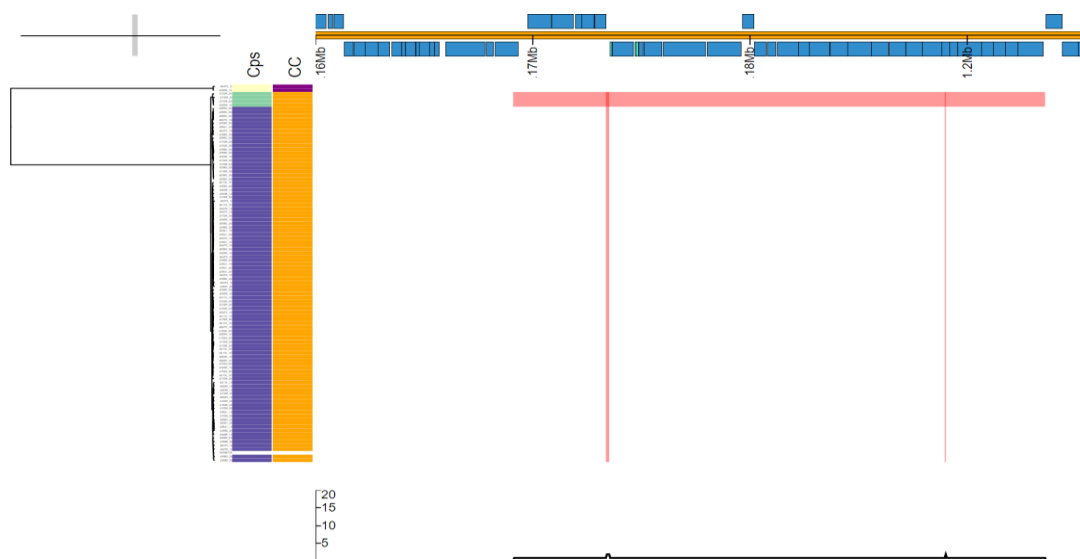
There was no evidence of recombination in the serotype Ia isolate 46472_1#308

Fig 3.8c The genomic region of the *hvgA* in serotype Ia isolate 46472_1#308



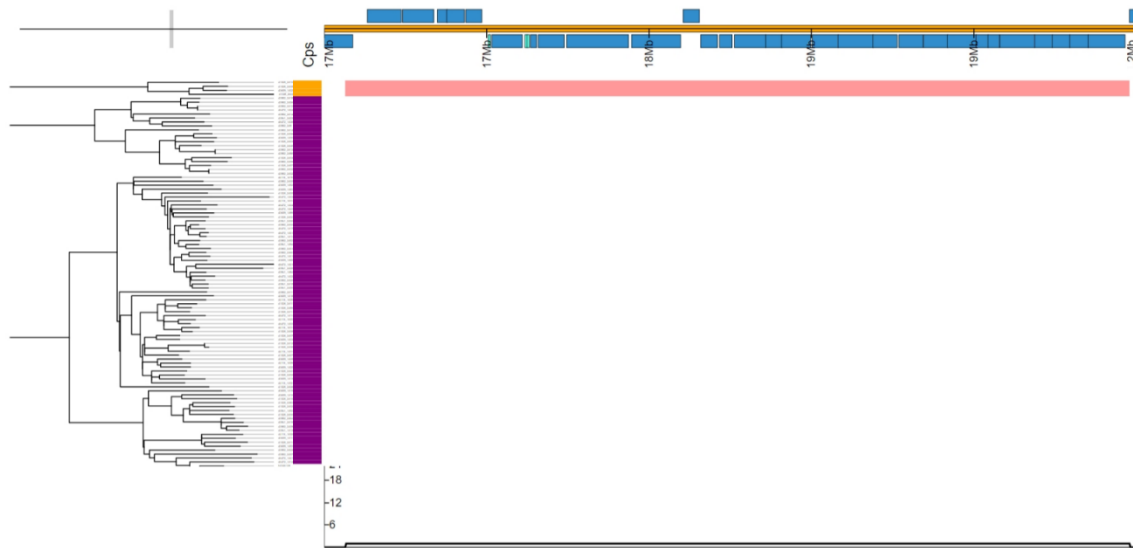
The analysis showed evidence of recombination among the *hvgA* positive serotype IV isolates.

Fig 3.8d Recombination block within the serotype IV isolates in phylogenetic tree with other *hvgA* positive isolates.



Further analysis of all CC17 isolates showed evidence of recombination within the genomic regions of serotype IV isolates and not in other CC17 isolates (Figure 3.8e).

Fig 3.8e Recombination block within the serotype IV isolates but absent in other CC17 serotype III isolates.



Prevalence of select surface protein antigen which are potential vaccine targets, in colonizing GBS isolates in pregnant women from Sub-Saharan countries and Southeast Asian countries

At least one of the alpha family protein genes (*alp1*, *alp2/3*, *alphaC*, and *rib*) was detected in 94.4% (770/816) of colonizing isolates, although being lower in Mali (79.2%; 80/101) and Bangladesh (80.9%; 72/89). Of the alp-like family protein genes, the *alp2/3* gene was more prevalent in India (60.6%; 83/137) compared with other countries where *alp2/3* prevalence ranged from 8.6% (9/105) in South Africa to 21.5% (23/107) in Nigeria (Figure 3.9 and Table 3.2). In contrast, the *rib* gene was more commonly prevalent in Kenya (37.3%; 41/110) and Mozambique (47.3%; 52/110), while being uncommon in isolates from India (2.2%; 3/137) and Bangladesh (3.4%; 3/89). Overall, the *rib* gene was highly prevalent in serotype III (85.9%; 122/142) isolates, ranging from 75% (15/20) in Mali to 100% (31/31) in South Africa, whilst being uncommon in serotype III isolates from India (14.2%; 1/7). All isolates that were *hgvA* gene positive (101/101) included at least one of the four alpha family proteins.

In each country, almost all isolates (98.4%; 803/816) harbored at least one of the three pilus island protein genes (PI-1, PI-2a and PI-2b), with the pilus 1/2a genes being most common (38.2%; 312/816) (Table 3.2).

Figure 3. 7: Prevalence of the alpha surface protein genes in colonizing GBS isolates from Southeast Asia and African countries.

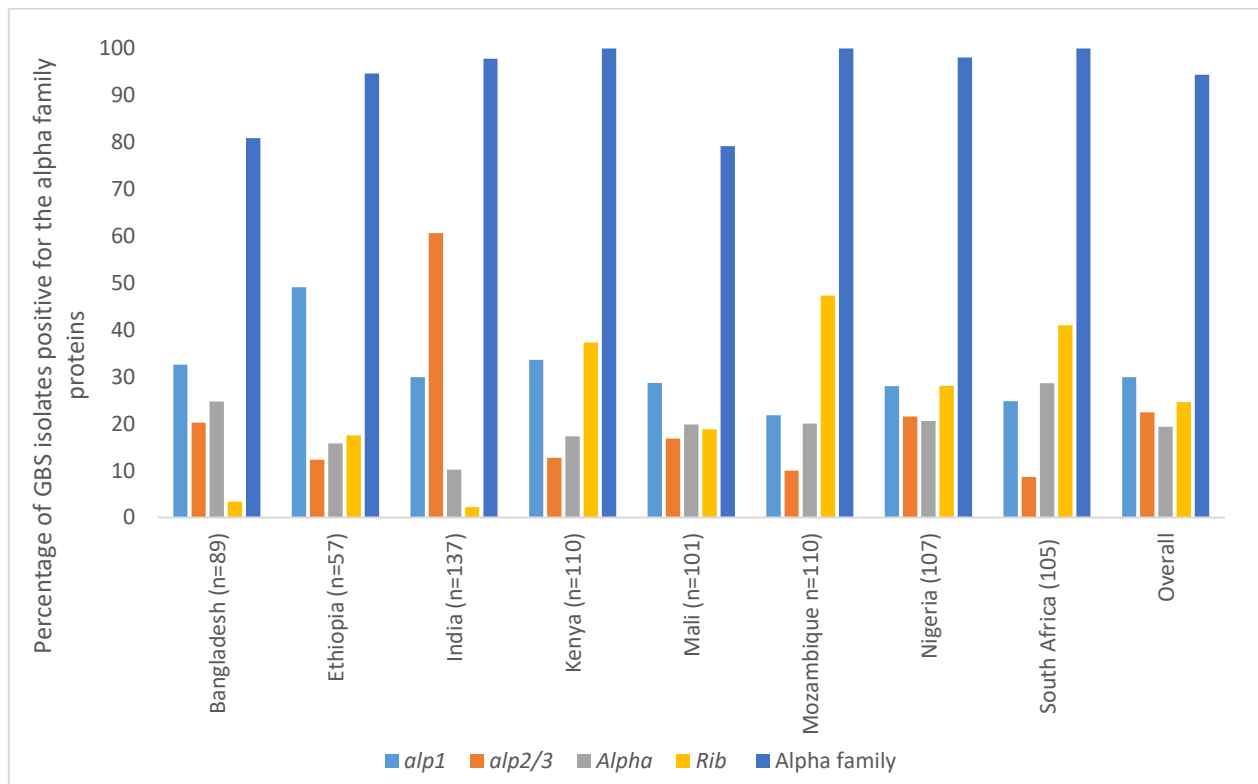


Table 3. 2: The prevalence of surface proteins in GBS maternal colonizing isolates from eight low and middle-income countries of Southeast Asia and Africa

Surface protein genes	Bangladesh (n=89)	Ethiopia (n=57)	India (n=137)	Kenya (n=110)	Mali (n=101)	Mozambique n=110)	Nigeria (107)	South Africa (105)	Overall (n=816)
<i>alp1</i>	29 (32.6%)	28 (49.1%)	41 (29.9%)	37 (33.6%)	29 (28.7%)	24 (21.8%)	30 (28.0%)	26 (24.8)	244 (29.9%)
<i>alp2/3</i>	18 (20.2%)	7 (12.3%)	83 (60.6%)	14 (12.7%)	17 (16.8%)	12 (10.9%)	23 (21.5%)	9 (8.6%)	183 (22.4%)
<i>alphaC</i>	22 (24.7%)	9 (15.8%)	14 (10.2%)	19 (17.3%)	20 (19.8%)	22 (20%)	22 (20.6%)	30 (28.6%)	158 (19.36%)
<i>rib</i>	3 (3.4%)	10 (17.5%)	3 (2.2%)	41 (37.3%)	19 (18.8%)	52 (47.3%)	30 (28.1%)	43 (40.95%)	201 (24.6%)
Alpha family	72 (80.9%)	54 (94.7%)	134 (97.8%)	110 (100%)	80 (79.2%)	110 (100%)	105 (98.1%)	105 (100%)	770 (94.4%)
<i>hvgA</i>	0	4 (7.0%)	0	31 (28.2%)	9 (8.9%)	30 (27.3%)	6 (5.6%)	21 (20.0%)	101 (12.4%)
<i>srr1</i>	83 (93.3%)	53 (93.0%)	137 (100%)	75 (68.2%)	87 (86.1%)	76 (69.1%)	84 (78.5%)	74 (70.5%)	669 (82%)
<i>srr2</i>	4 (4.5%)	2 (3.5%)	0	31 (28.2%)	9 (8.9%)	34 (30.9%)	6 (5.6%)	24 (22.9%)	110 (13.5%)
Serine family proteins	87 (97.8%)	55 (96.5%)	137 (100%)	106 (96.4%)	96 (95%)	110 (100%)	90 (84.1%)	98 (93.3%)	779 (95.5%)
PI-1	10 (11.2%)	3 (5.3%)	55 (40.2%)	2 (1.8%)	16 (15.8%)	0	19 (17.8%)	0	105 (12.8%)
PI-2a	9 (10.1%)	10 (17.5%)	7 (5.1%)	33 (30%)	17 (16.8%)	44 (40%)	21 (19.6%)	38 (36.2%)	179 (21.9%)
PI-2b	12 (13.5%)	0	2 (1.5%)	0	8 (7.9%)	0	1 (0.9%)	0	23 (2.8%)

pilus 1 and 2a	39 (43.8%)	36 (63.2%)	34 (24.8%)	38 (34.5%)	34 (33.7%)	30 (27.3%)	59 (55.1%)	42 (40.0%)	312 (38.2%)
pilus 1 and 2b	15 (16.9%)	6 (10.5%)	39 (28.4%)	34 (30.9%)	28 (27.7%)	34 (30.9%)	7 (6.5%)	24 (22.86%)	187 (22.9%)
pilus family	85 (95.5%)	55 (96.5%)	136 (99.3%)	107 (97.3%)	101 (100%)	108 (98.2%)	107 (100%)	104 (99.1%)	803 (98.4%)
<i>gbs2106</i>	89 (100%)	57 (100%)	137 (100%)	110 (100%)	101 (100%)	110 (100%)	107 (100%)	105 (100%)	816 (100%)

The prevalence of serine-rich repeat family genes (*srr1* and *srr2*) was greater than 94.0% in all countries, except Nigeria (84.1%; 90/107). The serine-rich repeat protein genes were observed in 72.3% (86/119) of serotypes II isolates, albeit only in 34.6% (9/26) of serotype II isolates from Nigeria (Table 3.3). Among the two serine-rich repeat genes, the *srr2* gene was not detected in India, while the *srr1* gene was detected in all Indian isolates. Furthermore, the *srr1* gene was common in non-CC17 at each site (>68.5%) but low (1.0%) in CC17 strains (Fig 3.10). The putative protein *gbs2106* gene was detected in all GBS colonizing isolates.

Figure 3. 8: Prevalence of the alpha surface protein genes in CCs of colonizing GBS isolates from Southeast Asia and African countries.

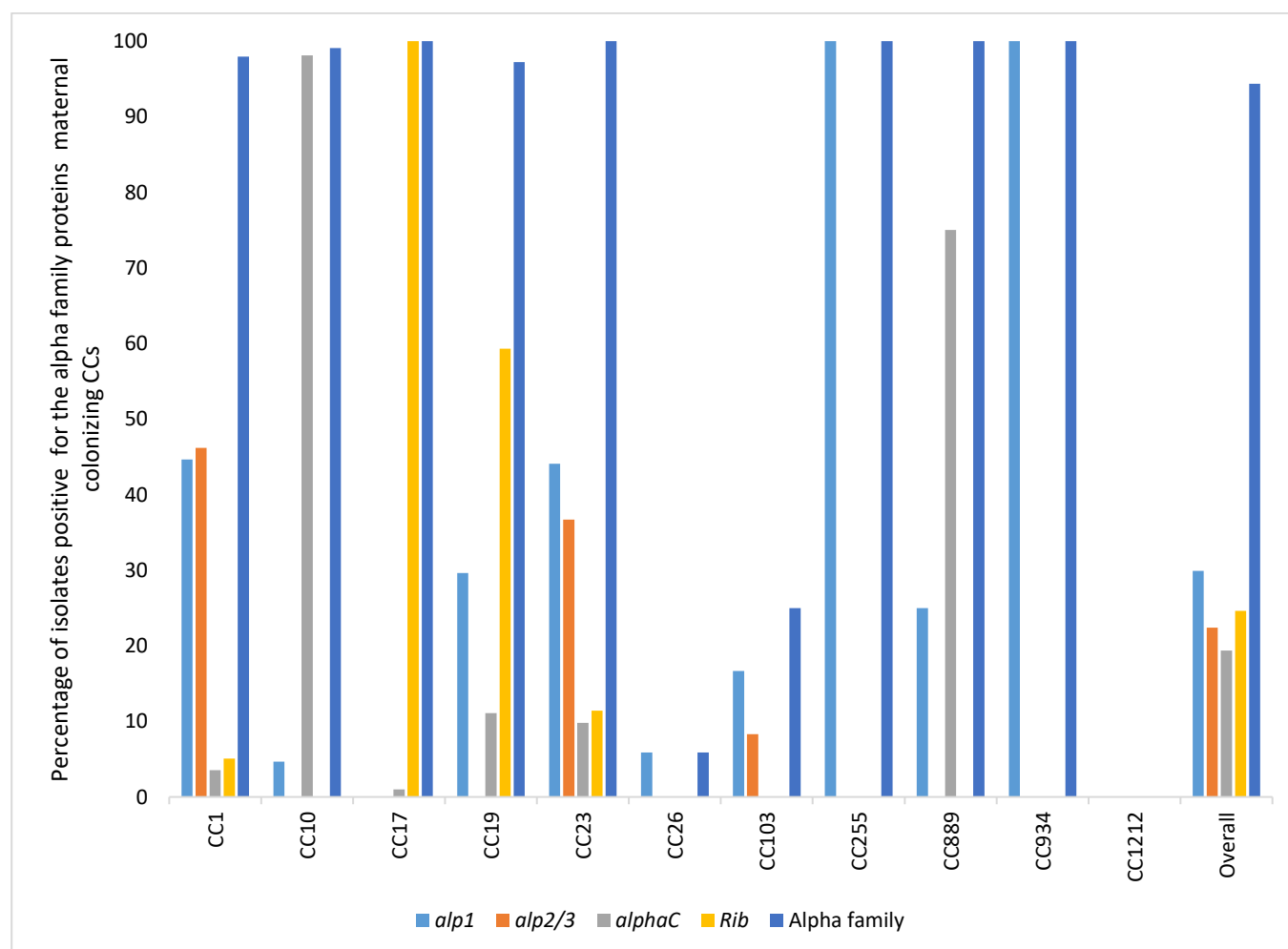


Table 3. 3: Surface proteins targets association with colonizing GBS serotypes in pregnant women from Southeast Asia and Africa.

	Ia (n=220)	Ib (n=64)	II (n=119)	III (n=142)	IV (n=32)	V (n=188)	VI (n=37)	VII (n=14)	Total (n=816)
<i>alp1</i>	122 (55.5%)	6 (9.4%)	20 (16.8%)	1 (0.7%)	19 (59.4%)	47 (25%)	29 (78.4%)	0	244 (29.9%)
<i>alp2/3</i>	70 (31.8%)	11 (17.2%)	3 (2.5%)	13 (9.2%)	1 (3.1%)	72 (38.3%)	0	13 (92.9%)	183 (22.4%)
<i>alphaC</i>	10 (4.6%)	46 (71.9%)	63 (52.9%)	2 (1.4%)	6 (18.8%)	25 (13.3%)	5 (13.5%)	1 (7.1%)	158 (19.4%)
<i>rib</i>	5 (2.3%)	2 (3.1%)	36 (30.3%)	122 (85.9%)	4 (12.5%)	29 (15.4%)	3 (8.1%)	0	201 (24.6%)
alpha family	202 (91.8%)	64 (100%)	117 (98.3%)	136 (95.8%)	30 (93.8%)	170 (90.4%)	37 (100%)	14 (100%)	770 (94.4%)
<i>hvgA</i>	2 (0.9%)	0	0	95 (66.9%)	4 (12.5%)	0	0	0	101 (12.4%)
<i>srr1</i>	219 (99.6%)	64 (100%)	86 (72.3%)	43 (30.3%)	20 (62.5%)	186 (98.9%)	37 (100%)	14 (100%)	669 (82%)
<i>srr2</i>	0	0	0	99 (69.7%)	11 (34.4%)	0	0	0	110 (13.5%)
serine rich repeat family	219 (99.6%)	64 (100%)	86 (72.3%)	142 (100%)	31 (96.9%)	186 (98.9%)	37 (100%)	14 (100%)	779 (95.5%)
PI-1	64 (29.1%)	13 (20.3%)	11 (9.2%)	8 (5.6%)	0	1 (0.5%)	8 (21.6%)	0	105 (12.9%)
PI-2a	117 (53.2%)	0	4 (3.4%)	0	0	56 (29.8%)	2 (5.4%)	0	179 (21.9%)
PI-2b	21 (9.6%)	0	0	2 (1.4%)	0	0	0	0	23 (2.8%)
pilus 1 and 2a	1 (0.5%)	49 (76.6%)	95 (79.8%)	35 (24.7%)	1 (3.1%)	130 (69.2%)	1 (2.7%)	0	312 (38.2%)
pilus 1 and 2b	14 (6.4%)	0	2 (1.7%)	98 (69%)	32 (100%)	0	27 (73%)	14 (100%)	187 (22.9%)
pilus family	217 (98.6%)	62 (96.9%)	112 (94.1%)	142 (100%)	32 (100%)	187 (99.5%)	37 (100%)	14 (100%)	803 (98.4%)

Prevalence of antimicrobial resistance genes in maternal colonizing GBS isolates in pregnant women from low- and middle-income countries.

The tetracycline resistance genes were the most common AMR genes overall (88.9%; 725/816), with prevalence ranging from 59.1% (81/137) in India to 100% (101/101) in Mali (Figure 3.11). The high prevalence of tetracycline resistance genes was primarily mediated by *tetM* (86.2%; 703/816) and *tetSM* (87.3%; 712/816) (Appendix 7). Within clonal complexes, tetracycline resistance genes varied from 58.3% (7/12) among CC889 strains to 100% (99/99) among CC17 strains in maternal isolates (Figure 3.12). Although the overall prevalence of tetracycline resistance was high in CC23 (76.7%; 188/245), the prevalence of tetracycline resistance genes was lower in CC23 isolates from India (23.1%; 15/65) and Bangladesh (66.7%; 4/6), compared to African countries where the prevalence in CC23 isolates ranged from 97.0% (32/33) in Nigeria to 100% (38/38) in South Africa, with the exception of Ethiopia (69.2%; 9/13).

The *aac6-aph2* gene conferring gentamicin resistance was detected in Nigeria (15.9%; 17/107) and Mali (3.0%; 3/101), with 95.5% (21/22) being CC19 isolates. Erythromycin resistance mediated by the *mefA* and *msrD* genes, was detected in 7.7% (63/816) of the isolates, including 31.5% (28/89) in Bangladesh, 20.9% (23/110) in Mozambique, whilst being present in <5% of isolates from Ethiopia, Kenya, Mali, Nigeria, and India. The prevalence of aminoglycoside resistance genes (*aph3-III* or *ant6-Ia*) was highest in women from Nigeria (16.8%; 18/107), followed by Mali (7.9%; 8/101) and least prevalent among women from Bangladesh (3.4%; 3/89). Chloramphenicol resistance genes (conferred by the *catpc194* and *catQ* genes) were more common in isolates from Bangladesh (9.0%; 8/89) compared with Nigeria (2.8%; 3/107) or India 2.9% (4/137). The macrolide-lincosamide-streptogramin B (MLS_B) resistance genes (*ermA*, *ermB*, *ermT*, *mefA*, *msrD*, *lsaC*, *lsaE*, *lnuB*, or *lnuC*) were detected in all countries, with prevalence ranging from 5.3% (3/57) in Ethiopia to 34.8% (31/89) in Bangladesh. Among the MLS_B genes, *ermA* was the most prevalent, detected in 7.7% (63/816) of isolates, mostly in CC19 (31.5%; 34/108), CC1 (12.8%; 25/195), while in other CCs the prevalence was low (<3%).

Figure 3. 9: The prevalence of antimicrobial resistance GBS isolates colonizing pregnant women from African countries and Southeast Asian countries.

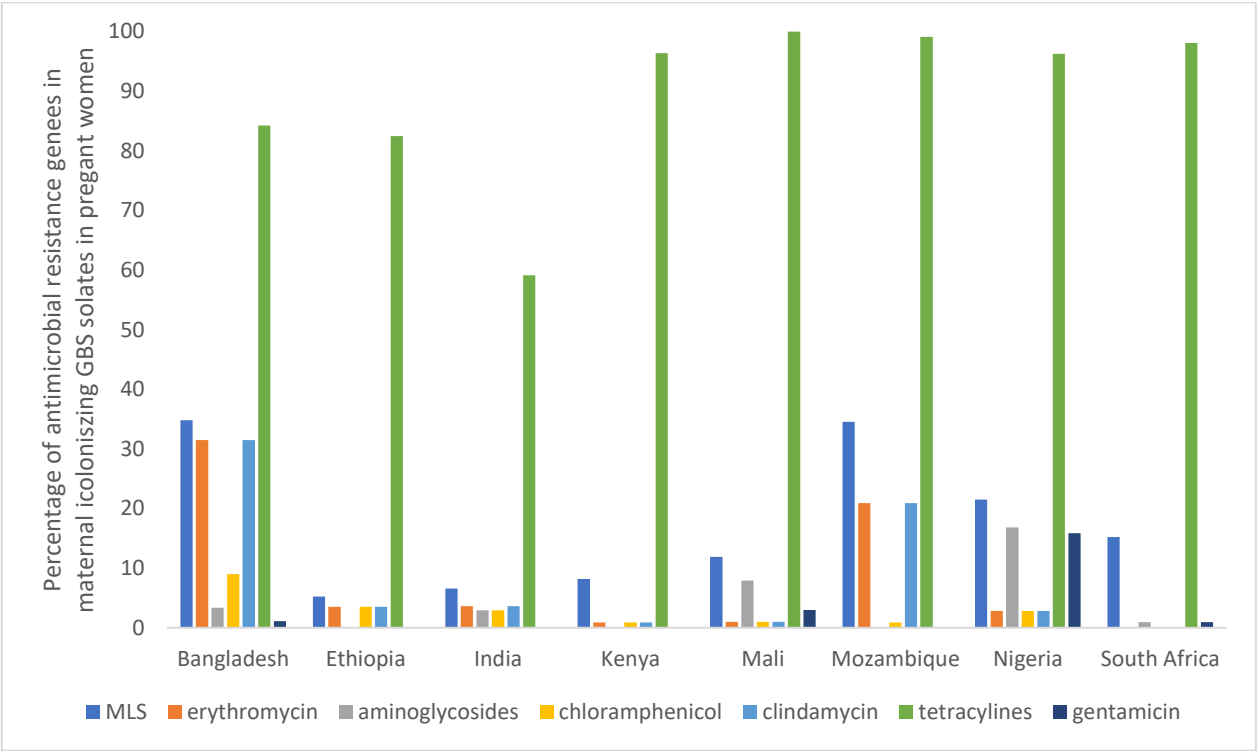
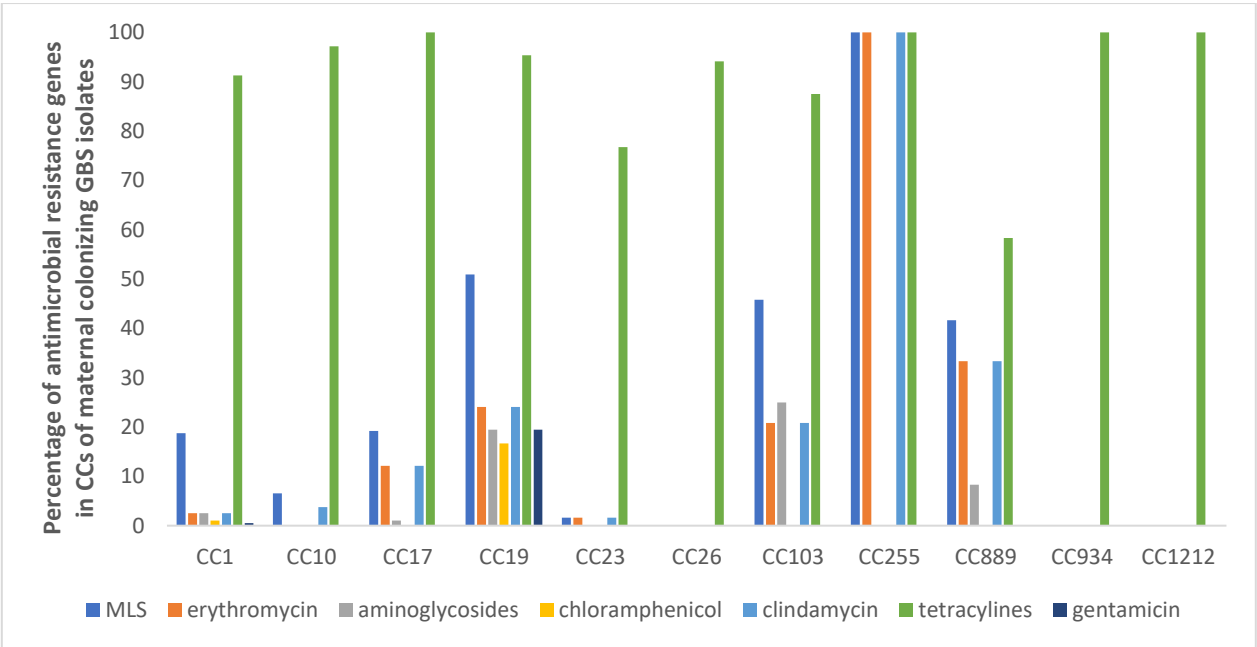


Figure 3. 10: The prevalence of antimicrobial resistance GBS CCs colonizing pregnant women from African countries and Southeast Asian countries.



3.4 Discussion

The findings from our study indicate heterogeneity, beyond the serotype distribution in the whole genome sequence of maternal GBS colonizing isolates across sub-Saharan African and Southeast Asian countries. The geographical differences in GBS genotype, and prevalence of virulence surface proteins such as *hvgA* gene, possibly contributes to the reported regional heterogeneity of higher incidence of early onset invasive GBS disease in sub-Saharan Africa compared with Southeast Asian countries (77).

Our findings illustrate that serotypes are associated with distinct CCs in different countries, which exhibits different profiles of surface protein antigen targets, antimicrobial resistance genes and invasive potential. In African countries, serotype III is associated with CC17, a lineage significantly associated with invasive GBS disease in infants (265). Conversely, in Southeast Asian countries with a low prevalence of serotype III, serotype III is associated with other CCs, such as CC23, typically associated with asymptomatic carriage and limited invasive potential (118). The low prevalence of serotype III and the absence of the *hvgA* gene in serotype III isolates from Southeast Asian countries, coupled with high prevalence of serotype III and the *hvgA* gene in isolates from African countries may partly account for the higher incidence of EOD in African compared with Southeast Asian countries.

Of the ten CCs identified in this study, five (CC1, CC10, CC19, CC17, and CC23) were previously identified as the most common CCs globally (261), four (CC1, CC10, CC19, and CC23) of which were detected across all countries, indicating global dissemination of these clones. Nevertheless, there were geographic restricted clonal complexes, including uncommon CCs (CC255, CC889 and CC1212) detected exclusively in Bangladesh, indicating the existence of unique GBS lineages within the country. The results could be due to the role of country-specific factors, including antibiotic use, host immune status and co-colonization by the vaginal microbiome, in shaping the genetic population structure of GBS strains within a particular geographical region. According to the PubMLST database, ST255 which belongs to the CC255 lineage, was previously detected in Europe, the United Kingdom and the Netherlands; while ST889 which is associated with CC889, was identified in Australia, Canada, and the United Kingdom(117).

Our study also identified the CC17-related pathogenic surface protein *hvgA* gene in non-CC17 isolates, specifically serotype Ia /CC934 from Ethiopia. The detection of the *hvgA* gene in serotype Ia/CC934 spanning different years of sample collection in the same geographic region, coupled

with the absence of recombination within the *hvgA* genomic region, suggest that GBS CC934 is a historically adapted clone circulating in Ethiopia. In contrast to the previously reported *hvgA* gene carrying isolates of serotype IV/CC17 due to capsular switching recombination within the capsular gene coding region (266, 267), there was no evidence of recombination within the capsular coding region of serotype Ia/CC9364 isolates. Our results show that serotype Ia/CC934 isolates represent an emerging hypervirulent GBS clone that acquired the *hvgA* gene from CC17 lineage. Currently the CC934 lineage is geographically restricted to Ethiopia, indicative of localized evolutionary processes. The dissemination and acquisition of genetic elements has been shown to promote the expansion of GBS clones. For example, the CC17 lineages have expanded through the acquisition of a novel phage with genes for coding for cell surface proteins (240).

The country-specific analysis allowed us to identify lineages with potential increased invasiveness and differing prevalence of protein antigen vaccine targets which could impact on the efficacy of vaccines currently under clinical development (70). A protein-alum adjuvant vaccine that incorporates the alpha family genes (*alp1*, *alp2/3*, *alphaC* and *rib*) would cover 94.0% of all maternal colonizing isolates. Nevertheless, the prevalence of alpha family genes was notably lower in Bangladesh (80.9%) and Mali (84.2%) compared with other sites (>94.7%), due to the presence of CC103 which was associated with lower prevalence of alpha family genes (40%). Furthermore, the relative presence of specific alp-like proteins varied between countries, with the prevalence of *rib* gene as an example ranging from 2.2% (3/137) in India to 47.3% (52/110) in Mozambique. The alpha protein family genes were detected in all CC17 and CC23 strains, although being absent in CC1212 isolates from Bangladesh. Our findings illustrate the importance of studying GBS sequences beyond the traditional capsular serotypes for predicting vaccine coverage in different countries.

Consistent with other studies, the *rib* gene that is part of the GBS-NN protein vaccine (260, 268) was present in all CC17 strains. Serotypes VI and VII, which are not covered by the hexavalent vaccine (269), carried at least one of the alp family genes and at least one of the pilus family genes. Our results also show that a protein adjuvant vaccine containing the pilus island protein genes (PI-1, PI-2a and PI-2b) would cover over 95.5% of maternal colonizing isolates in all countries. The lowest prevalence of pilus island protein genes was found in CC889 strains (83.3%). Other studies have suggested the loss of the pilus island gene PI-1 in CC17 along with the acquisition of mobile

genetic elements in the CC17 population (257). A lower antibody titer to the surface protein gene *gbs2106* has been positively associated with susceptibility to invasive GBS disease in HIV-infected individuals (263, 270). The surface protein gene *gbs2106* was identified in all GBS isolates suggesting a potential universal vaccine target that could offer protection against all serotypes detected in pregnant women from Southeast Asia and Africa.

The GBS isolates are well-known for their high prevalence of tetracycline resistance genes (123), which is corroborated by our study findings where the prevalence of the *tetM* gene was greater than 3% in all countries, except in India (59.1%). The acquisition of the *tetM* genes is hypothesized to have contributed to the emergence of human pathogenic GBS clones (123). Overall, a notable number of isolates with resistance genes to aminoglycosides, gentamicin, and MLS were identified in CC19 isolates. The gentamycin resistance genes were detected only in Nigeria (15.9%; 17/107) and Mali (2%; 3/101) and Bangladesh (1.1%; 1/89), suggesting possible adaptive mechanisms to local selection pressures by antibiotics.

The major strength of our study is the analysis of large samples of isolates from different LMICs, where there is little information on GBS sequence types. In addition, our study provides valuable insights into the genetic types of colonizing isolates in pregnant women, whose newborns are at high risk for invasive GBS disease. Limitations of the study include the unavailability of phenotypic data to corroborate *in silico* data on antimicrobial resistance patterns to determine the clinical relevance thereof, particularly since GBS remains susceptible to beta-lactams. Moreover, the study was a cross-sectionally designed, which hindered further analysis on genome changes over time.

In conclusion, our study provides information on the genome population structure of GBS colonizing isolates, including CCs associated virulence factors and vaccine target antigens including the *hvgA* and the alpha family proteins in Southeast Asian and African countries. The study revealed differences in the profile of CCs, surface protein antigen vaccine targets and AMR genes between different countries. The major finding of the study was the absence of *hvgA* gene in colonizing isolates from Southeast Asian countries compared with higher prevalence thereof in African countries. Therefore, the observed difference in the prevalence of virulence genes among GBS CCs, may partly account for the higher incidence of EOD in African compared with Southeast Asian countries (77) .

In the next chapter, the prevalence of colonization by more than one serotype was determined in pregnant women six study sites in South Africa, Kenya, Ethiopia, Bangladesh, Mali, and Nigeria.

Chapter 4 Use of microarray to determine prevalence of Group b streptococcus colonization by multiple serotypes in pregnant women from low- and middle-income countries

4.1 Abstract

Background: *Streptococcus* (GBS) causes stillbirths, preterm births, and invasive disease in newborns. There are ten GBS serotypes (Ia, Ib, and II-IX) classified based on their distinct capsular polysaccharide. Although the distribution of GBS serotypes in maternal colonization has been extensively studied, there is a paucity of data on the prevalence of colonization by multiple GBS serotypes.

Methods: A total of 733 vaginal swabs in which GBS was identified by culture method were randomly selected from six study sites in South Africa, Kenya, Ethiopia, Bangladesh, Mali, and Nigeria). The samples were analyzed for multiple GBS serotype colonization by microarray method.

Findings: Vaginal colonization by more than one GBS serotype was identified in 10% (72/702) of the women in whom GBS was identified by culture on vaginal swabs, including 93% (67/72) who were colonized by two serotypes. The highest prevalence of multiple GBS serotype colonization was observed in Mali (14.4%; n=20/139), while the lowest prevalence of multi-serotype colonization was in Bangladesh (3.7%; n=2/54). The most common combinations of serotypes associated with colonization by more than one serotype were GBS Ia and V (15.3%; n=11/72), followed by Ia and III (13.9%; n=10/72) and Ia and II (12.5%; n=9/72). Different GBS lineages of the same serotype were detected among nine (1.28%, 9/702) vaginal swabs.

Conclusion: Vaginal colonization by more than one GBS serotype is low in pregnant women.

4.2 Introduction

Group B *Streptococcus* (GBS; *Streptococcus agalactiae*) is a major cause of invasive disease in newborns globally, affecting 0.5-3 newborns per 1000 live births (4, 64). Early-onset invasive GBS disease (EOD; 0-7 days of life) caused by acquisition of GBS in the peripartum period from women colonized in the vagina or rectum (4, 271, 272). The prevalence of GBS recto-vaginal colonization

in pregnant women varies by geography, ranging between 0 to 40% (252). There are GBS serotypes (Ia, Ib, and II to IX) which are characterized based on the polysaccharide composition of the cell wall. GBS serotypes are identified based on the immunological reactivity of the surface capsular polysaccharides or detection of the capsular gene determinants (26). Six GBS serotypes, including Ia, Ib, II, III, IV and V constitute 98% of maternal colonizing serotypes and 93 to 99% of invasive GBS disease cases in sub-Saharan Africa (77, 272, 273).

A GBS hexavalent polysaccharide-protein conjugate vaccine is being developed for pregnant women, aimed at protecting their young offspring against invasive GBS disease (77, 274). Theoretically, vaccinating individuals with a serotype-specific capsular polysaccharide vaccine targeting only the dominant GBS serotypes, may result in replacement colonization and disease due to non-vaccine serotypes (84, 275, 276).

Several studies have investigated the distribution of GBS serotypes in maternal colonization, (252, 277) however, there is limited information on the frequency of concurrent colonization by multiple serotypes. The detection of colonization by multiple GBS serotypes is hampered by the traditional method of culture of GBS and serotyping by Latex agglutination method of only the dominant colony, likely impacts on the sensitivity to detect colonization by multiple serotypes. (83).

We used micro-array to evaluate the prevalence of GBS colonization by multiple serotypes in African women at the time of delivery of their offspring across 6 countries.

4.3 Results

GBS serotypes were detected by microarray in 95.8% (702/733) of the vaginal swabs that tested positive for GBS by PCR. The overall prevalence of colonization by more than one GBS serotype was 10.3% (72/702) in vaginal swabs in which GBS was detected. Colonization by more than one GBS serotype differed by geography, ranging from 3.7% (2/54) in Bangladesh to 14.4% (20/139) in Mali among GBS-positive women (Table 4.1). Of the 72 vaginal swabs with more than one GBS serotype, 93% (67/72), 5.6% (4/72) and 1.4% (1/72) had two, three and four serotypes identified, respectively. Overall, there were 150 individual serotypes identified among the 72 samples that were colonized with more than one serotype. In order of frequency, the serotypes identified in the 72 samples with more than one serotype were Ia 28% (42/150), followed by serotypes V 24% (36/150), II 16.7% (25/150), III 16% (24/150), Ib 7.3% (11/150), IV 7.3%, (11/150 and VI 0.7% (1/150).

Table 4. 1: Prevalence of multiple GBS serotypes in countries of low and middle-income countries.

Country	^a Prevalence of multiple serotypes in GBS-positive samples	^b Prevalence of GBS colonization in the parent study	^c Prevalence of multiple serotypes in the overall population
Kenya	9.7% (11/113; 95% CI:5.5-16.6)	17.7%	1.72%
Ethiopia	10.3% (7/68; 95% CI:5-19.8)	10%	1.03%
Mali	14.4% (20/139; 95% CI: 9.5-21.2)	36.9%	5.31%
Bangladesh	3.7% (2/54; 95% CI: 1-12.5)	12.6%	0.47%
Nigeria	11.7% (12/103; 95% CI:6.8-19.2)	19.5%	2.28%
South Africa	8.9% (20/225; 95% CI: 5.8-13.3)	28%	2.49%
Overall	10.3% (72/702; 95% CI: 8.2-13)	21%	2.2%

^a Prevalence of multiple serotypes detected in GBS positive samples in this study using microarray method and data are presented as % (n/N; 95% Confidence interval)

^b Prevalence of GBS colonization observed in the parent GBS prevalence study in which vaginal swabs were tested for GBS using culture methods (222).

^c Prevalence of multiple serotypes in women was calculated as a product of: ^aprevalence of multiple serotypes in GBS-positive samples multiplied by the ^bprevalence of GBS colonization in the parent study.

The most common serotype combination of concurrent serotypes in with more than one serotype were Ia and V (15.3%, 11/72), Ia and III (13.9%, 10/72) and Ia and II (12.5%, 9/72) (Figure 4.1). Nine (1.28%, 9/702) vaginal swabs were identified to being colonized by different strains of the same serotype, including serotype Ia (n=2), Ib (n=2), III (n=3) and V (n=2). The relative abundance

of dominant serotypes in samples with more than one serotype ranged from 51% to 99%, while the relative density of minor serotypes ranged from 1% to 47% (Figure 4.2). The ratio of serotype abundance between the most abundant and less abundant serotypes in samples with more than one serotype varied, ranging from 99:1 to 1.2:1 (Table 4.2).

Figure 4. 1: Combinations of serotypes in vaginal swabs colonized with more than one serotype.

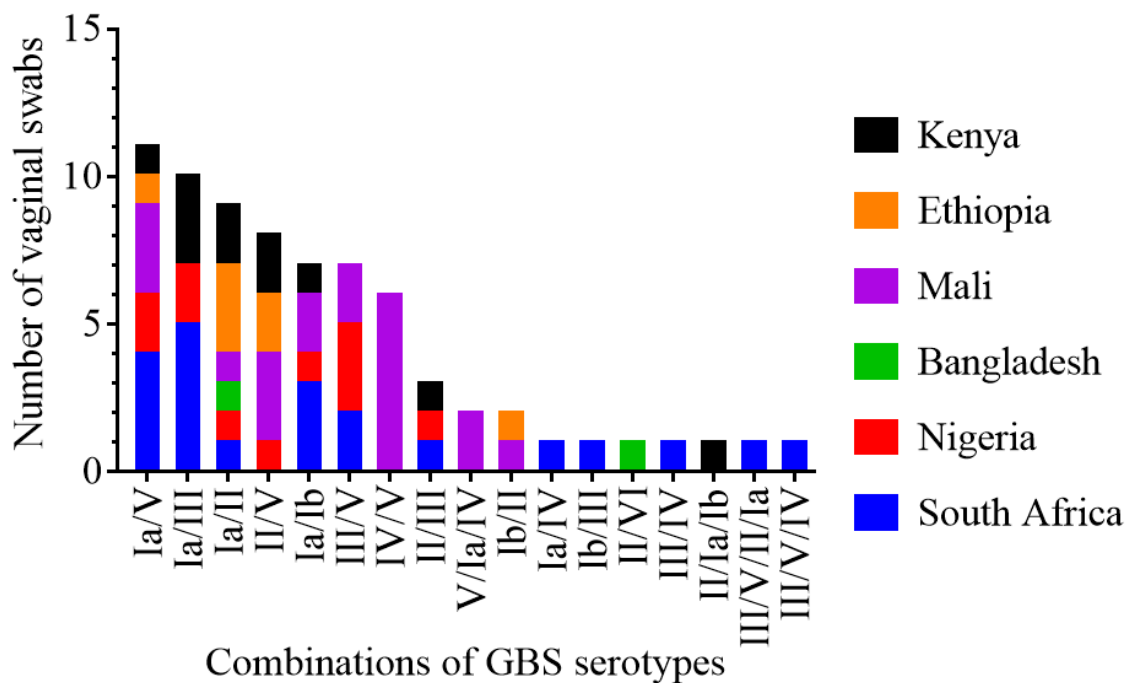


Figure 4. 2: Difference in relative abundance of serotypes among samples colonized with more than one serotype. The serotypes were arranged according to their relative abundance in samples with serotype 1 having the highest relative abundance, followed by serotype 2, serotype 3 or serotype 4.

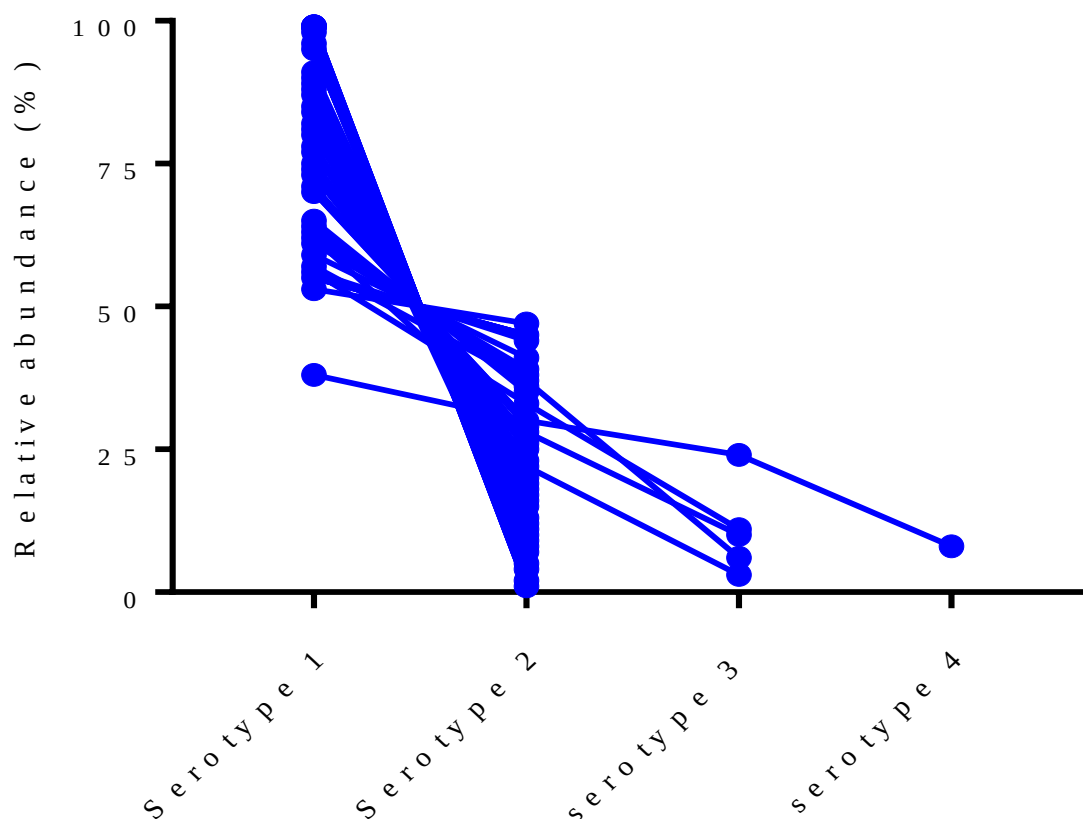


Table 4. 2: Serotypes in samples colonized with more than one serotype.

Sample	Array serotypes	Site	Total number of serotypes detected
2-1-2858	VI (75%) + II (25%)	Bangladesh	2
8-1-8079	III (55%) + V (45%)	Nigeria	2
11-1-11323	II (95%) + V (5%)	Mali	2
11-1-11805	V (99%) + Ia (1%)	Mali	2

11-1-11146	IV (89%) + V (11%)	Mali	2
11-1-11915	IV (73%) + V (27%)	Mali	2
3-1-3811	V (88%) + II (12%)	Ethiopia	2
3-1-3867	V (95%) + II (5%)	Ethiopia	2
1-1-1737	V (84%) + Ia (16%)	Kenya	2
1-1-1743	V (96%) + II (4%)	Kenya	2
1-1-1993	V (87%) + II (13%)	Kenya	2
11-1-11080	V (85%) + IV (15%)	Mali	2
11-1-11183	V (73%) + IV (27%)	Mali	2
11-1-11240	V (98%) + IV (2%)	Mali	2
11-1-11312	IV (59%) + V (41%)	Mali	2
11-1-11322	V (99%) + Ia (1%)	Mali	2
11-1-11334	V (75%) + Ia (22%) + IV (3%)	Mali	3
11-1-11340	III (70%) + V (30%)	Mali	2
11-1-11455	V (62%) + IV (28%) + Ia (10%)	Mali	3
7-1-7319	Ia (96%) + V (4%)	Nigeria	2
7-1-7587	V (82%) + III (18%)	Nigeria	2
12-1-12800	V (90%) + III (8%)	South Africa	2
12-1-12901	V (95%) + Ia (5%)	South Africa	2
12-1-12029	IV (81%) + III (19%)	South Africa	2
12-1-12516	III (73%) + II (27%)	South Africa	2
11-1-11025	II (89%) + V (11%)	Mali	2
11-1-11502	V (70%) + III (30%)	Mali	2
12-1-12215	III (96%) + Ia (4%)	South Africa	2

12-1-12218	IV (95%) + Ia (5%)	South Africa	2
12-1-12318	III (65%) + Ia (35%)	South Africa	2
12-1-12400	III (78%) + Ia (22%)	South Africa	2
12-1-12821	III (96%) + Ib (4%)	South Africa	2
12-1-12896	III (84%) + Ia (16%)	South Africa	2
12-1-12941	III (57%) + V (37%) + IV (6%)	South Africa	3
3-1-3378	Ia (81%) + II (19%)	Ethiopia	2
1-1-1487	II (56%) + Ia (33%) + Ib (11%)	Kenya	3
1-1-1945	II (75%) + III (25%)	Kenya	2
11-1-11136	II (98%) + V (2%)	Mali	2
11-1-11305	II (91%) + Ia (9%)	Mali	2
11-1-11388	II (99%) + Ib (1%)	Mali	2
12-1-12158	Ia (90%) + II (10%)	South Africa	2
11-1-11368	Ib (61%) + Ia (39%)	Mali	2
11-1-11507	Ib (55%) + Ia (45%)	Mali	2
8-1-8709	Ib (96%) + Ia (4%)	Nigeria	2
12-1-12090	Ia (71%) + Ib (29%)	South Africa	2
2-1-2414	II (80%) + Ia (20%)	Bangladesh	2
3-1-3240	Ib (90%) + II (10%)	Ethiopia	2
3-1-3734	Ia (87%) + II (13%)	Ethiopia	2
3-1-3962	Ia (75%) + II (25%)	Ethiopia	2
1-1-1115	Ia (62%) + Ib (38%)	Kenya	2
1-1-1532	Ia (98%) + II (2%)	Kenya	2
1-1-1640	Ia (77%) + III (23%)	Kenya	2

11-1-11717	V (56%) + Ia (44%)	Mali	2
8-1-8361	Ia (99%) + II (1%)	Nigeria	2
12-1-12073	Ia (87%) + V (7%)	South Africa	2
12-1-12161	Ia (95%) + Ib (5%)	South Africa	2
12-1-12210	III (85%) + Ia (15%)	South Africa	2
12-1-12237	Ia (55%) + V (45%)	South Africa	2
12-1-12297	Ib (53%) + Ia (47%)	South Africa	2
12-1-12500	V (81%) + III (19%)	South Africa	2
12-1-12066	Ia (63%) + V (37%)	South Africa	2
3-1-3980	Ia (55%) + V (45%)	Ethiopia	2
8-1-8140	Ia (55%) + III (45%)	Nigeria	2
1-1-1231	III (64%) + Ia (36%)	Kenya	2
1-1-1342	Ia (74%) + III (26%)	Kenya	2
8-1-8489	V (82%) + Ia (17%)	Nigeria	2
7-1-7238	V (98%) + III (2%)	Nigeria	2
7-1-7183	Ia (82%) + III (18%)	Nigeria	2
1-1-1475	II (99%) + Ia (1%)	Kenya	2
7-1-7190	III (38%) + V (30%) + II (24%) + Ia (8%)	Nigeria	4
7-1-7420	II (98%) + V (2%)	Nigeria	2
7-1-7894	II (98%) + III (2%)	Nigeria	2

4.4 Discussion

The main finding of this study is that colonization by more than one GBS serotype in pregnant women of low and middle-income countries is low, but present in 2% of women. The prevalence of colonization by more than one serotype varied geographically, with the lowest prevalence being 0.47% in Bangladesh and the highest in Malian women (5.3%). Furthermore, almost all serotypes (Ia, Ib, II to V) detected in women colonized by more than one serotype are covered by the current hexavalent polysaccharide-protein conjugate vaccine (274). Although the frequency of colonization with more than one serotype is low, the detection of multiple serotypes both before and after GBS vaccine licensure is crucial to understanding the future impact of GBS vaccines.

The major strength of our study is that we report population-based estimates of colonization by more than one serotype, rather than only the prevalence of multiple serotypes in GBS colonized samples. Based on the prevalence of GBS colonization from the parent GBS study, from which our study samples were collected, we estimated the overall prevalence of colonization by more than one serotype in pregnant women to be $\leq 5.3\%$. The observed prevalence of multiple serotypes is similar to estimates from Australia (2.5%) and Kenya (2.4%) albeit being higher than estimates from Ghana (0.25%), detected using culture and PCR methods (96, 278). Similar to maternal GBS colonization rates, which vary by region (251, 252), there were notable country-specific differences in colonization by more than one serotype, reflecting differences in the risk posed by multiple serotypes in different geographical regions. The prevalence of multiple serotypes in pregnant women from Kenya was 1.7%, slightly lower than a previous report of 2.4% in the same country using culture and latex agglutination methods on recto-vaginal swabs (96). The use of vaginal swabs rather than recto-vaginal swabs may have contributed to the low prevalence of colonization by more than one serotype observed in our study.

The overall prevalence of multiple serotypes observed in GBS-positive vaginal swabs in this study was comparable to the results of a recent systematic review which reported 10% (95 CI: 4-19) in GBS-positive samples taken from the same anatomical site (279). In contrast to our findings, Khatami *et al* observed a lower prevalence (6.6%) of multiple serotypes in GBS-positive swabs in the USA using the PCR method (83). Recently, in the Gambia, multiple serotypes were detected in 32% of rectovaginal swabs colonized with GBS using the RAPD PCR method. With the increase in the use of molecular methods to detect GBS serotypes, multiple GBS serotypes may be reported

more frequently, as conventional culture and latex methods often overlook colonization by more than one serotype by not testing multiple GBS colonies.

The most common serotypes in vaginal swabs colonized by more than one serotype were GBS III, V and Ia, which are the most prevalent maternal colonizing serotypes globally (252). A recent study also identified serotypes III, V and Ia as predominant serotypes in vaginal swabs colonized by more than one serotype (279). Of note, almost all serotypes detected in vaginal swabs colonized by more than one serotype, including serotypes Ia, Ib and II-V, are covered by the current GBS polysaccharide conjugate vaccines (51, 280, 281). Therefore, it is highly unlikely that serotype replacement will impact GBS vaccinations as the serotype in multiple serotype carriage with the potential for serotype switching and evasion of vaccine immunity are covered in the current GBS vaccines.

Our results show the utility of microarray in detecting multiple serotypes in vaginal swabs, up to 4 serotypes were detected in one (0.14%) sample and three serotypes in four (0.57%) samples. Microarray has been previously used to study for nasopharyngeal colonization by multiple *Streptococcus pneumoniae* serotypes, and was shown to be superior compared with the standard culture methods in detecting colonization by multiple serotypes (48.8% vs 11.2%) (282-284). The microarray is, however, costly to set up, but suitable for detecting minor serotypes carried at low density which are unlikely to be detected by the standard culture method (285).

In samples with multiple serotypes, colonization by two serotypes was more common than with three or four serotypes, which supports the notion that colonization by one or two serotypes reduces the chances of colonization by other serotypes (278, 279). Other studies have shown that serotypes differ in their competitiveness for colonization, with serotypes having lower metabolic needs tending to dominate the niche (245, 286). Although there is evidence that GBS strains could undergo capsular switching (287, 288), which theoretically may lead to serotype replacement, the frequency of multiple serotypes required to initiate capsular serotype switching and serotype replacement was low in this study. Nine (1.28%, 9/702) vaginal swabs were colonized by two strains of the same serotype, indicating that GBS strains are diverse and expanding further into different lineages.

The major limitations of this study were small sample size at some study sites and the cross-sectional design of the study. Given the dynamic and transient nature of vaginal GBS colonization, longitudinal studies are warranted to determine whether GBS colonization by more than one

serotype is a short-term or a long-term phenomenon. In conclusion, our results suggest that the prevalence of multiple serotypes in pregnant women is low, and it is highly unlikely that GBS vaccines will be hampered by capsular switching and serotype replacement. Continued monitoring and surveillance of multiple serotypes and their distribution is crucial to determine the impact of future GBS vaccines.

The next chapter will focus on the association between the vaginal Group B *Streptococcus* colonization and the vaginal microbiome in pregnant women as there is paucity of data on the role of the vaginal microbiome in vaginal GBS colonization.

Chapter 5 Association between vaginal Group B *Streptococcus* colonization and the vaginal microbiome in pregnant women from low and middle-income countries of sub-Saharan Africa

5.1 Abstract

Background. Group B *Streptococcus* (GBS; *Streptococcus agalactiae*) vaginal colonization in pregnant women is a major risk factor for their newborns to develop invasive GBS disease. The vaginal microbiome might play a role in influencing vaginal colonization by GBS.

Methods: We investigated the relationship between the vaginal microbiome and GBS colonisation in 282 pregnant women from Sub-Saharan African countries (Ethiopia, Kenya and South Africa), using an open array quantitative polymerase chain reaction (qPCR) method targeting 52 organisms.

Finding Overall, the most common organisms in women colonised by GBS were *Veillonella parvula* (87.2%; n=123/141), *Lactobacillus inners* (82.3%; n=116/141), *Ureaplasma urealyticum* (70.2%; n=99/141), and *Atopobium vaginae* (60.3%; n=85/141). Conversely, the most common organisms in women without GBS colonisation were *Lactobacillus inners* (73.8%; n=104/14), *Veillonella parvula* (71.6%; n=101/141), *Ureaplasma urealyticum* (59.6%; n=84/141), and *Gardnerella vaginalis* (59.6%; n=84/141). Three organisms, *Veillonella parvula* (adjusted odds ratio, 2.89; 95% confidence interval, 1.53-5.43), *Staphylococcus aureus* (aOR, 12.3; 95% CI, 5.31-28.24) and *Candida albicans* (aOR, 3.98; 95% CI, 2.33-6.79) showed positive association with vaginal GBS colonisation. Conversely, there was no inverse association with GBS colonization. The vaginal microbiome dominated by *Lactobacillus crispatus* (CST I, 3.9%; n=11/282), *Lactobacillus iners* (CST III, 7.1%; n=21/282), anaerobic bacteria (CST IV, 86.5%; n=246/282) and *Lactobacillus jensenii* (CST V, 2.1%; n=5/282) were identified, with CST V being detected only in women not colonized with GBS.

Interpretation: Vaginal GBS colonization is influenced by the presence or absence of other bacterial organisms colonizing the vagina in pregnant women, which suggest that interventions directed toward the vaginal microbiome may reduce the risk of invasive GBS disease in women and their newborns.

5.2 Introduction

The gram-positive bacterium Group B *Streptococcus* (GBS) causes invasive disease in 1 to 3 newborns per 1000 live births, manifesting as bacteremia, sepsis, pneumonia, and meningitis (4). The major risk factor for invasive GBS disease during the first week of life (i.e. early onset) is recto-vaginal colonization with GBS in pregnant women at the time of labor. Group B *Streptococcus* colonization in pregnant women is most commonly asymptomatic affecting 10 to 35% of the women (251, 252). Nevertheless, recto-vaginal GBS colonization in pregnant women, in addition to being a risk factor for early-onset invasive disease (EOD, 0-6 days of age) in newborns, has also been associated with premature labor and stillbirths (139, 289).

There is a paucity of data on the vaginal microbiome in relation to GBS colonization. A few studies report a relative reduction in vaginal *Lactobacillus* species among women colonized with GBS (204, 205), whereas no such association was observed by others (207, 290). A study in non-pregnant women reported that vaginal GBS colonization was positively associated with specific taxa such as *Prevotella bivia* and genus *Veillonella*, whilst being negatively associated with *Megasphaera* 1 and bacterial vaginosis-associated bacteria (BVAB) 1 (200). Most of the studies to date have been conducted in high-income countries in Europe and America (200, 247, 290), with a few from low and middle-income countries (194). Furthermore, the sequencing methods commonly used to identify vaginal organisms are less sensitive in detecting low density organisms that could potential influence the vaginal microbiome, conversely the open array assays allows the determination of absolute concentrations of target organisms in study samples (221).

In a study by Ravel *et al.* (179), the vaginal microbiome was classified into five distinct community state types (CSTs); four of which are dominated by *Lactobacillus* species, including *Lactobacillus crispatus* (CST I), *Lactobacillus gasseri* (CST II), *Lactobacillus inners* (CST III), and *Lactobacillus jensenii* (CST V). In contrast, the CST IV group is composed of diverse organisms that are characterized by significant proportions of strict and facultative anaerobic bacteria, but with a paucity of *Lactobacillus* species (179). A diverse vaginal microbiome (CST IV) has previously been associated with an increased risk of HIV acquisition and transmission (291, 292); however, the effect thereof on vaginal GBS colonization has not been ascertained.

The primary objective of this study was to investigate the association between vaginal GBS colonization and the vaginal microbiome in African women at the time of delivery of their

offspring; and to identify the specific vaginal organisms positively or negatively associated with vaginal GBS colonization.

5.3 Results

Characteristics of study participants

A total of 282 vaginal swabs from pregnant women were analyzed, 99% (278/282) were of black African descent. The median age of the women was 24 years [interquartile range (IQR): 21-27] in Ethiopia, 24 years [IQR: 21-29] in Kenya, and 26 years [IQR: 21-32] in South Africa). The median age, pregnancy history and body mass index were similar between women colonized with GBS (n=141) and those not colonized with GBS (n=141); Table 5.1. The country-stratified analysis between women with and without GBS colonization revealed some differences in demographics (Table 5.2), hence, country was included in the multivariable logistic regression model.

Table 5. 1: Demographic and clinical characteristics of the study population

Characteristics	Total participants (n=282)	GBS vaginal colonized (n=141)	No GBS vaginal colonization (n=141)	P-value
Age in years	24 (21-29)	25 (22-30)	24 (21-28)	0.17
Ethnicity (Black)	278 (99%)	139 (98.6%)	139 (98.5%)	1
Number of previous pregnancies	1 (0-2)	1 (0-2)	1 (0-2)	0.82
Previous spontaneous abortion	30/162 (18.5%)	18/85 (21.2%)	12/77 (15.6%)	0.36
Complications during pregnancy, labor or delivery	98 (34.8%)	47 (33.3%)	51 (36.2%)	0.62
Vaginal douche during pregnancy	20/274 (7.3%)	10/137 (7.3%)	10 /137 (7.3%)	0.99

Diet (mostly carbohydrates)	165/280 (58.9%)	84/140 (60.0%)	81/140 (57.9%)	0.72
Body mass index (kg/m ²) ^a	25.3 (22.9-28.9)	25.2 (23 -29.4)	25.4 (22.8-28.1)	0.52
Co-morbidities during pregnancy	55/280 (19.6%)	27/140 (19.3%)	28/140 (20%)	0.88
Evidence of foetal distress	55/279 (19.7%)	27/141 (19.2%)	28 /138 (20.3%)	0.81
Water source (indoor tap)	167/281 (59.4%)	86 (61.4%)	81 (57.5%)	0.46
Household contact with livestock	42/280 (15.0%)	22/139 (15.8%)	20/141 (14.2%)	0.70

Data are proportions of column total (%) or median (interquartile range). ^aThe body mass index (BMI) is the weight in kilograms divided by the square of the height in meters.

Table 5. 2: Site specific analyses of clinical and demographic characteristics of participants in Ethiopia, Kenya, and South Africa

	Ethiopia			Kenya			South Africa		
Characteristics	GBS colonized (n=47) ^a	No GBS colonization (n=47) ^b	P-value	GBS colonized (n=47) ^c	No GBS colonization (n=47) ^d	P-value	GBS colonized (n=47) ^e	No GBS colonization (n=47) ^f	P-value
Median age in years (IQR)	24 (20-28)	23 (21-26)	0.631	24 (21-28)	24 (21-30)	0.208	28 (23-33)	23 (20-29)	0.005
Ethnicity (Black)	44 (93.6%)	41 (87.2%)	0.30	47 (100%)	47 (100%)	1	45 (95.7%)	45 (95.7%)	1
Prior pregnancy	18 (38.3%)	20/46 (43.5%)	0.612	31 (66%)	34 (72.3%)	0.503	36 (76.6%)	23 (48.9%)	0.007
Median number of prior pregnancies (IQR)	0 (0-1)	0 (0-1)	0.576	1 (0-2)	1 (0-3)	0.036	1 (1-3)	0 (0-2)	0.005
Previous intrauterine deaths/stillbirths	1/18 (5.6%)	0	0	0	2 (5.9%)	0	0	0	0
Previous spontaneous abortion	1/20 (20%)	1/20 (20%)	0.939	7 (22.6%)	5 (14.7%)	0.417	10 (27.8%)	6 (26.1%)	0.887
Complications during current pregnancy, labor or delivery	14 (29.8%)	18 (38.3%)	0.385	13 (27.7%)	13 (27.7%)	1	20 (42.6%)	20 (42.6%)	1
1. Episiotomy	13 (27.7%)	17 (36.2%)	0.377	6 (12.8%)	5 (10.6%)	0.32	8 (17%)	12 (25.5%)	0.316

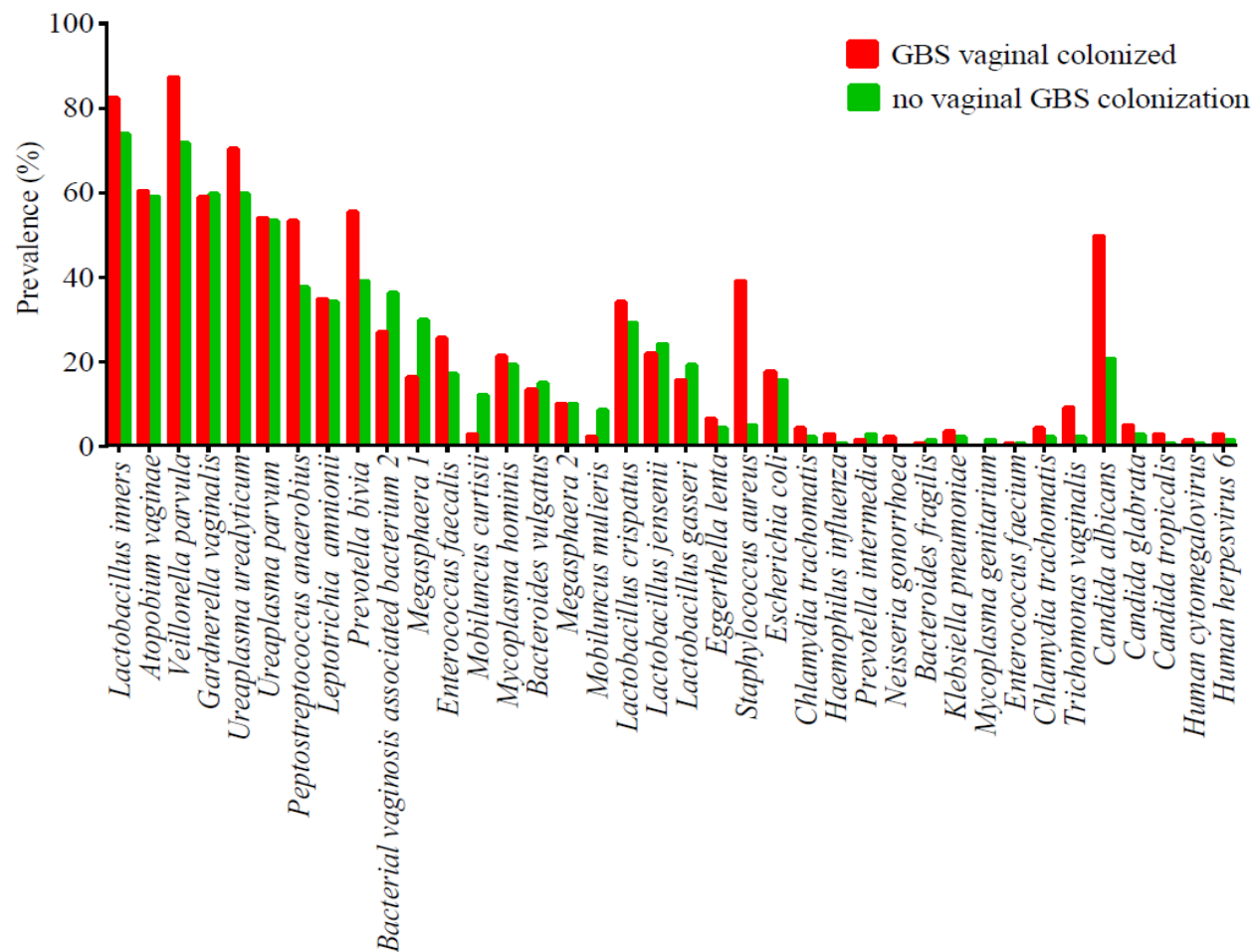
2 Perineal laceration/tear	1 (2.1%)	1 (2.1%)	1	8 (17%)	7 (14.9%)	0.778	9 (19.2%)	6 (12.8%)	0.778
Vaginal douche during pregnancy	1 (2.1%)	0	0	38 (80.9%)	37 (78.7%)	0.797	0	0	0
Diet (mostly carbohydrates)	8 (17.4%)	7 (15.2%)	0.778	47 (100%)	47 (100%)	1	29 (61.7%)	27 (57.45%)	0.674
Mid-upper arm circumference (cm)	25 (22-25)	23.7 (22-25)	0.346	25 (24-27)	26 (24-28)	0.569	30 (26-33)	28 (25-33)	0.399
Median body mass (kg/m ²) (IQR)	25 (24-27)	25 (24-27)	0.429	25 (23-27)	26 (23-29)	0.239	28.5 (22.2-33.5)	25.1 (21.1-32.3)	0.288
Co-morbidities during pregnancy	1/46 (2.2%)	3/46 (6.5%)	0.33	17 (36.17)	18 (38.3%)	0.831	9 (19.2%)	7 (14.9%)	0.584
Evidence of foetal distress	0	2/44(4.6%)	0	14 (29.8%)	13 (27.7%)	0.82	13 (27.7%)	13 (27.7%)	1
Water source (indoor tap)	40/46 (87%)	36 (76.6%)	0.202	14 (29.8%)	12 (25.5%)	0.512	15 (31.9%)	14 (29.8%)	0.823
Household have contact with livestock	0	0	0	21 (44.7%)	20 (42.6%)	0.835	1 (2.1%)	0	0

a,b,c,d,e,f Data are column percentage. IQR: Inter quartile range. The p-values were calculated by univariate logistic regression.

Prevalence and association of vaginal organisms with GBS colonization

Of the 52 organisms targeted by the qPCR panel, 39 (75%) were detected in at least one vaginal swab sample. Overall, the most common organisms detected in vaginal swabs were *Veillonella parvula* (79.4%; n=224/282), *Lactobacillus inners* (78%; n=220/282), *Ureaplasma urealyticum* (64.9%; n=183/282), and *Atopobium vaginae* (59.6%; n=168/282). Bacterial organisms were the most common organisms in pregnant women, whereas fungi and viruses were uncommon. Viruses detected included human herpesvirus (2.1%; n=6/282) and human cytomegalovirus (1.1%; n=3/282) (Table 5.3). Stratified by GBS colonization status, the most common organisms in GBS-colonized women were *Veillonella parvula* (87.2%; n=123/141), *Lactobacillus inners* (82.3%; n=116/141), *Ureaplasma urealyticum* (70.2%; n=99/141), *Atopobium vaginae* (60.3%; n=85/141), *Prevotella bivia* (55.3%; n=78/141), and *Ureaplasma parvum* (53.9%; n=76/141). The most common organisms in women without GBS colonization were *Lactobacillus inners* (73.8%; n=104/141), *Veillonella parvula* (71.6%; n=101/141), *Ureaplasma urealyticum* (59.6%; n=84/141) and *Gardnerella vaginalis* (59.6%; n=84/141) (Figure 5.1).

Figure 5. 1: The prevalence of vaginal organisms among GBS colonized women and those without GBS colonization. Vaginal organisms included bacteria, protozoa (*Trichomonas vaginalis*), fungi (*Candida* species) and viruses (Human cytomegalovirus and Human herpesviruses).



Organisms that were more common in women colonized by GBS compared with GBS-uncolonized women included *Veillonella parvula* (87% vs 72%; aOR, 2.89; 95% CI, 1.53-5.43), *Staphylococcus aureus* (39% vs. 5%; aOR, 12.3; 95% CI, 5.31-28.24) and *Candida albicans* (50% vs. 21%; aOR, 3.98; 95% CI, 2.33-6.79) after adjustment for country of sampling and age of participants. Also positively associated with GBS colonization, albeit not significant after Bonferroni correction (i.e p-value >.0013) for multiplicity were *Peptostreptococcus anaerobius* ((53% vs 38%; aOR, 2.04; 95% CI, 1.24-3.35; p=0.005), *Prevotella bivia* (55% vs 39%; aOR, 2.01; 95% CI, 1.24-3.28), *Ureaplasma urealyticum* (70% vs 60%; aOR; 1.7, 95% CI, 1.02-2.81; p=0.040), and *Trichomonas vaginalis* (9% vs 2%; aOR; 4.53; 95% CI, 1.26-16.32; p=0.021).

Organism less common in women colonized with GBS compared with GBS uncolonized women included *Megasphaera 1* (16s % vs 30%; aOR, 0.41; 95% CI, 0.22-0.75; p=0.004), *Mobiluncus curtisii* (3% vs 12%; aOR, 0.21; 95% CI, 0.06-0.66; p=0.007) and *Mobiluncus mulieris* (2% vs 9%; aOR, 0.22; 95% CI, 0.05-0.82; p=0.024) , albeit difference not being significant after Bonferroni correction for multiplicity (Table 5.3). Site-specific analysis showed consistency of findings to the overall analysis, particularly for *Staphylococcus aureus* being positively associated with GBS colonisation in Ethiopia (39% vs 5%; aOR, 21.7; 95% CI, 4.7-100.6; p= 0.0001), and Kenya (34% vs 4%; aOR, 9.8; 95% CI, 2.07-46.5; p =0.004), respectively (Table 5.4-5.6).

Table 5. 3: Association between vaginal organisms and GBS colonization status in pregnant women.

Target	Total (n=282) ^a	GBS colonized (n=141) ^b	No GBS colonisation (n=141) ^c	Odds ratio (95% CI)	p-value	Adjusted odds ratio (95% CI)	p-value
Bacterial organisms							
<i>Lactobacillus crispatus</i>	89 (31.6%)	48 (34%)	41 (29.1%)	1.26 (0.76-2.08)	0.37	1.30 (0.78-2.18)	0.31
<i>Lactobacillus gasseri</i>	49 (17.4%)	22 (15.6%)	27 (19.2%)	0.78 (0.42-1.45)	0.433	0.81 (0.43-1.52)	0.51
<i>Lactobacillus inners</i>	220 (78%)	116 (82.3%)	104 (73.8%)	1.65 (0.93-2.93)	0.086	1.68 (0.92-3.06)	0.088
<i>Lactobacillus jensenii</i>	65 (23.1%)	31 (22%)	34 (24.1%)	0.89 (0.51-1.54)	0.672	0.86 (0.49-1.50)	0.586
<i>Atopobium vaginae</i>	168 (59.6%)	85 (60.3%)	83 (58.9%)	1.06 (0.66-1.71)	0.808	1.08 (0.65-1.78)	0.766
<i>Bacteroides fragilis</i>	3 (1.1%)	1 (0.7%)	2 (1.4%)	0.50 (0.04-5.54)	0.569	0.5 (0.04-5.7)	0.579
<i>Bacteroides vulgatus</i>	40 (14.2%)	19 (13.5%)	21 (14.9%)	0.89 (0.46-1.74)	0.733	0.91 (0.46-1.78)	0.771
Bacterial vaginosis associated bacterium 2	89 (31.6%)	38 (27%)	51 (36.2%)	0.65 (0.39-1.08)	0.097	0.63 (0.37-1.05)	0.079
<i>Eggerthella lenta</i>	15 (5.3%)	9 (6.4%)	6 (4.3%)	1.53 (0.53-4.43)	0.429	1.50 (0.51-4.42)	0.459
<i>Gardnerella vaginalis</i>	167 (59.2%)	83 (58.9%)	84 (59.6%)	0.97 (0.6-1.56)	0.904	0.99 (0.61-1.65)	0.99
<i>Ureaplasma parvum</i>	151 (53.6%)	76 (53.9%)	75 (53.2%)	1.03 (0.64-1.64)	0.905	1.05 (0.65-1.68)	0.855
<i>Ureaplasma urealyticum</i>	183 (64.9%)	99 (70.2%)	84 (59.6%)	1.60 (0.98-2.62)	0.062	1.7 (1.02-2.81)	0.040
<i>Mobiluncus curtisii</i>	21 (7.5%)	4 (2.8%)	17 (12.1%)	0.21 (0.07-0.65)	0.007	0.21 (0.06-0.66)	0.007
<i>Mycoplasma hominis</i>	57 (20.2%)	30 (21.3%)	27 (19.2%)	1.14 (0.64-2.04)	0.657	1.12 (0.62-2.03)	0.699
<i>Mobiluncus mulieris</i>	15 (5.3%)	3 (2.1%)	12 (8.5%)	0.23 (0.06-0.85)	0.027	0.22 (0.05-0.82)	0.024
<i>Mycoplasma genitalium</i>	2 (0.71%)	0	2 (1.4%)	0			

Megasphaera 1	65 (23.1%)	23 (16.3%)	42 (29.8%)	0.46 (0.26-0.82)	0.008	0.41 (0.22-0.75)	0.004
Megasphaera 2	28 (9.9%)	14 (9.9%)	14 (9.9%)	1 (0.46-2.18)	1	0.90 (0.40-2.02)	0.799
Peptostreptococcus anaerobius	128 (45.4%)	75 (53.2%)	53 (37.6%)	1.89 (1.17-3.03)	0.009	2.04 (1.24-3.35)	0.005
Prevotella bivia	133 (47.2%)	78 (55.3%)	55 (39%)	1.94 (1.21-3.11)	0.006	2.01 (1.24-3.28)	0.005
Prevotella intermedia	6 (2.1%)	2 (1.4%)	4 (2.8%)	0.49 (0.09-2.73)	0.418	0.43 (0.07-2.45)	0.346
Leptotrichia amnionii	97 (34.4%)	49 (34.8%)	48 (34%)	1.03 (0.63-1.69)	0.9	1.03 (0.63-1.69)	0.923
Veillonella parvula	224 (79.4%)	123 (87.2%)	101 (71.6%)	2.71 (1.46-5.01)	0.002	2.89 (1.53-5.43)	0.0010
Chlamydia trachomatis	9 (3.2%)	6 (4.3%)	3 (2.1%)	2.04 (0.50-8.34)	0.319	2.27 (0.54-9.49)	0.263
Escherichia coli	47 (16.7%)	25 (17.7%)	22 (15.6%)	1.17 (0.62-2.18)	0.632	1.16 (0.61-2.22)	0.644
Enterococcus faecalis	60 (21.3%)	36 (25.5%)	24 (17%)	1.67 (0.94-2.98)	0.082	1.66 (0.93-2.98)	0.087
Enterococcus faecium	2 (0.7%)	1 (0.7%)	1 (0.7%)	1 (0.06-16.15)	1	0.98 (0.06-16.04)	0.99
Haemophilus influenza	5 (1.8%)	4 (2.8%)	1 (0.7%)	4.09 (0.45-37.04)	0.211	4.24 (0.46-38.8)	0.201
Klebsiella pneumoniae	8 (2.8%)	5 (3.6%)	3 (2.1%)	1.69 (0.4-7.22)	0.478	1.84 (0.43-7.96)	0.411
Staphylococcus aureus	62 (22%)	55 (39%)	7 (5%)	12.24 (5.33-28.1)	0.0001	12.3 (5.31-28.24)	0.0001
(Protozoa) Trichomonas vaginalis	16 (5.7%)	13 (9.2%)	3 (2.1%)	4.67 (1.3-16.77)	0.018	4.53 (1.26-16.32)	0.021
Neisseria gonorrhoea	3 (1.1%)	3 (2.1%)	0				
Fungi (Candida species)							
Candida albicans	99 (35.1%)	70 (49.7%)	29 (20.6%)	3.81 (2.25-6.44)	0.0001	3.98 (2.33-6.79)	0.0001
Candida glabrata	11 (3.9%)	7 (5%)	4 (2.8%)	1.79 (0.51-6.25)	0.362	1.77(0.50-6.2)	0.375
Candida tropicalis	5 (1.8%)	4 (2.8%)	1 (0.7%)	4.09 (0.45-37.04)	0.211	4.10 (0.45-37.23)	0.21
Chlamydia trachomatis	9 (3.2%)	6 (4.3%)	3 (2.1%)	2.04 (0.5-8.3)	0.319	2.17 (0.53-8.97)	0.28
Viruses							

Human cytomegalovirus	3 (1.1%)	2 (1.4%)	1 (0.7%)	2.01 (0.18-22.47)	0.569	2.29 (0.20-25.73)	0.501
Human herpesvirus 6	6 (2.1%)	4 (2.8%)	2 (1.4%)	2.03 (0.37-11.26)	0.418	2.01 (0.36-11.21)	0.428

Analysis was performed using univariate logistic regression and multivariate logistic regression analysis with adjustment for site of sample collection and age of participants. ^{a,b,c} Data are column %. Data are n (%) or median (IQR) as appropriate. CI: confidence interval. Bonferroni adjusted p values $\leq .0013$ taking multiple comparisons into consideration were considered statistically significant (bold).

Table 5. 4: Site-specific analysis of the association between vaginal organisms and GBS colonization in Ethiopia

Target	GBS colonised ^a	No GBS colonisation ^b	OR (95% CI)	p-value	aOR	p-values
bacterial organisms						
<i>Lactobacillus crispatus</i>	21 (44.7%)	21 (44.7%)	1 (0.4-2.3)	1	1 (0.45-2.27)	0.987
<i>Lactobacillus gasseri</i>	10 (21.3%)	15 (31.9%)	0.57 (0.2-1.5)	0.246	0.58 (0.23-1.47)	0.253
<i>Lactobacillus inners</i>	31 (66%)	24 (51.1%)	1.9 (0.8-4.26)	0.145	1.8 (0.79-4.22)	0.156
<i>Lactobacillus jensenii</i>	12 (25.5%)	15 (31.9%)	0.73 (0.3-1.8)	0.495	0.7 (0.29-1.78)	0.477
<i>Atopobium vaginae</i>	15 (31.9%)	25 (53.2%)	0.41 (0.18-0.95)	0.039	0.41 (0.18-0.96)	0.04
<i>Bacteroides fragilis</i>	0	2 (4.3%)	0			
<i>Bacteroides vulgatus</i>	8 (17%)	9 (19.2%)	0.87 (0.3-2.47)	0.789	0.87 (0.3-2.5)	0.797
Bacterial vaginosis associated bacterium 2	8 (17%)	15 (31.9%)	0.437 (0.16-1.16)	0.097	0.44 (0.17-1.17)	0.1
<i>Eggerthella lenta</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.5)	1	0.99 (0.06-16.3)	0.993
<i>Gardnerella vaginalis</i>	16 (34%)	21 (44.7%)	0.6 (0.28-1.47)	0.292	0.65 (0.28-1.5)	0.313
<i>Ureaplasma parvum</i>	23 (48.9%)	19 (40.4%)	0	0	1.38 (0.6-3.15)	0.445
<i>Ureaplasma urealyticum</i>	28 (59.6%)	21 (44.7%)	1.8 (0.8-4.13)	0.15	1.8 (0.79-4.1)	0.158
<i>Mobiluncus curtisii</i>	0	3 (6.4%)	0			
<i>Mycoplasma hominis</i>	0	1 (2.1%)	0	0	0.48 (0.11-2.12)	0.335
<i>Mobiluncus mulieris</i>	0	1 (2.1%)	0			
<i>Megasphaera 1</i>	3 (6.4%)	9 (19.2%)	1 (0.46-2.18)	0.076	0.29 (0.07-1.15)	0.079
<i>Megasphaera 2</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.5)	1	0.9 (0.05-15.7)	0.959
<i>Peptostreptococcus anaerobius</i>	21 (44.7%)	11 (23.4%)	2.6 (1-6.4)	0.032	2.65 (1.09-6.45)	0.031
<i>Prevotella bivia</i>	20 (42.6%)	16 (34%)	1.4 (0.6-3.31)	0.397	1.4 (0.61-3.217)	0.419

<i>Prevotella intermedia</i>	0	3 (6.4%)	0			
<i>Leptotrichia amnionii</i>	12 (25.5%)	9 (19.2%)	1.4 (0.5-3.9)	0.459	1.46 (0.54-3.88)	0.453
<i>Veillonella parvula</i>	41 (87%)	40 (85.1%)	1.19 (0.37-3.9)	0.765	1.1 (0.33-3.878)	0.859
<i>Escherichia coli</i>	13 (27.7%)	10 (21.3%)	1.4 (0.54-3.6)	0.473	1.4 (0.5-3.69)	0.463
<i>Enterococcus faecalis</i>	14 (29.8%)	10 (21.3%)	1.57 (0.6.4)	0.346	1.6 (0.6-4)	0.351
<i>Enterococcus faecium</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.5)	1	1 (0.06-17.7)	0.968
<i>Haemophilus influenza</i>	1 (2.1%)	0				
<i>Klebsiella pneumoniae</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.5)	1	1 (0.06-16.97)	0.988
<i>Staphylococcus aureus</i>	55 (39%)	7 (5%)	21.6 (5.7-99.3)	0.0001	21.7 (4.7-100.6)	0.0001
(Protozoa) <i>Trichomonas vaginalis</i>	3 (6.4%)	0				
Fungi (Candida species)						
<i>Candida albicans</i>	10 (21.3%)	12 (25.5%)	1.27 (0.49-3.3)	0.626	1.3 (0.49-3.4)	0.6
<i>Candida glabrata</i>	1 (2.1%)	0				
Viruses						
Human cytomegalovirus	0	1 (2.1%)	0			
Human herpesvirus 6	1 (2.1%)	0				

Analysis was performed using univariate and multivariate logistic regression analysis, with adjustment for age of participants. ^{a,b}Data are row %. Data are n (%) or median (IQR). OR: odds ratio. aOR: adjusted odds ratio. CI: confidence interval. P-values ≤ 0.0013 are considered statistically significant (bold).

Table 5. 5: Site-specific analysis of the association between vaginal organisms and GBS colonization in Kenya.

Target	GBS colonised ^a	No GBS colonisation ^b	OR (95% CI)	p-value	aOR	p-values
<i>Lactobacillus</i> spp						
<i>Lactobacillus crispatus</i>	11 (23.4%)	11 (23.4%)	1 (0.38-2.6)	1	0.99 (0.37-2.6)	0.976
<i>Lactobacillus gasseri</i>	6 (12.8%)	6 (12.8%)	1 (0.3-3.36)	1	0.74 (0.21-2.58)	0.637
<i>Lactobacillus inners</i>	44 (93.6%)	40 (85.1%)	2.5 (0.6-10.6)	0.193	2.8 (0.67-11.7)	0.159
<i>Lactobacillus jensenii</i>	7 (14.9%)	10 (21.3%)	0.6 (0.2-1.8)	0.423	0.59 (0.19-1.78)	0.347
<i>Atopobium vaginae</i>	34 (72.3%)	21 (44.7%)	3.2 (1.37-7.65)	0.007	3.2 (1.3-7.7)	0.011
<i>Bacteroides fragilis</i>	1 (2.1%)	0				
<i>Bacteroides vulgatus</i>	3 (6.4%)	1 (2.1%)	3.1 (0.3-31.3)	0.33	3.2 (0.3-32)	0.98
Bacterial vaginosis associated bacterium 2	10 (21.3%)	11 (23.4%)	0.88 (0.33-2.3)	0.804	0.9 (0.32-2.49)	0.837
<i>Eggerthella lenta</i>	2 (4.3%)	1 (2.1%)	2 (0.18-23.3)	0.565	6 (0.3-114.4)	0.231
<i>Gardnerella vaginalis</i>	32 (68%)	29 (61.7%)	1.3 (0.57-3.09)	0.517	1.15 (0.48-2.78)	0.752
<i>Ureaplasma parvum</i>	26 (55.3%)	27 (57.5%)	0.92 (0.84-2.07)	0.835	0.86 (0.37-1.99)	0.718
<i>Ureaplasma urealyticum</i>	38 (80.9%)	32 (68.1%)	1.98 (0.76-5.12)	1.98	1.66 (0.6-4.47)	0.319
<i>Mobiluncus curtisii</i>	2 (4.3%)	2 (4.3%)	1 (0.13-7.4)	1	0.73 (0.1-5.6)	0.76
<i>Mycoplasma hominis</i>	15 (31.9%)	10 (21.3%)	1.7 (0.68-4.40)	0.246	1.77 (0.66-4.79)	0.255
<i>Mobiluncus mulieris</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.47)	1	1.17 (0.06-21.9)	0.918
<i>Mycoplasma genitalium</i>	0	2 (4.3%)	1.04 (0.70-1.57)	0.835	0	
<i>Megasphaera</i> 1	5 (10.6%)	8 (17%)	0.58 (0.17-1.93)	0.374	0.65 (0.18-2.3)	0.513
<i>Megasphaera</i> 2	8 (17.1%)	3 (6.4%)	3.01 (0.75-12.1)	0.122	4.1 (0.88-19.1)	0.072

<i>Peptostreptococcus anaerobius</i>	22 (46.8%)	14 (29.8%)	2.07 (0.88-4.84)	0.092	2.05 (0.85-4.9)	0.108
<i>Prevotella bivia</i>	26 (55.3%)	14 (29.8%)	2.92 (1.25-6.82)	0.013	2.87 (1.2-6.89)	0.018
<i>Prevotella intermedia</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.47)	1	2.67 (0.09-79.7)	0.571
<i>Leptotrichia amnionii</i>	23 (48.9%)	13 (27.7%)	2.5 (1.1-5.9)	0.036	2.6 (1.08-6.4)	0.034
<i>Veillonella parvula</i>	39 (83%)	27 (57.5%)	3.61 (1.38-9.39)	0.008	3.8 (1.4-10.1)	0.007
<i>Chlamydia trachomatis</i>	2 (4.3%)	1 (2.1%)	2.04 (0.17-23.3)	0.565	1.79 (0.15-20.9)	0.642
<i>Escherichia coli</i>	10 (21.3%)	9 (19.2%)	1.14 (0.42-3.1)	0.797	1.09 (0.39-3.04)	0.866
<i>Enterococcus faecalis</i>	8 (17%)	2 (4.3%)	4.6 (0.9-23)	0.062	4.4 (0.86-22.1)	0.076
<i>Haemophilus influenza</i>	2 (4.3%)	1 (2.1%)	2 (0.17-23.3)	0.565	1.5 (0.12-17.5)	0.761
<i>Klebsiella pneumoniae</i>	4 (8.5%)	2 (4.3%)	2.09 (0.36-12)	0.408	2 (0.34-11.98)	0.428
<i>Staphylococcus aureus</i>	16 (34%)	2 (4.3%)	11.6 (2.49-54.1)	0.002	9.8 (2.07-46.5)	0.004
<i>Neisseria gonorrhoea</i>	0	1 (2.1%)	1	0		
(Protozoa) <i>Trichomonas vaginalis</i>	8 (17%)	2 (4.3%)	4.6 (0.92-23.03)	0.062	6.7 (1.1-39.8)	0.037
Fungi (Candida species)						
<i>Candida albicans</i>	34 (72.3%)	8 (17%)	12.75 (4.7-34.4)	0.0001	12.8 (4.6-35.7)	0.0001
<i>Candida glabrata</i>	5 (10.6%)	3 (6.4%)	1.7 (0.39-7.76)	0.464	2.2 (0.46-10.09)	0.327
<i>Candida tropicalis</i>	3 (6.4%)	1 (2.1%)	3.13 (0.3-31.3)	0.33	2.78 (0.27-28.9)	0.392

Analysis was performed using univariate and multivariate logistic regression analysis, with adjustment for age of participants and number of prior pregnancies. ^{a,b}Data are column percentage. Data are n (%) or median (IQR). OR: odds ratio. aOR: adjusted odds ratio. CI: confidence interval. Bonferroni-adjusted p-values $\leq .0013$ taking multiple comparisons into consideration were considered statistically significant (bold).

Table 5. 6: Site-specific analysis of the association between vaginal organisms and GBS colonization in South Africa.

Target	GBS colonised ^a	No GBS colonisation ^b	OR (95% CI)	p-value	aOR	p-values
bacterial organisms						
<i>Lactobacillus crispatus</i>	16 (34%)	9 (19.2%)	2.18 (0.85-5.60)	0.106	3.5 (1.2-10)	0.022
<i>Lactobacillus gasseri</i>	6 (12.8%)	6 (12.8%)	1 (0.30-3.36)	1	1.14 (0.31-4.16)	0.84
<i>Lactobacillus inners</i>	41 (87.2%)	40 (85.1%)	1.20 (0.37-3.87)	0.765	1.2 (0.33-4.31)	0.779
<i>Lactobacillus jensenii</i>	12 (25.5%)	9 (19.1%)	1.45 (0.54-3.85)	0.459	1.3 (0.48-3.7)	0.585
<i>Atopobium vaginae</i>	36 (76.6%)	37 (78.7%)	0.88 (0.33-2.34)	0.804	1 (0.36-2.8)	0.99
<i>Bacteroides vulgatus</i>	8 (17%)	11 (23.4%)	0.67 (0.24-1.86)	0.443	0.7 (0.25-2.01)	0.546
Bacterial vaginosis associated bacterium 2	20 (42.6%)	25 (53.2%)	0.65 (0.29-1.47)	0.303	0.59 (0.24-1.4)	0.232
<i>Eggerthella lenta</i>	6 (12.8%)	4 (8.5%)	1.57 (0.41-5.98)	0.506	1.7 (0.42-7.06)	0.448
<i>Gardnerella vaginalis</i>	35 (74.5%)	34 (72.3%)	1.11 (0.45-2.79)	0.815	1.2 (0.47-3.27)	0.669
<i>Ureaplasma parvum</i>	27 (57.5%)	29 (61.7%)	0.84 (0.37-1.91)	0.674	0.97 (0.4-2.33)	0.944
<i>Ureaplasma urealyticum</i>	33 (70.2%)	31 (66%)	1.21 (0.51-2.90)	0.658	1.38 (0.54-3.5)	0.5
<i>Mobiluncus curtisii</i>	2 (4.3%)	12 (25.5%)	0.13 (0.03-0.62)	0.01	0.12 (0.02-0.65)	0.013
<i>Mycoplasma hominis</i>	12 (25.5%)	11 (23.4%)	1.12 (0.44-2.88)	0.81	1.26 (0.46-3.5)	0.654
<i>Mobiluncus mulieris</i>	2 (4.3%)	10 (21.3%)	0.16 (0.03-0.80)	0.025	0.15 (0.03-0.8)	0.027
<i>Megasphaera 1</i>	15 (31.9%)	25 (53.2%)	0.41 (0.18-0.95)	0.039	0.36 (0.15-0.9)	0.028
<i>Megasphaera 2</i>	5 (10.6%)	10 (21.3%)	0.44 (0.14-1.4)	0.166	0.22 (0.06-0.86)	0.029
<i>Peptostreptococcus anaerobius</i>	32 (68.1%)	28 (59.6%)	1.45 (0.62-3.37)	0.391	1.98 (0.77-5.07)	0.156
<i>Prevotella bivia</i>	32 (68.1%)	25 (53.2%)	1.88 (0.81-4.35)	0.141	2.18 (0.88-5.39)	0.091
<i>Prevotella intermedia</i>	1 (2.1%)	0	1	0		
<i>Leptotrichia amnionii</i>	14 (29.8%)	26 (55.3%)	0.34 (0.15-0.80)	0.013	0.25 (0.09-0.65)	0.004
<i>Veillonella parvula</i>	43 (91.5%)	34 (72.3%)	4.11 (1.23-13.75)	0.022	4.3 (1.2-15.5)	0.024
<i>Chlamydia trachomatis</i>	4 (8.5%)	2 (4.3%)	2.09 (0.36-12.02)	0.408	3.1 (0.49-19.95)	0.228

<i>Escherichia coli</i>	2 (4.3%)	3 (6.4%)	0.65 (0.10-4.09)	0.648	0.5 (0.07-3.39)	0.478
<i>Enterococcus faecalis</i>	14 (29.8%)	12 (25.5%)	1.24 (0.50-3.06)	0.645	1.18 (0.45-3.11)	0.731
<i>Haemophilus influenza</i>	1 (2.1%)	0	1	0		
<i>Staphylococcus aureus</i>	16 (34%)	3 (6.4%)	7.57 (2.03-28.22)	0.003	5.99 (1.5-23.89)	0.011
(Protozoa) <i>Trichomonas vaginalis</i>	2 (4.3%)	1 (2.1%)	2.04 (0.18-23.35)	0.565	2.4 (0.19-31.25)	0.496
Fungi (Candida species)						
<i>Candida albicans</i>	24 (51.1%)	11 (23.40%)	3.41 (1.41-8.27)	0.007	3.5 (1.4-9)	0.009
<i>Candida glabrata</i>	1 (2.1%)	1 (2.1%)	1 (0.06-16.47)	1	1.04 (0.05-19.8)	0.981
<i>Candida tropicalis</i>	1 (2.1%)	0	1	0		
Viruses						
Human cytomegalovirus	1 (2.1%)	0	1	0		

Analysis was performed using univariate and multivariate logistic regression analysis, with adjustment for age of participants and number of prior pregnancies. ^{a,b,c}Data are column percentage. Data are n (%) or median (IQR). OR: odds ratio. aOR: adjusted odds ratio. CI: confidence interval. Bonferroni-adjusted p-values $\leq .0013$ taking multiple comparisons into consideration were considered statistically significant.

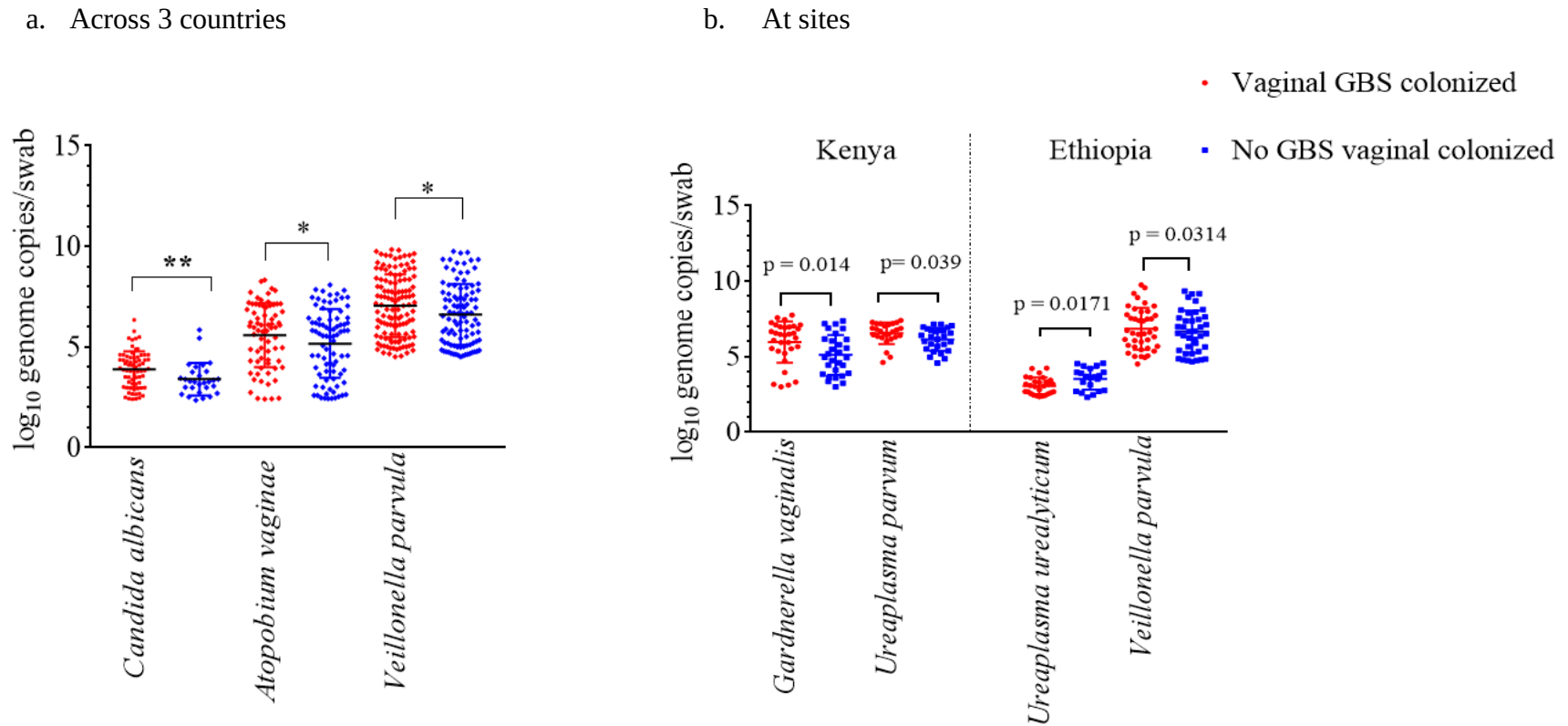
Abundance of vaginal organisms in GBS colonized and non-colonized

The abundance (genome copies per swab) of the organisms was compared between GBS colonized women and those not colonized. The median abundance of *Candida albicans* (3.91 vs 3.30 log₁₀ copies/swab; p=0.007), *Atopobium vaginae* (5.6 vs 5.1 log₁₀ copies/swab; p=0.046), and *Veillonella parvula* (p=0.03, Wilcoxon rank-sum test) was higher among women colonized with GBS compared with those not colonized (Figure 5.2a). Differences in the abundance of vaginal organisms at sites are shown in Figure 5.2b.

Analysis of community state types

Samples were assigned to four CSTs, including CST I (*Lactobacillus crispatus*–dominant group), CST III (*Lactobacillus inners*–dominant group), CST IV (mixed organism) and CST V (*Lactobacillus jensenii*–dominant group). The CST IV-A (14%; 39/282) consisted of anaerobic bacteria accompanied by a low cell count of *Lactobacillus* species, while CST IV-B (73%; 207/282) consisted of anaerobic bacteria. The CST V (1.8%; 5/282) was detected exclusively in women not colonized by GBS (Table 5.7). There was non-significant difference in the overall proportion of CST detected between GBS colonized and uncolonized women (p=0.06). The CST with the large proportion of women being positive for GBS was CST IV-A (56.4%; n=22/39), followed by CST I (54.6%; n=6/11), CST IV-B (51.7%; n=107/207), and CST III (30%; n=6/20).

Figure 5. 2: The density of vaginal organisms with significant differences between women colonized with GBS and non-colonized women



In Kenya, there were significantly high median density of *Gardnerella vaginalis* (median number 6.46 vs 4.95 log₁₀ copies/swab) and *Ureaplasma parvum* (6.67 vs 6.34 log₁₀ copies/swab) in GBS colonized compared with non-colonized women. In Ethiopia, *Ureaplasma urealyticum* (3.08 vs. 3.67 log₁₀ copies/swab) was significantly less in GBS colonized than GBS non-colonized women, whilst *Veillonella parvula* (6.62 vs 6.60 log₁₀ copies/swab) was significantly high in GBS colonized than in GBS non-colonized.

Table 5. 7: Distribution of community state types grouped by GBS status.

Community State Type	Prevalence (N=282)	GBS colonised (N=141)	GBS non-colonised (N=141)	p-value
CST I	11 (3.9%)	6 (4.3%)	5 (3.5%)	1
CST III	20 (7.0%)	6 (4.3%)	14 (9.9%)	0.10
CST IVA	39 (14%)	22 (15.6%)	17 (12.1%)	0.49
CST IV B	207 (73%)	107 (76%)	100 (71%)	0.42
CST V	5 (1.8%)	0 (0%)	5 (3.5%)	0.06

CST: Community state types

5.4 Discussion

The main finding of our study is a positive association between vaginal GBS colonization and specific taxa of the vaginal microbiome such as *Staphylococcus aureus*, *Candida albicans* and *Veillonella parvula*, suggesting that vaginal GBS colonization is influenced by the vaginal organisms. There were also significant differences in the abundance of *Candida albicans*, *Atopobium vaginae* and *Veillonella parvula* between women colonized by GBS and those not colonized. Overall, our findings indicate that vaginal organisms are likely important determinants of vaginal GBS colonization.

The individual vaginal organisms which showed positive association with GBS after Bonferroni correction for multiplicity, included *Veillonella parvula* (aOR, 2.89; p=0.001), *Candida albicans* (aOR, 3.98; p=0.0001) and *Staphylococcus aureus* (aOR, 12.3; p=0.0001). *Candida albicans* has been positively associated with vaginal GBS colonization in non-pregnant women from Kenya, Rwanda and South Africa (194). Furthermore, in vitro studies have shown that GBS has a higher tropism for *Candida albicans* and effectively attaches to vaginal epithelial cells in the presence of *Candida albicans* (211, 293). Colonization by *Candida albicans* may be associated with an increased risk of vaginal GBS colonization (294-296).

Previous studies have shown a mutualistic relationship between *Prevotella bivia* and *Bacteroides fragilis* with GBS colonization in non-pregnant women (200). However, no association was found between vaginal GBS colonization and *Prevotella bivia* or *Bacteroides fragilis* in our study. A negative association between *Lactobacillus* species and GBS colonization has been reported in earlier studies, with oral *Lactobacillus reuteri* being shown to significantly reduce rectovaginal GBS colonization in pregnant women (297, 298). In contrast, no association was found between vaginal GBS colonization and the presence or absence of the lactobacillus species in our study. Most (99%; 260/282) of the participants had *Lactobacillus* species in their vaginal microbiome, mostly *Lactobacillus inners*. Compared to other *Lactobacillus* species, *Lactobacillus inners* is less protective against vaginal colonization by anaerobic bacteria (299). Another study showed an inverse association between *Lactobacillus crispatus* and vaginal GBS colonization in pregnant women (OR; 0.5, p < 0.001), suggesting that *Lactobacillus crispatus* may prevent vaginal GBS colonization (300). However, *Lactobacillus crispatus* was not associated with GBS colonization in our study, but being detected in 31.6% of the pregnant women.

This study, using an open array qPCR-based method, demonstrated significant differences in the quantities of vaginal organisms between women colonized with GBS and women not colonized. Previous studies have utilized culturing methods that are less sensitive to fastidious organisms and may miss about 80% of organisms in the vaginal microbiome (301). In this study, women with GBS colonization had a high abundance of *Candida albicans* ($p=0.007$), *Atopobium vaginae* ($p=0.046$) and *Veillonella parvula* ($p=0.03$) compared to women without GBS colonization.

Ureaplasma species were highly prevalent in pregnant women of this study, which underscores the importance of *Ureaplasma* species as potential prevention targets. *Ureaplasma* species commonly colonize the female genital tract and have been identified as one of the leading pathogens of early-onset neonatal sepsis in South Africa (201) and community-acquired infection in young infants from South Asia (302). The high prevalence of different organisms suggest that organisms associated with vaginal GBS colonization form a high-risk community in the vaginal microbiome which potentially predisposes women to vaginal infections or negative pregnancy outcomes.

The CST V (0% vs 4.3%) was detected exclusively in women not colonized with GBS, suggesting a possible role of the vaginal microbiome dominated by *Lactobacillus gasseri* in combating vaginal GBS colonization. The possible role of *Lactobacillus* species in preventing vaginal GBS colonization is supported by other *in-vitro* studies showing that supernatants of *Lactobacillus* species inhibit GBS viability, biofilm formation and host cell invasion (303). The group CST IVB consisting of anaerobic bacteria or bacterial species associated with bacterial vaginosis, was highly prevalent in our study population (77.3%), suggesting that women of African descent usually harbour diverse vaginal organisms which likely predisposes them to colonization by GBS.

The major strength of this study is investigating organisms at the species level and analyzing the vaginal organisms using molecular methods capable of detecting low-abundance organisms that may have an impact on vaginal GBS colonization. The study samples analyzed in this study were from diverse populations, representing women from the Sub-Saharan African region, a region with paucity of data on vaginal microbiome composition. The limitations of our study include the lack of generalizability of the results to other racial groups, as most of the participants were black African women. Since the vaginal microbiome is inherently dynamic, a longitudinal study would be more informative in examining the relationship between vaginal microbiota and GBS colonization, including understanding the impact of GBS vaginal carriage on the vaginal microbiome.

Despite the limitations mentioned above, the significance of our findings is that we have provided proof of concept for vaginal organisms that may be associated with reduced risk and increased risk of vaginal GBS colonization. The introduction of a GBS vaccine is likely to create a vacant niche in the vaginal microbiome that could be occupied by other organisms. Further work is needed to determine whether organisms associated with GBS colonization are reduced or unmasked and become more prevalent following GBS vaccine introduction.

Conclusions

Overall, the associations between vaginal GBS colonization and the vaginal organisms show that the vaginal microbiome influences vaginal GBS colonization. Further investigations are therefore warranted to determine whether interventions directed towards the vaginal microbiome would impact the risk of GBS disease in women and their offspring.

Chapter 6.0 Integrated Discussion and Conclusion

There were several major findings in this research study. The thesis detailed the sequence types and clonal complexes of colonizing GBS isolates in African and Southeast Asian countries, regions with limited genetic sequence information on GBS. Overall, our analysis of whole-genome sequences of colonizing GBS isolates revealed heterogeneity in GBS STs and CCs as well as variations in surface protein antigen profiles and antimicrobial resistance genes in low and middle income countries. Our study further showed prevalence variation of serotype III, which is also characterized by the absence of the *hvgA* gene in Southeast Asian countries. Conversely, serotype III isolates and the *hvgA* gene were highly prevalent in African countries, highlighting the geographical differences in invasiveness of GBS serotypes.

Previous studies have reported varying prevalence of rectovaginal GBS colonization in pregnant women, ranging from 6.8% to 15.3% in Southeast Asian countries, and from 18.1% to 26.7% in African countries (66). The incidence risk (per 1000 live births) of invasive GBS disease in infants also varies, with the lowest risk observed in Southeast Asia (0.30) and the highest in Africa (1.12) (4). The differences in colonization rates and incidence risk of invasive GBS disease between African and Southeast Asian countries can be explained in part by variations in colonizing sequence types and clonal complexes, which are characterized by differences in invasive potential and surface protein profile.

Our results also showed regional heterogeneity in intra-ST or CC associations. For instance, in African countries, serotype Ia was predominantly associated with ST23, ranging from 30.77% (4/13) in Mali to 96.55% (28/29) in Kenya. Conversely, in Bangladesh and India, serotype Ia was predominantly associated with ST103 (42.86%; 12/28) and ST25 (78.69%; 48/61), respectively. In Southeast Asian countries, serotype III isolates were mostly associated with GBS ST23, a lineage frequently associated with carriage. In contrast, in African countries, serotype III was associated with ST17, a clone significantly associated with invasive disease in infants. The differences in colonizing sequence types between countries or regions can be attributed to several factors, including selection pressure from country-specific factors such as local antibiotic use and population-level immunity to GBS.

In our thesis, we identified the *hvgA* gene in non-CC17 isolates, particularly in isolates of serotype Ia, which belong to a rare lineage CC934. Further analysis revealed no evidence of recombination

within the *hvgA* genomic region of serotype Ia isolates and indicated that CC934 is a historically adapted clone that is geographically restricted to the African region. To date, only 3 isolates of ST934 have been submitted to the public repository (Pubmlst), from colonized women from Ethiopia (119). Localized evolutionary processes may have triggered the emergence of the hypervirulent clone. Although other studies have shown that GBS lineages such as CC17 have expanded through the acquisition of a new phage encoding surface proteins, the expansion of the serotype Ia CC934 lineage appears to be associated with the acquisition of the *hvgA* virulence gene. Additionally, the presence of the *hvgA* gene in serotype IV isolates was attributed to recombination within the capsular gene region, which encompasses a 31.3kb region. The identification of virulence genes in other non-native serotypes underscores the importance of considering non-capsular determinants, such as surface protein genes and the *hvgA* virulence gene, in vaccine development.

The results also show the importance of incorporating serotype IV in vaccine development to prevent CC17 isolates from potential escape of vaccine-induced immunity through capsular switching. The major CCs (CC1, CC10, CC17, CC19 and CC23) identified in women from low and middle-income countries were also detected in high-income countries, showing that there is a global dispersion of GBS sequence types. Although our study showed that low and middle-income countries have a similar GBS population structure to high-income countries, the prevalence of STs and CCs differed geographically. In particular, the less prevalent CCs such as CC255, CC889 and CC1212 were geographically restricted to Bangladesh, indicating the existence of site-specific unique GBS lineages. We showed that Bangladesh (80.9%) and Mali (84.2%) had a lower prevalence of alpha family genes compared to other sites (>94.0%), due to the presence of CC103 associated with a low (40%) prevalence of the surface genes. The results underscore the importance of studying GBS sequences beyond the traditional capsular serotypes for predicting vaccine coverage in different countries. Furthermore, we detected gbs2106 surface protein genes in all GBS isolates, showing that gbs2106 is a potential vaccine candidate antigen that can be targeted against all GBS serotypes expressing the protein.

In addition to characterizing the genetic variability of GBS isolates using whole genome sequencing methods, our study investigated the association between vaginal GBS colonization and the vaginal microbiome in pregnant women from three sub-Saharan African countries. Based on previous studies that found an association between vaginal organisms and GBS in high income countries, our study hypothesis was that the vaginal microbiome influences vaginal GBS

colonization. Our results showed that vaginal GBS colonization is influenced by the vaginal microbiome. The three vaginal organisms that were positively associated with vaginal GBS colonization included *Veillonella parvula* (adjusted odds ratio, 2.89; 95% confidence interval, 1.53-5.43), *Staphylococcus aureus* (aOR, 12.3; 95% CI, 5.31-28.24), and *Candida albicans* (aOR, 3.98; 95% CI, 2.33-6.79). The three organisms that showed a negative association with vaginal GBS colonization included *Megasphaera 1* (aOR, 0.41; 95% CI, 0.22-0.75; p=0.004), *Mobiluncus curtisii* (aOR, 0.21; 95% CI, 0.06-0.66; p=0.007) and *Mobiluncus mulieris* (aOR, 0.22; 95% CI, 0.05-0.82; p=0.024), albeit not significant after Bonferroni correction for multiplicity.

Our results suggest that vaginal organisms positively associated with GBS colonization may amplify the risk of vaginal GBS colonization. Conversely, organisms negatively associated with vaginal GBS colonization may have an antagonistic relationship with the vaginal microbiome, and potentially being protective against vaginal GBS colonization. In contrast to other studies that found an association between vaginal GBS colonization and *Lactobacillus* species, our study did not identify an association between vaginal GBS colonization and *Lactobacillus* species, indicating that vaginal GBS colonization is not dependent on the presence or absence of *Lactobacillus* species. Women colonized with GBS had a significantly higher density of vaginal colonization with *Candida albicans*, *Atopobium vaginalis*, and *Veillonella parvula* than non-colonized women (p<0.05 for each), indicating that vaginal organism density may serve as a potential marker for vaginal GBS colonization. Although our study revealed the association between vaginal GBS colonization and vaginal organisms, it is unclear whether the organisms positively associated with vaginal GBS colonization will proliferate or decline after the introduction of a GBS vaccine. A longitudinal study would be needed to provide further insight into the relationship between the vaginal microbiome and vaginal GBS colonization, including determining the vaginal microbiome that predisposes to vaginal GBS colonization and the impact of GBS vaginal carriage on the vaginal microbiome.

We also investigated the prevalence of vaginal colonization with more than one GBS serotype in pregnant women from six sub-Saharan African countries using the micro-array method. Colonization with more than one GBS serotype is not routinely reported in serotyping laboratories as only a single colony is characterized on the culture plate. In addition, the culture and latex agglutination methods used in routine laboratories for GBS serotyping lacks sensitivity to multiple serotypes when only a single colony is selected. Our results showed a low prevalence ($\leq 5.3\%$) of

colonization with more than one serotype in pregnant women from low and middle-income countries.

Several areas for future research were identified in this study. Future studies should investigate the intricate relationship between genetic sequence types and their invasive potential to better inform on risk posed by sequence types. Furthermore, future studies should aim to determine the genetic population structure of GBS in invasive GBS isolates in low and middle income countries of Africa and Southeast Asia. The observed *hvgA*-positive CC934 represents a gap in our understanding of the evolutionary processes and factors that drive the emergence and expansion of GBS lineages. Although our study revealed an association between the vaginal microbiome and GBS colonization, the potential impact of a GBS vaccine on vaginal organisms positively associated with GBS colonization is unknown. Therefore, future research studies should investigate whether the introduction of a GBS vaccine influences the abundance and dynamics of the vaginal microbiome. Furthermore, sequencing methods should be explored to detect other organism that were not targeted in this study. A major limitation of the study was the unavailability of clinical samples and data from HIV infected women as they were excluded from the primary study. A South African study have shown that serotype III predominated in HIV-infected women (11/27, 40.7% of all serotypes vs. 13.6% in HIV-uninfected women), while serotype Ia was predominant in HIV-uninfected women (13/22, 59.1% of all serotypes vs. 29.6% in HIV-infected women). There is paucity of data on GBS sequence types in HIV infected and uninfected women,

In conclusion, regional variation in GBS STs and CCs may partly account for the regional differences in the incidence of early-onset GBS disease in sub-Saharan Africa and Southeast Asia. Our results suggest that the vaginal microbiome plays a significant role in vaginal GBS colonization. However, further studies are warranted to determine whether interventions targeting the vaginal microbiome could prevent vaginal GBS colonization and thus reduce the incidence of early-onset GBS disease. The low prevalence of multiple GBS serotypes in pregnant women from low and middle income countries, suggests that capsular switching or serotype replacement is less likely to affect GBS vaccine efficacy in pregnant women. Our data on the diversity of GBS sequence types and the vaginal microbiome composition in pregnant women can serve as baseline data for future studies investigating changes in GBS genetic types, multiple GBS serotypes and the vaginal microbiome.

Chapter 7.0 References

1. Edmond KM, Kortsalioudaki C, Scott S, Schrag SJ, Zaidi AKM, Cousens S, Heath PT. Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *The Lancet*. 2012;379(9815):547-56.
2. Koenig JM, Keenan WJ. Group B streptococcus and early-onset sepsis in the era of maternal prophylaxis. *Pediatric clinics of North America*. 2009;56(3):689-708.
3. Verani J, McGee L, Schrag S. Division of bacterial diseases. National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention (CDC) Prevention of perinatal group B streptococcal disease: revised guidelines from CDC. 2010:1-36.
4. Madrid L, Seale AC, Kohli-Lynch M, Edmond KM, Lawn JE, Heath PT, Madhi SA, Baker CJ, Bartlett L, Cutland C, Gravett MG. Infant Group B Streptococcal Disease Incidence and Serotypes Worldwide: Systematic Review and Meta-analyses. *Clinical infectious diseases*. 2017;65(suppl_2):S160-S72.
5. Seale AC, Bianchi-Jassir F, Russell NJ, Kohli-Lynch M, Tann CJ, Hall J, Madrid L, Blencowe H, Cousens S, Baker CJ, Bartlett L. Estimates of the Burden of Group B Streptococcal Disease Worldwide for Pregnant Women, Stillbirths, and Children. *Clinical infectious diseases*. 2017;65(suppl_2):S200-S19.
6. Furuta A, Brokaw A, Manuel G, Dacanay M, Marcell L, Seepersaud R, Rajagopal L, Adams Waldorf K. Bacterial and host determinants of group b streptococcal infection of the neonate and infant. *Frontiers in Microbiology*. 2022;13:820365.
7. Nishihara Y, Dangor Z, French N, Madhi S, Heyderman R. Challenges in reducing group B Streptococcus disease in African settings. *Arch Dis Child*. 2017;102(1):72-7.
8. Le Doare K, Heath PT. An overview of global GBS epidemiology. *Vaccine*. 2013;31:D7-D12.
9. Schuchat A. Epidemiology of group B streptococcal disease in the United States: shifting paradigms. *Clinical microbiology reviews*. 1998;11(3):497-513.
10. Lin SM, Zhi Y, Ahn KB, Lim S, Seo HS. Status of group B streptococcal vaccine development. *Clinical and experimental vaccine research*. 2018;7(1):76-81.
11. Okike IO, Johnson AP, Henderson KL, Blackburn RM, Muller-Pebody B, Ladhani SN, et al. Incidence, etiology, and outcome of bacterial meningitis in infants aged < 90 days in the United Kingdom and Republic of Ireland: prospective, enhanced, national population-based surveillance. *Clinical Infectious Diseases*. 2014;59(10):e150-e7.

12. Song JY, Lim JH, Lim S, Yong Z, Seo HS. Progress toward a group B streptococcal vaccine. *Human vaccines & immunotherapeutics*. 2018;1-13.
13. Shabayek S, Spellerberg B. Group B Streptococcal Colonization, Molecular Characteristics, and Epidemiology. *Frontiers in microbiology*. 2018;9:437.
14. Schrag SJ, Verani JR. Intrapartum antibiotic prophylaxis for the prevention of perinatal group B streptococcal disease: experience in the United States and implications for a potential group B streptococcal vaccine. *Vaccine*. 2013;31:D20-D6.
15. Nanduri SA, Petit S, Smelser C, Apostol M, Alden NB, Harrison LH, et al. Epidemiology of Invasive Early-Onset and Late-Onset Group B Streptococcal Disease in the United States, 2006 to 2015: Multistate Laboratory and Population-Based Surveillance. *JAMA Pediatr*. 2019;173(3):224-33.
16. Huang J, Lin XZ, Zhu Y, Chen C. Epidemiology of group B streptococcal infection in pregnant women and diseased infants in mainland China. *Pediatrics and neonatology*. 2019;60(5):487-95.
17. Madhi SA, Dangor Z, Heath PT, Schrag S, Izu A, Sobanjo-Ter Meulen A, Dull PM. Considerations for a phase-III trial to evaluate a group B Streptococcus polysaccharide-protein conjugate vaccine in pregnant women for the prevention of early- and late-onset invasive disease in young-infants. *Vaccine*. 2013;31 Suppl 4:D52-7.
18. Cutland CL, Schrag SJ, Thigpen MC, Velaphi SC, Wadula J, Adrian PV, Kuwanda L, Groome MJ, Buchmann E, Madhi SA. Increased Risk for Group BStreptococcusSepsis in Young Infants Exposed to HIV, Soweto, South Africa, 2004–20081. *Emerging Infectious Diseases*. 2015;21(4):638-45.
19. Maldonado, Yvonne, Victor Nizet, Elizabeth D. Barnett, Kathryn M. Edwards, and Richard Malley, eds. Remington and Klein's Infectious Diseases of the Fetus and Newborn Infant, *E-Book*. Elsevier Health Sciences, 2024.
20. El Aila NA, Tency I, Claeys G, Saerens B, Cools P, Verstraelen H, Temmerman M, Verhelst R, Vaneechoutte M. Comparison of different sampling techniques and of different culture methods for detection of group B streptococcus carriage in pregnant women. *BMC infectious diseases*. 2010 Dec;10:1-8.
21. Guo D, Xi Y, Wang S, Wang Z. Is a positive Christie-Atkinson-Munch-Peterson (CAMP) test sensitive enough for the identification of Streptococcus agalactiae. *BMC Infectious Diseases*. 2019;19(1):7.

22. Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, Msadek T, Zouine M, Couvé E, Lalioui L, Poyart C, Trieu-Cuot P. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Molecular microbiology*. 2002 Sep;45(6):1499-513.
23. Doran KS, Engelson EJ, Khosravi A, Maisey HC, Fedtke I, Equils O, Michelsen KS, Arditi M, Peschel A, Nizet V. Blood-brain barrier invasion by group B *Streptococcus* depends upon proper cell-surface anchoring of lipoteichoic acid. *The Journal of clinical investigation*. 2005 Sep 1;115(9):2499-507.
24. Chaffin DO, Mentele LM, Rubens CE. Sialylation of group B streptococcal capsular polysaccharide is mediated by cpsK and is required for optimal capsule polymerization and expression. *Journal of bacteriology*. 2005;187(13):4615-26.
25. Cieslewicz MJ, Chaffin D, Glusman G, Kasper D, Madan A, Rodrigues S, Fahey J, Wessels MR, Rubens CE. Structural and genetic diversity of group B streptococcus capsular polysaccharides. *Infection and immunity*. 2005 May;73(5):3096-103.
26. Slotved H-C, Kong F, Lambertsen L, Sauer S, Gilbert GL. Serotype IX, a proposed new *Streptococcus agalactiae* serotype. *Journal of clinical microbiology*. 2007;45(9):2929-36.
27. Slotved HC, Kong F, Lambertsen L, Sauer S, Gilbert GL. Serotype IX. *Journal of clinical microbiology*. 2007;45(9):2929-36.
28. Nuccitelli A, Rinaudo CD, Maione D. Group B *Streptococcus* vaccine: state of the art. *Ther Adv Vaccines*. 2015;3(3):76-90.
29. Alhhazmi A, Pandey A, Tyrrell GJ. Identification of group B streptococcus capsule type by use of a dual phenotypic/genotypic assay. *Journal of clinical microbiology*. 2017;55(9):2637-50.
30. Slotved H-C, Fuursted K, Kavalari ID, Hoffmann S. Molecular Identification of Invasive Non-typeable Group B *Streptococcus* Isolates From Denmark (2015 to 2017). *Frontiers in cellular and infection microbiology*. 2021;11:239.
31. Baker CJ, Kasper DL. Correlation of maternal antibody deficiency with susceptibility to neonatal group B streptococcal infection. *N Engl J Med*. 1976;294(14):753-6.
32. Madhi SA, Cutland CL, Jose L, Koen A, Govender N, Wittke F, Olugbosi M, Sobanjo-ter Meulen A, Baker S, Dull PM, Narasimhan V. Safety and immunogenicity of an investigational maternal trivalent group B streptococcus vaccine in healthy women and their infants: a randomised phase 1b/2 trial. *The Lancet infectious diseases*. 2016 Aug 1;16(8):923-34.

33. Kwatra G, Adrian PV, Shiri T, Buchmann EJ, Cutland CL, Madhi SA. Natural acquired humoral immunity against serotype-specific group B *Streptococcus* rectovaginal colonization acquisition in pregnant women. *Clinical microbiology and infection*. 2015;21(6):568 e13-21.
34. Le Doare K, Kampmann B, Vekemans J, Heath PT, Goldblatt D, Nahm MH, Baker C, Edwards MS, Kwatra G, Andrews N, Madhi SA. Serocorrelates of protection against infant group B streptococcus disease. *The Lancet Infectious Diseases*. 2019 1;19(5):e162-71.
35. Madhi SA, Anderson AS, Absalon J, Radley D, Simon R, Jongihlati B, Strehlau R, Van Niekerk AM, Izu A, Naidoo N, Kwatra G. Potential for maternally administered vaccine for infant group B streptococcus. *New England Journal of Medicine*. 2023 Jul 20;389(3):215-27..
36. Jalalifar S, Havaei SA, Motallebirad T, Moghim S, Fazeli H, Esfahani BN. Determination of surface proteins profile, capsular genotyping, and antibiotic susceptibility patterns of Group B *Streptococcus* isolated from urinary tract infection of Iranian patients. *BMC Research Notes*. 2019;12(1):437.
37. Pietrocola G, Arciola CR, Rindi S, Montanaro L, Speziale P. *Streptococcus agalactiae* non-pilus, cell wall-anchored proteins: involvement in colonization and pathogenesis and potential as vaccine candidates. *Frontiers in immunology*. 2018;9:602.
38. Wilkinson HW, Eagon R. Type-specific antigens of group B type Ic streptococci. *Infection and Immunity*. 1971;4(5):596-604.
39. Bellais S, Six A, Fouet A, Longo M, Dmytruk N, Glaser P, et al. Capsular Switching in Group B *Streptococcus* CC17 Hypervirulent Clone: A Future Challenge for Polysaccharide Vaccine Development. *Journal of Infectious Diseases*. 2012;206(11):1745-52.
40. Dos Santos NFB, da Silva LR, Costa F, de Mattos DM, de Carvalho E, Ferreira LCS, Ferreira RCC. Immunization with a recombinant BibA surface protein confers immunity and protects mice against group B *Streptococcus* (GBS) vaginal colonization. *Vaccine*. 2020;38(33):5286-96.
41. Shabayek S, Abdalla S, Abouzeid AM. Serotype and surface protein gene distribution of colonizing group B streptococcus in women in Egypt. *Epidemiology and infection*. 2014;142(1):208-10.
42. Stålhammar-Carlemalm M, Stenberg L, Lindahl G. Protein rib: a novel group B streptococcal cell surface protein that confers protective immunity and is expressed by most strains causing invasive infections. *The Journal of experimental medicine*. 1993;177(6):1593-603.

43. Mitchell TJ. The pathogenesis of streptococcal infections: from tooth decay to meningitis. *Nature Reviews Microbiology*. 2003;1(3):219-30.
44. Aznar E, Strazielle N, Costa L, Poyart C, Tazi A, Gherzi-Egea J-F, guignot j. The hypervirulent Group B Streptococcus HvgA adhesin promotes brain invasion through transcellular crossing of the choroid plexus. *bioRxiv*. 2024:2024.03. 26.586743.
45. Tazi A, Disson O, Bellais S, Bouaboud A, Dmytruk N, Dramsi S. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. *Journal of experimental medicine*. 2010;207(11):2313-22.
46. Teatero S, McGeer A, Low DE, Li A, Demczuk W, Martin I, Fittipaldi N. Characterization of invasive group B streptococcus strains from the greater Toronto area, Canada. *J Clin Microbiol*. 2014;52(5):1441-7.
47. Lamy MC, Dramsi S, Billoët A, Réglier-Poupet H, Tazi A, Raymond J, Guérin F, Couvé E, Kunst F, Glaser P, Trieu-Cuot P. Rapid detection of the “highly virulent” group B Streptococcus ST-17 clone. *Microbes and infection*. 2006 Jun 1;8(7):1714-22.
48. Manning SD, Springman AC, Lehotzky E, Lewis MA, Whittam TS, Davies HD. Multilocus sequence types associated with neonatal group B streptococcal sepsis and meningitis in Canada. *Journal of clinical microbiology*. 2009;47(4):1143-8.
49. Larsson C, Lindroth M, Nordin P, Stålhammar-Carlemalm M, Lindahl G, Krantz I. Association between low concentrations of antibodies to protein α and Rib and invasive neonatal group B streptococcal infection. *Archives of Disease in Childhood-Fetal and Neonatal Edition*. 2006;91(6):F403-F8.
50. Kekic D, Gajic I, Opavski N, Kojic M, Vukotic G, Smitran A. Trends in molecular characteristics and antimicrobial resistance of group B streptococci: a multicenter study in Serbia, 2015–2020. *Scientific Reports*. 2021;11(1):540.
51. Bianchi-Jassir F, Paul P, To KN, Carreras-Abad C, Seale AC, Jauneikaite E, Madhi SA, Russell NJ, Hall J, Madrid L, Bassat Q. Systematic review of Group B Streptococcal capsular types, sequence types and surface proteins as potential vaccine candidates. *Vaccine*. 2020;38(43):6682-94.
52. Banks C, Lindbom BJ, Kitson G, Darsley M, Fischer PB. Preclinical development of a novel Group B Streptococcus (GBS) vaccine candidate for maternal immunization based upon the alpha-like protein family of GBS surface proteins (Alp). *Birth Defects Research*. 2023;115(9):933-44.

53. Rinaudo CD, Telford JL, Rappuoli R, Seib KL. Vaccinology in the genome era. *J Clin Invest.* 2009;119(9):2515-25.
54. Fabbrini M, Rigat F, Rinaudo CD, Passalacqua I, Khacheh S, Creti R, et al. The Protective Value of Maternal Group B Streptococcus Antibodies: Quantitative and Functional Analysis of Naturally Acquired Responses to Capsular Polysaccharides and Pilus Proteins in European Maternal Sera. *Clinical infectious diseases.* 2016;63(6):746-53.
55. Maione D, Margarit I, Rinaudo CD, Massignani V, Mora M, Scarselli M, et al. Identification of a universal Group B streptococcus vaccine by multiple genome screen. *Science.* 2005;309(5731):148-50.
56. Pawlowski A, Lannergård J, Gonzalez-Miro M, Cao D, Larsson S, Persson JJ, et al. A group B Streptococcus alpha-like protein subunit vaccine induces functionally active antibodies in humans targeting homotypic and heterotypic strains. *Cell Reports Medicine.* 2022;3(2).
57. Banks C, Lindbom BJ, Kitson G, Darsley M, Fischer PB. Preclinical development of a novel Group B Streptococcus (GBS) vaccine candidate for maternal immunization based upon the alpha-like protein family of GBS surface proteins (Alp). *Birth Defects Res.* 2023;115(9):933-44.
58. Dominguez K, Randis TM. Toward the development of a protein-based group B Streptococcus vaccine. *Cell Rep Med.* 2022;3(2):100536.
59. Baker CJ. Group B streptococcal infections. *Clinics in perinatology.* 1997;24(1):59-70.
60. Yadeta TA, Worku A, Egata G, Seyoum B, Marami D, Berhane Y. Vertical transmission of group B Streptococcus and associated factors among pregnant women: a cross-sectional study, Eastern Ethiopia. *Infection and drug resistance.* 2018:397-404.
61. Doran KS, Nizet V. Molecular pathogenesis of neonatal group B streptococcal infection: no longer in its infancy. *Molecular microbiology.* 2004;54(1):23-31.
62. Schuchat A. Group B streptococcus. *The Lancet.* 1999;353(9146):51-6.
63. Patras KA, Nizet V. Group B Streptococcal Maternal Colonization and Neonatal Disease: Molecular Mechanisms and Preventative Approaches. *Front Pediatr.* 2018;6:27.
64. Melin P. Neonatal group B streptococcal disease: from pathogenesis to preventive strategies. Clinical microbiology and infection. *European Society of Clinical Microbiology and Infectious Diseases.* 2011;17(9):1294-303.
65. Russell NJ, Seale AC, O'Driscoll M, O'Sullivan C, Bianchi-Jassir F, Gonzalez-Guarin J, et al. Maternal Colonization With Group B Streptococcus and Serotype Distribution Worldwide: Systematic Review and Meta-analyses. *Clinical infectious diseases.* 2017;65(suppl_2):S100-S11.

66. Kwatra G, Cunningham MC, Merrill E, Adrian PV, Ip M, Klugman KP, Tam WH, Madhi SA. Prevalence of maternal colonisation with group B streptococcus: a systematic review and meta-analysis. *The Lancet Infectious Diseases*. 2016 Sep 1;16(9):1076-84..
67. Armer T, Clark P, Duff P, Saravanos K. Rapid intrapartum detection of group B streptococcal colonization with an enzyme immunoassay. *American journal of obstetrics and gynecology*. 1993;168(1 Pt 1):39-43.
68. Regan JA, Klebanoff MA, Nugent RP, Eschenbach DA, Blackwelder WC, Lou Y, Gibbs RS, Rettig PJ, Martin DH, Edelman R, VIP Study Group. Colonization with group B streptococci in pregnancy and adverse outcome. *American journal of obstetrics and gynecology*. 1996 Apr 1;174(4):1354-60.
69. Schafer R, Phillippi JC. Group B Streptococcal Bacteriuria in Pregnancy: An Evidence-Based, Patient-Centered Approach to Care. *Journal of Midwifery & Women's Health*. 2020;65(3):376-81.
70. Kobayashi M, Schrag SJ, Alderson MR, Madhi SA, Baker CJ, Sobanjo-ter Meulen A, Kaslow DC, Smith PG, Moorthy VS, Vekemans J. WHO consultation on group B Streptococcus vaccine development: report from a meeting held on 27–28 April 2016. *Vaccine*. 2019 Nov 28;37(50):7307-14.
71. Madzivhandila M, Adrian PV, Cutland CL, Kuwanda L, Schrag SJ, Madhi SA. Serotype distribution and invasive potential of group B streptococcus isolates causing disease in infants and colonizing maternal-newborn dyads. *PLoS One*. 2011;6(3):e17861.
72. Mukesi M, Iweriebor BC, Obi LC, Nwodo UU, Moyo SR, Okoh AI. Prevalence and capsular type distribution of Streptococcus agalactiae isolated from pregnant women in Namibia and South Africa. *BMC infectious diseases*. 2019;19:1-7.
73. Lee C-C, Hsu J-F, Janapatla RP, Chen C-L, Zhou Y-L, Lien R, Chiu C-H. Clinical and microbiological characteristics of group B streptococcus from pregnant women and diseased infants in intrapartum antibiotic prophylaxis era in Taiwan. *Scientific reports*. 2019;9(1):1-7.
74. Barcaite E, Bartusevicius A, Tameliene R, Kliucinskas M, Maleckiene I, Nadisauskiene R. Prevalence of maternal group B streptococcal colonisation in European countries. *Acta Obstetricia et Gynecologica Scandinavica*. 2008;87(3):260-71.
75. Russell NJ, Seale AC, O'Sullivan C, Le Doare K, Heath PT, Lawn JE, et al. Risk of Early-Onset Neonatal Group B Streptococcal Disease With Maternal Colonization Worldwide: Systematic Review and Meta-analyses. *Clinical infectious diseases*. 2017;65(suppl_2):S152-S9.

76. Africa CW, Kaambo E. Group B Streptococcus serotypes in pregnant women from the Western Cape Region of South Africa. *Frontiers in public health*. 2018;6:356.
77. Madrid L, Seale AC, Kohli-Lynch M, Edmond KM, Lawn JE, Heath PT, et al. Infant Group B Streptococcal Disease Incidence and Serotypes Worldwide: Systematic Review and Meta-analyses. *Clin Infect Dis*. 2017;65(suppl_2):S160-s72.
78. Teatero S, Athey TBT, Caesele PV, Horsman G, Alexander DC, Melano RG, et al. Emergence of Serotype IV Group B Streptococcus Adult Invasive Disease in Manitoba and Saskatchewan, Canada, Is Driven by Clonal Sequence Type 459 Strains. *Journal of clinical microbiology*. 2015;53(9):2919-26.
79. Dangor Z, Lala SG, Kwatra G, Madhi SA. Group B Streptococcus: developing a correlate of protection for a vaccine against neonatal infections. *Current opinion in infectious diseases*. 2016;29(3):262-7.
80. Seale AC, Koech AC, Sheppard AE, Barsosio HC, Langat J, Anyango E, et al. Maternal colonization with Streptococcus agalactiae and associated stillbirth and neonatal disease in coastal Kenya. *Nature microbiology*. 2016;1(7):16067.
81. Lewnard JA, Cowley LA. Serotype association with neonatal invasive potential, disease manifestation, and clinical outcome in group B Streptococcus: A Bayesian analysis. *bioRxiv*. 2018 20:421412.
82. To K-N, Powell O, Jamroz D, Kopunova R, Anastasiadou K, Faal A, et al. RAPD PCR detects co-colonisation of multiple group B streptococcus genotypes: A practical molecular technique for screening multiple colonies. *Journal of microbiological methods*. 2021;190:106322.
83. Khatami A, Randis TM, Tavares L, Gegick M, Suzman E, Ratner AJ. Vaginal co-colonization with multiple Group B Streptococcus serotypes. *Vaccine*. 2019;37(3):409-11.
84. Wyres KL, Lamberts LM, Croucher NJ, McGee L, Von Gottberg A, Liñares J, et al. Pneumococcal capsular switching: a historical perspective. *The Journal of infectious diseases*. 2013;207(3):439-49.
85. Kamng'ona AW, Hinds J, Bar-Zeev N, Gould KA, Chaguza C, Msefula C, et al. High multiple carriage and emergence of Streptococcus pneumoniae vaccine serotype variants in Malawian children. *BMC infectious diseases*. 2015;15:234.
86. Khan UB, Jauneikaite E, Andrews R, Chalker VJ, Spiller OB. Identifying large-scale recombination and capsular switching events in Streptococcus agalactiae strains causing disease in adults in the UK between 2014 and 2015. *Microbial Genomics*. 2022;8(3):000783.

87. Neemuchwala A, Teatero S, Athey TBT, McGeer A, Fittipaldi N. Capsular Switching and Other Large-Scale Recombination Events in Invasive Sequence Type 1 Group B Streptococcus. *Emerging infectious diseases*. 2016;22(11):1941-4.
88. Anthony BF, Okada DM, Hobel CJ. Epidemiology of group B Streptococcus: longitudinal observations during pregnancy. *Journal of Infectious Diseases*. 1978;137(5):524-30.
89. Anthony BF, Eisenstadt R, Carter J, Kim KS, Hobel CJ. Genital and intestinal carriage of group B streptococci during pregnancy. *Journal of Infectious Diseases*. 1981;143(6):761-6.
90. Hoogkamp-Korstanje JA, Gerards L, Cats B. Maternal carriage and neonatal acquisition of group B streptococci. *Journal of Infectious Diseases*. 1982;145(6):800-3.
91. Ferrieri P, Hillier S, Krohn M, Moore D, Paoletti L, Flores A. Characterization of vaginal & rectal colonization with multiple serotypes of group B streptococci using multiple colony picks. *Indian Journal of Medical Research*. 2004;119:208-12.
92. Perez-Ruiz M, Rodríguez-Granger J, Bautista-Marín M, Romero-Noguera J, Rosa-Fraile M. Genetic diversity of Streptococcus agalactiae strains colonizing the same pregnant woman. *Epidemiology & Infection*. 2004;132(2):375-8.
93. Whitney C, Daly S, Limpongsanurak S, Festin M, Thinn K, Chipato T, et al. The international infections in pregnancy study: group B streptococcal colonization in pregnant women. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2004;15(4):267-74.
94. El Aila NA, Tency I, Claeys G, Saerens B, De Backer E, Temmerman M, et al. Genotyping of Streptococcus agalactiae (group B streptococci) isolated from vaginal and rectal swabs of women at 35-37 weeks of pregnancy. *BMC infectious diseases*. 2009;9:1-11.
95. Slotved H-C, Dayie NT, Banini JA, Frimodt-Møller N. Carriage and serotype distribution of Streptococcus agalactiae in third trimester pregnancy in southern Ghana. *BMC pregnancy and childbirth*. 2017;17:1-7.
96. Jisuvei SC, Osoti A, Njeri MA. Prevalence, antimicrobial susceptibility patterns, serotypes and risk factors for group B streptococcus rectovaginal isolates among pregnant women at Kenyatta National Hospital, Kenya; a cross-sectional study. *BMC infectious diseases*. 2020;20(1):1-9.
97. Rosa-Fraile M, Spellerberg B. Reliable detection of group B streptococcus in the clinical laboratory. *Journal of clinical microbiology*. 2017;JCM. 00582-17.
98. Kobayashi M, Schrag SJ, Alderson MR, Madhi SA, Baker CJ, Sobanjo-Ter Meulen A, et al. WHO consultation on group B Streptococcus vaccine development: Report from a meeting held on 27-28 April 2016. *Vaccine*. 2016.

99. Rosini R, Campisi E, De Chiara M, Tettelin H, Rinaudo D, Toniolo C, et al. Genomic analysis reveals the molecular basis for capsule loss in the group B *Streptococcus* population. *PLoS One*. 2015;10(5):e0125985.
100. Ippolito DL, James WA, Tinnemore D, Huang RR, Dehart MJ, Williams J, et al. Group B streptococcus serotype prevalence in reproductive-age women at a tertiary care military medical center relative to global serotype distribution. *BMC infectious diseases*. 2010;10(1):1-11.
101. Poyart C, Tazi A, Réglier-Poupet H, Billoët A, Tavares N, Raymond J, Trieu-Cuot P. Multiplex PCR assay for rapid and accurate capsular typing of group B streptococci. *J Clin Microbiol*. 2007;45(6):1985-8.
102. Imperi M, Pataracchia M, Alfarone G, Baldassarri L, Orefici G, Creti R. A multiplex PCR assay for the direct identification of the capsular type (Ia to IX) of *Streptococcus agalactiae*. *J Microbiol Methods*. 2010;80(2):212-4.
103. Kong F, Ma L, Gilbert GL. Simultaneous detection and serotype identification of *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization. *Journal of medical microbiology*. 2005;54(Pt 12):1133-8.
104. Breeding KM, Ragipani B, Lee K-UD, Malik M, Randis TM, Ratner AJ. Real-time PCR-based serotyping of *Streptococcus agalactiae*. *Scientific reports*. 2016;6:38523-.
105. Rabaan AA, Saunar JV, Bazzi AM, Soriano JL. Modified use of real-time PCR detection of group B *Streptococcus* in pregnancy. *Journal of medical microbiology*. 2017;66(10):1516-20.
106. Kwatra G, Adrian PV, Shiri T, Buchmann EJ, Cutland CL, Madhi SA. Serotype-specific acquisition and loss of group B streptococcus recto-vaginal colonization in late pregnancy. *PLoS One*. 2014;9(6):e98778.
107. Jaksik R, Iwanaszko M, Rzeszowska-Wolny J, Kimmel M. Microarray experiments and factors which affect their reliability. *Biology Direct*. 2015;10(1):46.
108. Kretschy N, Somoza MM. Comparison of the Sequence-Dependent Fluorescence of the Cyanine Dyes Cy3, Cy5, DyLight DY547 and DyLight DY647 on Single-Stranded DNA. *PLOS ONE*. 2014;9(1):e85605.
109. Wen L, Wang Q, Li Y, Kong F, Gilbert GL, Cao B, et al. Use of a serotype-specific DNA microarray for identification of group B *Streptococcus* (*Streptococcus agalactiae*). *Journal of clinical microbiology*. 2006;44(4):1447-52.

110. Sheppard AE, Vaughan A, Jones N, Turner P, Turner C, Efstratiou A, et al. Capsular Typing Method for *Streptococcus agalactiae* Using Whole-Genome Sequence Data. *Journal of clinical microbiology*. 2016;54(5):1388-90.
111. Kandasamy R, Gurung M, Thapa A, Ndimah S, Adhikari N, Murdoch DR, et al. Multi-serotype pneumococcal nasopharyngeal carriage prevalence in vaccine naïve Nepalese children, assessed using molecular serotyping. *PLoS One*. 2015;10(2):e0114286.
112. Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD, et al. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *Journal of clinical microbiology*. 2011;49(5):1784-9.
113. Zurn K, Lander F, Hufnagel M, Monecke S, Berner R. Microarray Analysis of Group B Streptococci Causing Invasive Neonatal Early- and Late-onset Infection. *Pediatr Infect Dis J*. 2020;39(5):449-53.
114. Teatero S, McGeer A, Li A, Gomes J, Seah C, Demczuk W, et al. Population structure and antimicrobial resistance of invasive serotype IV group B Streptococcus, Toronto, Ontario, Canada. *Emerging infectious diseases*. 2015;21(4):585.
115. Jt A, Aso J O, At O, Aso Oo O. Molecular Epidemiological Status of Group B Streptococcus in Ile Ife South Western Nigeria. *Archives of Medicine*. 2018;10(3).
116. Spratt BG, Hanage WP, Li B, Aanensen DM, Feil EJ. Displaying the relatedness among isolates of bacterial species – the eBURST approach. *FEMS Microbiology Letters*. 2004;241(2):129-34.
117. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res*. 2018;3:124.
118. Teatero S, Ferrieri P, Martin I, Demczuk W, McGeer A, Fittipaldi N. Serotype Distribution, Population Structure, and Antimicrobial Resistance of Group B Streptococcus Strains Recovered from Colonized Pregnant Women. *Journal of clinical microbiology*. 2017;55(2):412-22.
119. Jones N, Bohnsack JF, Takahashi S, Oliver KA, Chan M-S, Kunst F, et al. Multilocus sequence typing system for group B streptococcus. *Journal of clinical microbiology*. 2003;41(6):2530-6.
120. Davies HD, Jones N, Whittam TS, Elsayed S, Bisharat N, Baker CJ. Multilocus sequence typing of serotype III group B streptococcus and correlation with pathogenic potential. *J Infect Dis*. 2004;189(6):1097-102.

121. Wang P, Tong J-j, Ma X-h, Song F-l, Fan L, Guo C-m, et al. Serotypes, Antibiotic Susceptibilities, and Multi-Locus Sequence Type Profiles of *Streptococcus agalactiae* Isolates Circulating in Beijing, China. *PLOS ONE*. 2015;10(3):e0120035.
122. Manning SD, Lewis MA, Springman AC, Lehotzky E, Whittam TS, Davies HD. Genotypic diversity and serotype distribution of group B streptococcus isolated from women before and after delivery. *Clinical infectious diseases*. Society of America. 2008;46(12):1829-37.
123. Da Cunha V, Davies MR, Douarre PE, Rosinski-Chupin I, Margarit I, Spinali S, et al. *Streptococcus agalactiae* clones infecting humans were selected and fixed through the extensive use of tetracycline. *Nature communications*. 2014;5:4544.
124. Bellais S, Six A, Fouet A, Longo M, Dmytruk N, Glaser P, et al. Capsular Switching in Group B *Streptococcus* CC17 Hypervirulent Clone: A Future Challenge for Polysaccharide Vaccine Development. *The Journal of Infectious Diseases*. 2012;206(11):1745-52.
125. Björnsdóttir ES, Martins ER, Erlendsdóttir H, Haraldsson G, Melo-Cristino J, Kristinsson KG, Ramirez M. Changing epidemiology of group B streptococcal infections among adults in Iceland: 1975-2014. *Clinical microbiology and infection*. 2016;22(4):379.e9-.e16.
126. Jamroz D, Bijlsma MW, de Goffau MC, van De Beek D, Kuijpers TW, Parkhill J, van Der Ende A, Bentley SD. Increasing incidence of group B streptococcus neonatal infections in the Netherlands is associated with clonal expansion of CC17 and CC23. *Scientific reports*. 2020 Jun 12;10(1):9539.
127. Marttinen P, Hanage WP, Croucher NJ, Connor TR, Harris SR, Bentley SD, Corander J. Detection of recombination events in bacterial genomes from large population samples. *Nucleic acids research*. 2012;40(1):e6.
128. Metcalf BJ, Chochua S, Gertz RE, Hawkins PA, Ricaldi J, Li Z, et al. Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive *Streptococcus agalactiae* recovered in the USA. *Clinical Microbiology and Infection*. 2017;23(8):574.e7-.e14.
129. Dixit A, Freschi L, Vargas R, Calderon R, Sacchettini J, Drobniewski F, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Scientific Reports*. 2019;9(1):5602.
130. Martins ER, Pedroso-Roussado C, Melo-Cristino J, Ramirez M, Infections PGftSoS. *Streptococcus agalactiae* causing neonatal infections in Portugal (2005–2015): diversification and emergence of a CC17/PI-2b multidrug resistant sublineage. *Frontiers in microbiology*. 2017;8:499.

131. Brzychczy-Wloch M, Gosiewski T, Bulanda M. Multilocus sequence types of invasive and colonizing neonatal group B streptococci in Poland. *Medical Principles and Practice*. 2014;23(4):323-30.
132. Meehan M, Cunney R, Cafferkey M. Molecular epidemiology of group B streptococci in Ireland reveals a diverse population with evidence of capsular switching. *European journal of clinical microbiology & infectious diseases*. *European Society of Clinical Microbiology*. 2014;33(7):1155-62.
133. Andersen M, Smith B, Murra M, Nielsen SY, Slotved H-C, Henriksen TB. Invasive group B Streptococcus strains and clinical characteristics in Danish infants from 1999 to 2009. *Frontiers in Microbiology*. 2022;13:1001953.
134. van Der Mee-Marquet N, Diene S, Barbera L, Courtier-Martinez L, Lafont L, Ouachée A, et al. Analysis of the prophages carried by human infecting isolates provides new insight into the evolution of Group B Streptococcus species. *Clinical Microbiology and Infection*. 2018;24(5):514-21.
135. Morozumi M, Chiba N, Igarashi Y, Mitsunashi N, Wajima T, Iwata S, Ubukata K. Direct identification of Streptococcus agalactiae and capsular type by real-time PCR in vaginal swabs from pregnant women. *Journal of Infection and Chemotherapy*. 2015;21(1):34-8.
136. Wu B, Su J, Li L, Wu W, Wu J, Lu Y, et al. Phenotypic and genetic differences among group B Streptococcus recovered from neonates and pregnant women in Shenzhen, China: 8-year study. *BMC microbiology*. 2019;19:1-9.
137. Ryu H, Park Y-J, Kim Y-k, Chang J, Yu JK. Dominance of clonal complex 10 among the levofloxacin-resistant Streptococcus agalactiae isolated from bacteremic patients in a Korean hospital. *Journal of Infection and Chemotherapy*. 2014;20(8):509-11.
138. Sigaúque B, Kobayashi M, Vubil D, Nhacolo A, Chauque A, Moaine B, et al. Invasive bacterial disease trends and characterization of group B streptococcal isolates among young infants in southern Mozambique, 2001–2015. *PLoS One*. 2018;13(1):e0191193.
139. Seale AC, Koech AC, Sheppard AE, Barsosio HC, Langat J, Anyango E, et al. Maternal colonization with Streptococcus agalactiae and associated stillbirth and neonatal disease in coastal Kenya. *Nature microbiology*. 2016;1(7):1-10.
140. Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC microbiology*. 2017;17(1):194.

141. Pop M. Genome assembly reborn: recent computational challenges. *Briefings in Bioinformatics*. 2009;10(4):354-66.
142. Baker M. De novo genome assembly: what every biologist should know. *Nature Methods*. 2012;9(4):333-7.
143. Liao X, Li M, Zou Y, Wu F-X, Wang J. Current challenges and solutions of de novo assembly. *Quantitative Biology*. 2019;7(2):90-109.
144. Sheppard SK, Guttman DS, Fitzgerald JR. Population genomics of bacterial host adaptation. *Nature Reviews Genetics*. 2018;19(9):549-65.
145. Morais P, Adachi H, Yu YT. Suppression of Nonsense Mutations by New Emerging Technologies. *International journal of molecular sciences*. 2020;21(12).
146. Hammarlöf DL, Kröger C, Owen SV, Canals R, Lacharme-Lora L, Wenner N, et al. Role of a single noncoding nucleotide in the evolution of an epidemic African clade of Salmonella. *Proceedings of the National Academy of Sciences*. 2018;115(11):E2614-E23.
147. Jain R, Rivera MC, Moore JE, Lake JA. Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol*. 2002;61(4):489-95.
148. Brochet M, Rusniok C, Couve E, Dramsi S, Poyart C, Trieu-Cuot P, et al. Shaping a bacterial genome by large chromosomal replacements, the evolutionary history of Streptococcus agalactiae. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(41):15961-6.
149. Tickler IA, Tenover FC, Dewell S, Le VM, Blackman RN, Goering RV, et al. Streptococcus agalactiae, Strains with Chromosomal Deletions Evade Detection with Molecular Methods. *Journal of clinical microbiology*. 2019;57(4):e02040-18.
150. McEwen SA, Collignon PJ. Antimicrobial resistance: a one health perspective. *Antimicrobial resistance in bacteria from livestock and companion animals*. 2018:521-47.
151. Zhu Y, Lin X-Z. Updates in prevention policies of early-onset group B streptococcal infection in newborns. *Pediatrics & Neonatology*. 2021;62(5):465-75.
152. Francois Watkins LK, McGee L, Schrag SJ, Beall B, Jain JH, Pondo T, et al. Epidemiology of Invasive Group B Streptococcal Infections Among Nonpregnant Adults in the United States, 2008-2016. *JAMA Internal Medicine*. 2019;179(4):479-88.
153. Hayes K, O'Halloran F, Cotter L. A review of antibiotic resistance in Group B Streptococcus: the story so far. *Critical Reviews in Microbiology*. 2020;46(3):253-69.

154. da Costa de Souza G, Roque-Borda CA, Pavan FR. Beta-lactam resistance and the effectiveness of antimicrobial peptides against KPC-producing bacteria. *Drug Development Research*. 2022;83(7):1534-54.
155. Bottalico L, Charitos IA, Potenza MA, Montagnani M, Santacroce L. The war against bacteria, from the past to present and beyond. *Expert Review of Anti-infective Therapy*. 2022;20(5):681-706.
156. Metcalf B, Chochua S, Gertz Jr R, Hawkins P, Ricaldi J, Li Z, et al. Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive *Streptococcus agalactiae* recovered in the USA. *Clinical Microbiology and Infection*. 2017;23(8):574. e7-. e14.
157. Seki T, Kimura K, Reid ME, Miyazaki A, Banno H, Jin W, et al. High isolation rate of MDR group B streptococci with reduced penicillin susceptibility in Japan. *J Antimicrob Chemother*. 2015;70(10):2725-8.
158. Leclercq R. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clinical infectious diseases*. 2002;34(4):482-92.
159. Malbruny B, Werno AM, Murdoch DR, Leclercq R, Cattoir V. Cross-resistance to lincosamides, streptogramins A, and pleuromutilins due to the *lsa* (C) gene in *Streptococcus agalactiae* UCN70. *Antimicrobial Agents and Chemotherapy*. 2011;55(4):1470-4.
160. Khan UB. Detailed genomic and antimicrobial resistance comparison of UK *Streptococcus agalactiae* isolates from adults to those of diverse global origins: Cardiff University; 2022.
161. Nakajima Y. Mode of action and resistance mechanisms of antimicrobial macrolides. *Macrolide antibiotics*. 2003:453-I.
162. Sharkey LK, Edwards TA, O'Neill AJ. ABC-F proteins mediate antibiotic resistance through ribosomal protection. *MBio*. 2016;7(2): 01975-15.
163. Oliveira L, Simões LC, Costa NS, Zadoks RN, Pinto TC. The landscape of antimicrobial resistance in the neonatal and multi-host pathogen group B *Streptococcus*: review from a One Health perspective. *Frontiers in Microbiology*. 2022;13:943413.
164. Ji W, Zhang L, Guo Z, Xie S, Yang W, Chen J, et al. Colonization prevalence and antibiotic susceptibility of Group B *Streptococcus* in pregnant women over a 6-year period in Dongguan, China. *PLOS ONE*. 2017;12(8):e0183083.

165. Kao Y, Tsai M-H, Lai M-Y, Chu S-M, Huang H-R, Chiang M-C, et al. Emerging serotype III sequence type 17 group B streptococcus invasive infection in infants: the clinical characteristics and impacts on outcomes. *BMC infectious diseases*. 2019;19(1):1-8.
166. Van Du V, Dung PT, Toan NL, Van Mao C, Bac NT, Van Tong H, et al. Antimicrobial resistance in colonizing group B Streptococcus among pregnant women from a hospital in Vietnam. *Scientific reports*. 2021;11(1):20845.
167. Botelho ACN, Oliveira JG, Damasco AP, Santos KT, Ferreira AFM, Rocha GT, et al. Streptococcus agalactiae carriage among pregnant women living in Rio de Janeiro, Brazil, over a period of eight years. *PLoS One*. 2018;13(5):e0196925.
168. Vinnemeier C, Brust P, Owusu-Dabo E, Sarpong N, Sarfo E, Bio Y, et al. Group B Streptococci serotype distribution in pregnant women in Ghana: assessment of potential coverage through future vaccines. *Tropical Medicine & International Health*. 2015;20(11):1516-24.
169. Verani JR, Schrag SJ. Group B streptococcal disease in infants: progress in prevention and continued challenges. *Clin Perinatol*. 2010;37(2):375-92.
170. Hays C, Louis M, Plainvert C, Dmytruk N, Touak G, Trieu-Cuot P, et al. Changing epidemiology of group B Streptococcus susceptibility to fluoroquinolones and aminoglycosides in France. *Antimicrobial agents and chemotherapy*. 2016;60(12):7424-30.
171. Périchon B, Courvalin P. VanA-type vancomycin-resistant Staphylococcus aureus. *Antimicrobial agents and chemotherapy*. 2009;53(11):4580-7.
172. Depardieu F, Perichon B, Courvalin P. Detection of the van alphabet and identification of enterococci and staphylococci at the species level by multiplex PCR. *Journal of clinical microbiology*. 2004;42(12):5857-60.
173. Connell SR, Tracz DM, Nierhaus KH, Taylor DE. Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob Agents Chemother*. 2003;47(12):3675-81.
174. Chopra I, Roberts M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*. 2001;65(2):232-60.
175. Neuman H, Koren O. The Pregnancy Microbiome. *Nestle Nutrition Institute workshop series*. 2017;88:1-9.
176. Conti C, Malacrino C, Mastromarino P. Inhibition of herpes simplex virus type 2 by vaginal lactobacilli. *Journal of physiology and pharmacology*. 2009;60 Suppl 6:19-26.

177. Atashili J, Poole C, Ndumbe PM, Adimora AA, Smith JS. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. *Aids*. 2008;22(12):1493-501.
178. Younes JA, Lievens E, Hummelen R, van der Westen R, Reid G, Petrova MI. Women and Their Microbes: The Unexpected Friendship. *Trends in microbiology*. 2018;26(1):16-32.
179. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, et al. Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences*. 2011;108 (Supplement 1):4680-7.
180. Sirichoat A, Buppasiri P, Engchanil C, Namwat W, Faksri K, Sankuntaw N, et al. Characterization of vaginal microbiota in Thai women. *Peer J*. 2018;6:e5977.
181. Li J, McCormick J, Bocking A, Reid G. Importance of vaginal microbes in reproductive health. *Reproductive sciences*. 2012;19(3):235-42.
182. Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, et al. Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology*. 2014;160(Pt 10):2272-82.
183. Borgdorff H, Tsvitsivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, et al. Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. *The ISME journal*. 2014;8(9):1781-93.
184. Fettweis JM, Serrano MG, Sheth NU, Mayer CM, Glascock AL, Brooks JP, et al. Species-level classification of the vaginal microbiome. *BMC Genomics*. 2012;13 Suppl 8(Suppl 8):S17.
185. Witkin SS, Mendes-Soares H, Linhares IM, Jayaram A, Ledger WJ, Forney LJ. Influence of vaginal bacteria and D- and L-lactic acid isomers on vaginal extracellular matrix metalloproteinase inducer: implications for protection against upper genital tract infections. *mBio*. 2013;4(4).
186. Tamarelle J, Thiébaud ACM, de Barbeyrac B, Bébéar C, Ravel J, Delarocque-Astagneau E. The vaginal microbiota and its association with human papillomavirus, Chlamydia trachomatis, Neisseria gonorrhoeae and Mycoplasma genitalium infections: a systematic review and meta-analysis. *Clinical Microbiology and Infection*. 2019;25(1):35-47.
187. Short C-ES, Brown RG, Quinlan R, Lee YS, Smith A, Marchesi JR, et al. Lactobacillus-depleted vaginal microbiota in pregnant women living with HIV-1 infection are associated with increased local inflammation and preterm birth. *Frontiers in cellular and infection microbiology*. 2021;9:19.

188. Zhang C, Liu Y, Gao W, Pan Y, Gao Y, Shen J, Xiong H. The direct and indirect association of cervical microbiota with the risk of cervical intraepithelial neoplasia. *Cancer medicine*. 2018;7(5):2172-9.
189. Petrova MI, van den Broek MFL, Balzarini J, Vanderleyden J, Lebeer S. Vaginal microbiota and its role in HIV transmission and infection. *FEMS microbiology reviews*. 2013;375:762-92.
190. Borgdorff H, van der Veer C, van Houdt R, Alberts CJ, de Vries HJ, Bruisten SM, et al. The association between ethnicity and vaginal microbiota composition in Amsterdam, the Netherlands. *PLoS One*. 2017;12(7):e0181135.
191. Serrano MG, Parikh HI, Brooks JP, Edwards DJ, Arodz TJ, Edupuganti L, et al. Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nature medicine*. 2019;25(6):1001-11.
192. Torrone EA, Morrison CS, Chen PL, Kwok C, Francis SC, Hayes RJ, et al. Prevalence of sexually transmitted infections and bacterial vaginosis among women in sub-Saharan Africa: An individual participant data meta-analysis of 18 HIV prevention studies. *PLoS medicine*. 2018;15(2):e1002511.
193. Chen X, Lu Y, Chen T, Li R. The Female Vaginal Microbiome in Health and Bacterial Vaginosis. *Frontiers in cellular and infection microbiology*. 2021;11:631972-.
194. Cools P, Jaspers V, Hardy L, Crucitti T, Delany-Moretlwe S, Mwaura M, et al. A Multi-Country Cross-Sectional Study of Vaginal Carriage of Group B Streptococci (GBS) and Escherichia coli in Resource-Poor Settings: Prevalences and Risk Factors. *PLoS One*. 2016;11(1):e0148052.
195. Lennard KS, Dabee S, Barnabas SL, Havyarimana E, Jaumdally SZ, Botha G, et al. High rates of bacterial vaginosis and Chlamydia in a low-income, high-population-density community in Cape Town. *Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie*. 2017;36(1):1-9.
196. Kenyon C, Colebunders R, Crucitti T. The global epidemiology of bacterial vaginosis: a systematic review. *American journal of obstetrics and gynecology*. 2013;209(6):505-23.
197. World Health Organization. Global health sector strategies on, respectively, HIV, viral hepatitis and sexually transmitted infections for the period 2022-2030. World Health Organization; 2022.
198. Cherpes TL, Kant JA, Melan MA, Hillier SL, Meyn LA, Cosentino LA. Genital Tract Shedding of Herpes Simplex Virus Type 2 in Women: Effects of Hormonal Contraception,

Bacterial Vaginosis, and Vaginal Group B Streptococcus Colonization. *Clinical Infectious Diseases*. 2005;40(10):1422-8.

199. Unemo M, Bradshaw CS, Hocking JS, de Vries HJC, Francis SC, Mabey D, et al. Sexually transmitted infections: challenges ahead. *The Lancet Infectious Diseases*. 2017;17(8):e235-e79.

200. Rosen GH, Randis TM, Desai PV, Sapra KJ, Ma B, Gajer P, et al. Group B Streptococcus and the vaginal microbiota. *The Journal of infectious diseases*. 2017;216(6):744-51.

201. Velaphi SC, Westercamp M, Moleleki M, Pondo T, Dangor Z, Wolter N, et al. Surveillance for incidence and etiology of early-onset neonatal sepsis in Soweto, South Africa. *PLoS One*. 2019;14(4):e0214077.

202. Saha SK, Schrag SJ, El Arifeen S, Mullany LC, Shahidul Islam M, Shang N, et al. Causes and incidence of community-acquired serious infections among young children in south Asia (ANISA). *Lancet* (London, England). 2018;392(10142):145-59.

203. Lennard K, Dabee S, Barnabas SL, Havyarimana E, Blakney A, Jaumdally SZ, et al. Microbial composition predicts genital tract inflammation and persistent bacterial vaginosis in South African adolescent females. *Infection and immunity*. 2018;86(1):e00410-17.

204. Kubota T, Nojima M, Itoh S. Vaginal bacterial flora of pregnant women colonized with group B streptococcus. *Journal of infection and chemotherapy. Japan Society of Chemotherapy*. 2002;8(4):326-30.

205. Altoparlak U, Kadanali A, Kadanali S. Genital flora in pregnancy and its association with group B streptococcal colonization. *International Federation of Gynaecology and Obstetrics*. 2004;87(3):245-6.

206. Moore DM, Meyn LA, Krohn MA, Hillier SL. Association of Sexual Activity with Colonization and Vaginal Acquisition of Group B Streptococcus in Nonpregnant Women. *American Journal of Epidemiology*. 2002;155(10):949-57.

207. Rick AM, Aguilar A, Cortes R, Gordillo R, Melgar M, Samayoa-Reyes G, et al. Group B Streptococci Colonization in Pregnant Guatemalan Women: Prevalence, Risk Factors, and Vaginal Microbiome. *Open forum infectious diseases*. 2017;4(1):ofx020.

208. Zárate G, Nader-Macias M. Influence of probiotic vaginal lactobacilli on in vitro adhesion of urogenital pathogens to vaginal epithelial cells. *Letters in applied microbiology*. 2006;43(2):174-80.

209. Marziali G, Foschi C, Parolin C, Vitali B, Marangoni A. In-vitro effect of vaginal lactobacilli against group B Streptococcus. *Microb Pathog*. 2019;136:103692.

210. Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, et al. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC Genomics*. 2010;11:488.
211. Pidwill GR, Rego S, Jenkinson HF, Lamont RJ, Nobbs AH. Coassociation between Group B Streptococcus and Candida albicans Promotes Interactions with Vaginal Epithelium. *Infect Immunity*. 2018;86(4).
212. Foster-Nyarko E, Kwambana B, Aderonke O, Ceesay F, Jarju S, Bojang A, et al. Associations between nasopharyngeal carriage of Group B Streptococcus and other respiratory pathogens during early infancy. *BMC microbiology*. 2016;16:97.
213. Antonio MAD, Hillier SL, Hawes SE. The Identification of Vaginal Lactobacillus Species and the Demographic and Microbiologic Characteristics of Women Colonized by These Species. *The Journal of Infectious Diseases*. 1999;180(6):1950-6.
214. Daniel Johannes Rönqvist P, Birgitta Forsgren-Brusk U, Elisabeth Grahn-Håkansson E. Lactobacilli in the female genital tract in relation to other genital microbes and vaginal pH. *Acta Obstetricia et Gynecologica Scandinavica*. 2006;85(6):726-35.
215. Virtanen S, Rantsi T, Virtanen A, Kervinen K, Nieminen P, Kalliala I, Salonen A. Vaginal Microbiota Composition Correlates Between Pap Smear Microscopy and Next Generation Sequencing and Associates to Socioeconomic Status. *Scientific Reports*. 2019;9(1):7750.
216. Martin DH. The microbiota of the vagina and its influence on women's health and disease. *The American journal of the medical sciences*. 2012;343(1):2-9.
217. Sharma M, Chopra C, Mehta M, Sharma V, Mallubhotla S, Sistla S, et al. An Insight into Vaginal Microbiome Techniques. *Life (Basel, Switzerland)*. 2021;11(11).
218. Sirichoat A, Sankuntaw N, Engchanil C, Buppasiri P, Faksri K, Namwat W, et al. Comparison of different hypervariable regions of 16S rRNA for taxonomic profiling of vaginal microbiota using next-generation sequencing. *Archives of microbiology*. 2021;203(3):1159-66.
219. Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM. Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. *J Clin Microbiol*. 2007;45(10):3270-6.
220. Grigorenko E, Fisher C, Patel S, Winkelman V, Williamson P, Chancey C, et al. Highly Multiplex Real-Time PCR-Based Screening for Blood-Borne Pathogens on an OpenArray Platform. *J Mol Diagn*. 2017;19(4):549-60.

221. Broccanello C, Gerace L, Stevanato P. QuantStudio(™) 12K Flex OpenArray(®) System as a Tool for High-Throughput Genotyping and Gene Expression Analysis. *Methods in molecular biology* (Clifton, NJ). 2020;2065:199-208.
222. Kwatra G, Izu A, Cutland C, Akaba G, Ali MM, Ahmed Z, Beck MM, et al. Prevalence of group B Streptococcus colonisation in mother–newborn dyads in low-income and middle-income south Asian and African countries: a prospective, observational study. *The Lancet Microbe*. (In Press). 2024.
223. Verani, Jennifer R., Lesley McGee, and Stephanie J. Schrag. "Prevention of perinatal group B streptococcal disease: revised guidelines from CDC, 2010." (2010): 1-36.
224. Afshar B, Broughton K, Creti R, Decheva A, Hufnagel M, Kriz P, et al. International external quality assurance for laboratory identification and typing of Streptococcus agalactiae (Group B streptococci). *Journal of clinical microbiology*. 2011;49(4):1475-82.
225. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. 2017.
226. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*. 2020;70(1):e102.
227. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-9.
228. Pruitt KD, Tatusova T, Brown GR, Maglott DR. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic acids research*. 2012;40(D1):D130-D5.
229. Kong F, Lambertsen LM, Slotved H-C, Ko D, Wang H, Gilbert GL. Use of phenotypic and molecular serotype identification methods to characterize previously nonserotypeable group B streptococci. *Journal of clinical microbiology*. 2008;46(8):2745-50.
230. Inouye M, Conway TC, Zobel J, Holt KE. Short read sequence typing (SRST): multi-locus sequence types from short reads. *BMC Genomics*. 2012;13(1):338.
231. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biology*. 2020;21(1):180.
232. Tonkin-Hill G, Lees JA, Bentley SD, Frost S, Corander J. RhierBAPS: An R implementation of the population clustering algorithm hierBAPS. 2018.
233. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Research*. 2019;47(11):5539-49.

234. Paoletti LC, Kasper DL. Surface structures of group B Streptococcus important in human immunity. *Gram-Positive Pathogens*. 2019;204-27.
235. Rosini R, Rinaudo CD, Soriani M, Lauer P, Mora M, Maione D, et al. Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Molecular microbiology*. 2006;61(1):126-41.
236. van Sorge NM, Quach D, Gurney MA, Sullam PM, Nizet V, Doran KS. The group B streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. *The Journal of infectious diseases*. 2009;199(10):1479-87.
237. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics*. 2022;8(1).
238. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*. 2016;2(4).
239. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PloS one*. 2010;5(3):e9490.
240. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic acids research*. 2014;43(3):e15-e.
241. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. *Nucleic acids research*. 2022;50(D1):D20-d6.
242. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic acids research*. 2021;49(W1):W293-W6.
243. Madhi Shabir A, Kwatra Gaurav, Richardson Simone I, Koen Anthonet L, Baillie Vicky, Cutland Clare L, et al. Durability of ChAdOx1 nCoV-19 (AZD1222) vaccine and hybrid humoral immunity against variants including omicron BA. 1 and BA. 4 6 months after vaccination (COV005): a post-hoc analysis of a randomised, phase 1b–2a trial. *The Lancet Infectious Diseases*. 2023;23(3):295-306.
244. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*. 1927;22(158):209-12.

245. Brugger SD, Frey P, Aebi S, Hinds J, Mühlemann K. Multiple Colonization with *S. pneumoniae* before and after Introduction of the Seven-Valent Conjugated Pneumococcal Polysaccharide Vaccine. *PLOS ONE*. 2010;5(7):e11638.
246. Newton R, Hinds J, Wernisch L. Empirical Bayesian models for analysing molecular serotyping microarrays. *BMC bioinformatics*. 2011;12(1):1-16.
247. Kubota T, Nojima M, Itoh S. Vaginal bacterial flora of pregnant women colonized with group B streptococcus. *Journal of infection and chemotherapy*. 2002;8:326-30.
248. Broccanello C, Gerace L, Stevanato P. QuantStudio™ 12K Flex OpenArray® System as a tool for high-throughput genotyping and gene expression analysis. *Quantitative Real-Time PCR*: Springer; 2020. p. 199-208.
249. Tozetto-Mendoza TR, Mendes-Correa MC, Moron AF, Forney LJ, Linhares IM, Ribeiro da Silva A, Jr., et al. The vaginal Torquetenovirus titer varies with vaginal microbiota composition in pregnant women. *PLoS One*. 2022;17(1):e0262672.
250. Sensini A, Tissi L, Verducci N, Orofino M, von Hunolstein C, Brunelli B, et al. Carriage of group B Streptococcus in pregnant women and newborns: a 2-year study at Perugia General Hospital. *Clinical Microbiology and Infection*. 1997;3(3):324-8.
251. Kwatra G, Cunningham MC, Merrall E, Adrian PV, Ip M, Klugman KP, et al. Prevalence of maternal colonisation with group B streptococcus: a systematic review and meta-analysis. *Lancet Infect Dis*. 2016;16(9):1076-84.
252. Russell NJ, Seale AC, O'Driscoll M, O'Sullivan C, Bianchi-Jassir F, Gonzalez-Guarin J, et al. Maternal Colonization With Group B Streptococcus and Serotype Distribution Worldwide: Systematic Review and Meta-analyses. *Clinical Infectious Diseases*. 2017;65(suppl_2):S100-S11.
253. Da Cunha V, Davies MR, Douarre P-E, Rosinski-Chupin I, Margarit I, Spinali S, et al. Streptococcus agalactiae clones infecting humans were selected and fixed through the extensive use of tetracycline. *Nature communications*. 2014;5(1):4544.
254. McGee L, Chochua S, Li Z, Mathis S, Rivers J, Metcalf B, et al. Multistate, Population-Based Distributions of Candidate Vaccine Targets, Clonal Complexes, and Resistance Features of Invasive Group B Streptococci Within the United States, 2015–2017. *Clinical Infectious Diseases*. 2020;72(6):1004-13.
255. Lu B, Li D, Cui Y, Sui W, Huang L, Lu X. Epidemiology of Group B streptococcus isolated from pregnant women in Beijing, China. *Clinical Microbiology and Infection*. 2014;20(6):O370-O3.

256. Ji W, Liu H, Madhi SA, Cunningham M, Zhang Z, Dangor Z, et al. Clinical and molecular epidemiology of invasive group B streptococcus disease among infants, China. *Emerging Infectious Diseases*. 2019;25(11):2021.
257. Teatero S, Ramoutar E, McGeer A, Li A, Melano RG, Wasserscheid J, et al. Clonal Complex 17 Group B Streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool. *Sci Rep*. 2016;6:20047.
258. Jamrozy D, Bijlsma MW, de Goffau MC, van de Beek D, Kuijpers TW, Parkhill J, et al. Increasing incidence of group B streptococcus neonatal infections in the Netherlands is associated with clonal expansion of CC17 and CC23. *Scientific Reports*. 2020;10(1).
259. Schindler Y, Rahav G, Nissan I, Treygerman O, Prajgrod G, Attia BZ, et al. Group B streptococcus virulence factors associated with different clinical syndromes: Asymptomatic carriage in pregnant women and early-onset disease in the newborn. *Frontiers in Microbiology*. 2023;14.
260. Pawlowski A, Lannergård J, Gonzalez-Miro M, Cao D, Larsson S, Persson JJ, et al. A group B Streptococcus alpha-like protein subunit vaccine induces functionally active antibodies in humans targeting homotypic and heterotypic strains. *Cell Reports Medicine*. 2022;3(2):100511.
261. Bianchi-Jassir F, Paul P, To K-N, Carreras-Abad C, Seale AC, Jauneikaite E, et al. Systematic review of Group B Streptococcal capsular types, sequence types and surface proteins as potential vaccine candidates. *Vaccine*. 2020;38(43):6682-94.
262. Dobrut A, Brzychczy-Włoch M. Immunogenic Proteins of Group B Streptococcus—Potential Antigens in Immunodiagnostic Assay for GBS Detection. *Pathogens*. 2021;11(1):43.
263. Dzanibe S, Adrian PV, Mlacha SZK, Dangor Z, Kwatra G, Madhi SA. Reduced Transplacental Transfer of Group B Streptococcus Surface Protein Antibodies in HIV-infected Mother–Newborn Dyads. *The Journal of Infectious Diseases*. 2017;215(3):415-9.
264. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*. 2011;28(4):464-9.
265. Gori A, Harrison OB, Mlia E, Nishihara Y, Chan JM, Msefula J, et al. Pan-GWAS of Streptococcus agalactiae Highlights Lineage-Specific Genes Associated with Virulence and Niche Adaptation. *mBio*. 2020;11(3).

266. Tazi A, Disson O, Bellais S, Bouaboud A, Dmytruk N, Dramsi S, et al. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. *Journal of Experimental Medicine*. 2010;207(11):2313-22.
267. Teatero S, Ramoutar E, McGeer A, Li A, Melano RG, Wasserscheid J, et al. Clonal Complex 17 Group B Streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool. *Scientific reports*. 2016;6(1):20047.
268. Brokaw A, Nguyen S, Quach P, Orvis A, Furuta A, Johansson-Lindbom B, et al. A Recombinant Alpha-Like Protein Subunit Vaccine (GBS-NN) Provides Protection in Murine Models of Group B Streptococcus Infection. *J Infect Dis*. 2022;226(1):177-87.
269. Berner R. Group B streptococcus vaccines: one step further. *The Lancet Infectious Diseases*. 2021;21(2):158-60.
270. Dzanibe S, Adrian PV, Kimaro Mlacha SZ, Madhi SA. Natural acquired group B Streptococcus capsular polysaccharide and surface protein antibodies in HIV-infected and HIV-uninfected children. *Vaccine*. 2016;34(44):5217-24.
271. Jauneikaite E, Kapatai G, Davies F, Gozar I, Coelho J, Bamford KB, et al. Serial clustering of late-onset group B streptococcal infections in the neonatal unit: a genomic re-evaluation of causality. *Clinical Infectious Diseases*. 2018;67(6):854-60.
272. Seale AC, Bianchi-Jassir F, Russell NJ, Kohli-Lynch M, Tann CJ, Hall J, et al. Estimates of the burden of group B streptococcal disease worldwide for pregnant women, stillbirths, and children. *Clinical infectious diseases*. 2017;65(suppl_2):S200-S19.
273. van Kassel MN, de Boer G, Teeri SA, Jamrozy D, Bentley SD, Brouwer MC, et al. Molecular epidemiology and mortality of group B streptococcal meningitis and infant sepsis in the Netherlands: a 30-year nationwide surveillance study. *The Lancet Microbe*. 2021;2(1):e32-e40.
274. Buurman ET, Timofeyeva Y, Gu J, Kim J-h, Kodali S, Liu Y, et al. A novel hexavalent capsular polysaccharide conjugate vaccine (GBS6) for the prevention of neonatal group B streptococcal infections by maternal immunization. *The Journal of infectious diseases*. 2019;220(1):105-15.
275. Levy C, Ouldali N, Caeymaex L, Angoulvant F, Varon E, Cohen R. Diversity of serotype replacement after pneumococcal conjugate vaccine implementation in Europe. *Elsevier*; 2019. p. 252-3. e3.
276. Wantuch PL, Avci FY. Invasive pneumococcal disease in relation to vaccine type serotypes. *Human vaccines & immunotherapeutics*. 2019;15(4):874-5.

277. Slotved HC, Møller JK, Khalil MR, Nielsen SY. The serotype distribution of *Streptococcus agalactiae* (GBS) carriage isolates among pregnant women having risk factors for early-onset GBS disease: a comparative study with GBS causing invasive infections during the same period in Denmark. *BMC Infect Dis*. 2021;21(1):1129.
278. Furfaro LL, Nathan EA, Chang BJ, Payne MS. Group B streptococcus prevalence, serotype distribution and colonization dynamics in Western Australian pregnant women. *Journal of medical microbiology*. 2019;68(5):728-40.
279. Barro C, Salloum M, Lim S, Delputte P, Le Doare K. Simultaneous carriage of multiple serotypes of Group B Streptococcus: Systematic review and meta-analysis. *Vaccine*. 2023;41(1):15-22.
280. Carreras-Abad C, Ramkhelawon L, Heath PT, Le Doare K. A vaccine against group B Streptococcus: recent advances. *Infection and Drug Resistance*. 2020;13:1263.
281. Swamy GK, Metz TD, Edwards KM, Soper DE, Beigi RH, Campbell JD, et al. Safety and immunogenicity of an investigational maternal trivalent group B streptococcus vaccine in pregnant women and their infants: Results from a randomized placebo-controlled phase II trial. *Vaccine*. 2020;38(44):6930-40.
282. Andrejko K, Ratnasiri B, Hausdorff WP, Laxminarayan R, Lewnard JA. Antimicrobial resistance in paediatric *Streptococcus pneumoniae* isolates amid global implementation of pneumococcal conjugate vaccines: a systematic review and meta-regression analysis. *The Lancet Microbe*. 2021;2(9):e450-e60.
283. Swarthout TD, Gori A, Bar-Zeev N, Kamng'ona AW, Mwalukomo TS, Bonomali F, et al. Evaluation of pneumococcal serotyping of nasopharyngeal-carriage isolates by latex agglutination, whole-genome sequencing (PneumoCaT), and DNA microarray in a high-pneumococcal-carriage-prevalence population in Malawi. *Journal of clinical microbiology*. 2020;59(1):e02103-20.
284. Turner P, Turner C, Jankhot A, Phakaudom K, Nosten F, Goldblatt D. Field Evaluation of Culture plus Latex Sweep Serotyping for Detection of Multiple Pneumococcal Serotype Colonisation in Infants and Young Children. *PLOS ONE*. 2013;8(7):e67933.
285. Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD, et al. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *Journal of clinical microbiology*. 2011;49(5):1784-9.
286. Dhoubhadel BG, Yasunami M, Nguyen HAT, Suzuki M, Vu TH, Thi Thuy Nguyen A, et al. Bacterial load of pneumococcal serotypes correlates with their prevalence and multiple

- serotypes is associated with acute respiratory infections among children less than 5 years of age. *PLoS One*. 2014;9(10):e110777.
287. Neemuchwala A, Teatero S, Athey TB, McGeer A, Fittipaldi N. Capsular Switching and Other Large-Scale Recombination Events in Invasive Sequence Type 1 Group B Streptococcus. *Emerg Infect Dis*. 2016;22(11):1941-4.
 288. Khan UB, Jauneikaite E, Andrews R, Chalker VJ, Spiller OB. Identifying large-scale recombination and capsular switching events in Streptococcus agalactiae strains causing disease in adults in the UK between 2014 and 2015. *Microb Genom*. 2022;8(3).
 289. Dangor Z, Lala SG, Cutland CL, Koen A, Jose L, Nakwa F, et al. Burden of Invasive Group B Streptococcus Disease and Early Neurological Sequelae in South African Infants. *PLOS ONE*. 2015;10(4):e0123014.
 290. Brzychczy-Włoch M, Pabian W, Majewska E, Zuk MG, Kielbik J, Gosiewski T, Bulanda MG. Dynamics of colonization with group B streptococci in relation to normal flora in women during subsequent trimesters of pregnancy. *The new microbiologica*. 2014;37(3):307-19.
 291. Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, et al. The vaginal microbiome and preterm birth. *Nature medicine*. 2019;25(6):1012-21.
 292. McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, Jaoko W, et al. Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. *The Lancet Infectious Diseases*. 2018;18(5):554-64.
 293. Shing SR, Ramos AR, Patras KA, Riestra AM, McCabe S, Nizet V, Coady A. The fungal pathogen *Candida albicans* promotes bladder colonization of group B Streptococcus. *Frontiers in cellular and infection microbiology*. 2020:437.
 294. Mashau RC, Meiring ST, Dramowski A, Magobo RE, Quan VC, Perovic O, et al. Culture-confirmed neonatal bloodstream infections and meningitis in South Africa, 2014–19: a cross-sectional study. *The Lancet Global Health*. 2022;10(8):e1170-e8.
 295. Pillay D, Naidoo L, Swe Swe-Han K, Mahabeer Y. Neonatal sepsis in a tertiary unit in South Africa. *BMC Infect Dis*. 2021;21(1):225.
 296. Zelellw DA, Dessie G, Worku Mengesha E, Balew Shiferaw M, Mela Merhaba M, Emishaw S. A Systemic Review and Meta-analysis of the Leading Pathogens Causing Neonatal Sepsis in Developing Countries. *Biomed Res Int*. 2021;2021:6626983.

297. Sharpe M, Shah V, Freire-Lizama T, Cates EC, McGrath K, David I, et al. Effectiveness of oral intake of *Lactobacillus rhamnosus* GR-1 and *Lactobacillus reuteri* RC-14 on Group B *Streptococcus* colonization during pregnancy: a midwifery-led double-blind randomized controlled pilot trial. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2021;34(11):1814-21.
298. Olsen P, Williamson M, Traynor V, Georgiou C. The impact of oral probiotics on vaginal Group B *Streptococcal* colonisation rates in pregnant women: A pilot randomised control study. *Women and Birth*. 2018;31(1):31-7.
299. Happel A-U, Kullin B, Gamiieldien H, Wentzel N, Zauchenberger CZ, Jaspan HB, et al. Exploring potential of vaginal *Lactobacillus* isolates from South African women for enhancing treatment for bacterial vaginosis. *PLOS Pathogens*. 2020;16(6):e1008559.
300. Starc M, Lučovnik M, Eržen Vrlič P, Jeverica S. Protective Effect of *Lactobacillus crispatus* against Vaginal Colonization with Group B *Streptococci* in the Third Trimester of Pregnancy. *Pathogens*. 2022;11(9).
301. Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP, Romero R. The vaginal microbiome: new information about genital tract flora using molecular based techniques. *International Journal of Obstetrics & Gynaecology*. 2011;118(5):533-49.
302. Saha SK, Schrag SJ, El Arifeen S, Mullany LC, Islam MS, Shang N, et al. Causes and incidence of community-acquired serious infections among young children in south Asia (ANISA): an observational cohort study. *The Lancet*. 2018;392(10142):145-59.
303. Shiroda M, Aronoff DM, Gaddy JA, Manning SD. The impact of *Lactobacillus* on group B streptococcal interactions with cells of the extraplacental membranes. *Microbial Pathogenesis*. 2020;148:104463.
304. Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, Loundou A, et al. High *Atopobium vaginae* and *Gardnerella vaginalis* vaginal loads are associated with preterm birth. *Clinical Infectious Diseases*. 2014;60(6):860-7.
305. Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, Garland SM. The Association of *Atopobium vaginae* and *Gardnerella vaginalis* with Bacterial Vaginosis and Recurrence after Oral Metronidazole Therapy. *The Journal of Infectious Diseases*. 2006;194(6):828-36.
306. Mendling W, Palmeira-de-Oliveira A, Biber S, Prasauskas V. An update on the role of *Atopobium vaginae* in bacterial vaginosis: what to consider when choosing a treatment? A mini review. *Archives of gynecology and obstetrics*. 2019;300(1):1-6.

307. Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, Loundou A, et al. High Atopobium vaginae and Gardnerella vaginalis vaginal loads are associated with preterm birth. *Clinical Infectious Diseases*. 2015;60(6):860-7.
308. Hočevár K, Maver A, Vidmar Šimic M, Hodžić A, Haslberger A, Premru Seršen T, Peterlin B. Vaginal Microbiome Signature Is Associated With Spontaneous Preterm Delivery. *Frontiers in Medicine*. 2019;6.
309. Krohn MA, Hillier SL, Lee ML, Rabe LK, Eschenbach DA. Vaginal Bacteroides species are associated with an increased rate of preterm delivery among women in preterm labor. *Journal of Infectious Diseases*. 1991;164(1):88-93.
310. Leszczynski P, van Belkum A, Pituch H, Verbrugh H, Meisel-Mikolajczyk F. Vaginal carriage of enterotoxigenic Bacteroides fragilis in pregnant women. *Journal of clinical microbiology*. 1997;35(11):2899-903.
311. Zaker N, Lertzman J, Embree J. Neonatal and maternal postpartum bacteroides bacteremia. *Canadian Journal of Infectious Diseases and Medical Microbiology*. 1999;10(5):358-61.
312. Mikamo H, Kawazoe K, Izumi K, Watanabe K, Ueno K, Tamaya T. Studies on the pathogenicity of anaerobes, especially Prevotella bivia, in a rat pyometra model. *Infectious diseases in obstetrics and gynecology*. 1998;6(2):61-5.
313. Obata-Yasuoka M, Ba-Thein W, Hamada H, Hayashi H. A multiplex polymerase chain reaction-based diagnostic method for bacterial vaginosis. *Obstetrics & Gynecology*. 2002;100(4):759-64.
314. Fethers K, Twin J, Fairley CK, Fowkes FJI, Garland SM, Fehler G, et al. Bacterial vaginosis (BV) candidate bacteria: associations with BV and behavioural practices in sexually-experienced and inexperienced women. *PloS one*. 2012;7(2):e30633-e.
315. Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, Jensen JS, et al. Composition of the Vaginal Microbiota in Women of Reproductive Age – Sensitive and Specific Molecular Diagnosis of Bacterial Vaginosis Is Possible? *PLOS ONE*. 2013;8(4):e60670.
316. O'Connell CM, Ferone ME. Chlamydia trachomatis genital infections. *Microbial cell*. 2016;3(9):390.
317. Alger LS, Lovchik JC, Hebel JR, Blackmon LR, Crenshaw MC. The association of Chlamydia trachomatis, Neisseria gonorrhoeae, and group B streptococci with preterm rupture of the membranes and pregnancy outcome. *American journal of obstetrics and gynecology*. 1988;159(2):397-404.

318. Kafil HS, Mobarez AM. Spread of enterococcal surface protein in antibiotic resistant *Enterococcus faecium* and *Enterococcus faecalis* isolates from urinary tract infections. *The open microbiology journal*. 2015;9:14.
319. Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN. Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. *Applied and environmental microbiology*. 2008;74(15):4898-909.
320. Meltzer MC, Desmond RA, Schwebke JR. Association of *Mobiluncus curtisii* with recurrence of bacterial vaginosis. *Sexually transmitted diseases*. 2008;35(6):611.
321. Dubbink JH, Verweij SP, Struthers HE, Ouburg S, McIntyre JA, Morre SA, Peters RP. Genital Chlamydia trachomatis and Neisseria gonorrhoeae infections among women in sub-Saharan Africa: A structured review. *International journal of STD & AIDS*. 2018;29(8):806-24.
322. Duijts, Liesbeth. "Infectious diseases and immune system in infants Risk factors and consequences." (2008).
323. Aung ET, Fairley CK, Williamson DA, Azzato F, Towns JM, Wigan R, et al. Treponema pallidum Detection at Asymptomatic Oral, Anal, and Vaginal Sites in Adults Reporting Sexual Contact with Persons with Syphilis. *Emerg Infect Dis*. 2023;29(10):2083-92.
324. Patel MA, Nyirjesy P. Role of Mycoplasma and Ureaplasma species in female lower genital tract infections. *Current infectious disease reports*. 2010;12(6):417-22.
325. Hillier SL, Krohn MA, Rabe LK, Klebanoff SJ, Eschenbach DA. The Normal Vaginal Flora, H₂O₂-Producing Lactobacilli, and Bacterial Vaginosis in Pregnant Women. *Clinical Infectious Diseases*. 1993;16(Supplement_4):S273-S81.
326. Zhu S, Viejo-Borbolla A. Pathogenesis and virulence of herpes simplex virus. *Virulence*. 2021 31;12(1):2670-702.
327. Salinas AM, Osorio VG, Pacha-Herrera D, Vivanco JS, Trueba AF, Machado A. Vaginal microbiota evaluation and prevalence of key pathogens in ecuadorian women: an epidemiologic analysis. *Scientific reports*. 2020;10(1):18358.
328. Freitas AC, Chaban B, Bocking A, Rocco M, Yang S, Hill JE, et al. The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women. *Scientific reports*. 2017;7(1):9212.
329. Onderdonk AB, Delaney ML, Fichorova RN. The Human Microbiome during Bacterial Vaginosis. *Clinical microbiology reviews*. 2016;29(2):223-38.

330. Mu X, Zhao C, Yang J, Wei X, Zhang J, Liang C, et al. Group B Streptococcus colonization induces Prevotella and Megasphaera abundance-featured vaginal microbiome compositional change in non-pregnant women. *PeerJ*. 2019;7:e7474-e.
331. Chaisilwattana P, Monif GR. In vitro ability of the group B streptococci to inhibit gram-positive and gram-variable constituents of the bacterial flora of the female genital tract. *Infect Dis Obstet Gynecol*. 1995;3(3):91-7.
332. Lin X, Xiao H-M, Liu H-M, Lv W-Q, Greenbaum J, Gong R, Zhang Q, Chen YC, Peng C, Xu XJ, Pan DY. Gut microbiota impacts bone via Bacteroides vulgatus-valeric acid-related pathways. *Nature communications*. 2023;14(1):6853.
333. Pandey PK, Verma P, Kumar H, Bavdekar A, Patole MS, Shouche YS. Comparative analysis of fecal microflora of healthy full-term Indian infants born with different methods of delivery (vaginal vs cesarean): Acinetobacter sp. prevalence in vaginally born infants. *J Biosci*. 2012;37(6):989-98.
334. Aldape MJ, Bayer CR, Rice SN, Bryant AE, Stevens DL. Comparative efficacy of antibiotics in treating experimental Clostridium septicum infection. *International journal of antimicrobial agents*. 2018;52(4):469-73.
335. Sharma NC, Efstratiou A, Mokrousov I, Mutreja A, Das B, Ramamurthy T. Diphtheria. *Nature Reviews Disease Primers*. 2019;5(1):1-18.
336. Davin-Regli A, Pagès J-M. Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment. *Frontiers in microbiology*. 2015;6:392.
337. Ramachandran VL, Ajay Kumar, Piyush Gupta, Sunil Gomber, Vg. Enterobacter aerogenes outbreak in a neonatal intensive care unit. *Pediatrics international*. 1999;41(2):157-61.
338. Whittaker R, Economopoulou A, Dias JG, Bancroft E, Ramliden M, Celentano LP. Epidemiology of invasive Haemophilus influenzae disease, Europe, 2007–2014. *Emerging infectious diseases*. 2017;23(3):396.
339. Effah CY, Sun T, Liu S, Wu Y. Klebsiella pneumoniae: an increasing threat to public health. *Annals of clinical microbiology and antimicrobials*. 2020;19(1):1-9.
340. Matereke LT, Okoh AI. Listeria monocytogenes virulence, antimicrobial resistance and environmental persistence: A review. *Pathogens*. 2020;9(7):528.
341. Drzewiecka D. Significance and roles of Proteus spp. bacteria in natural environments. *Microbial ecology*. 2016;72:741-58.

342. Harwich MD, Serrano MG, Fettweis JM, Alves JMP, Reimers MA, Buck GA, Jefferson KK. Genomic sequence analysis and characterization of *Sneathia amnii* sp. nov. *BMC genomics*. 2012;13(8):S4.
343. De Martino SJ, Mahoudeau I, Brettes JP, Piemont Y, Monteil H, Jaulhac B. Peripartum bacteremias due to *Leptotrichia amnionii* and *Sneathia sanguinegens*, rare causes of fever during and after delivery. *Journal of clinical microbiology*. 2004 Dec;42(12):5940-3.
344. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, et al. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PloS one*. 2012;7(6):e37818.
345. Weiser JN, Ferreira DM, Paton JC. *Streptococcus pneumoniae*: transmission, colonization and invasion. *Nat Rev Microbiol*. 2018;16(6):355-67.
346. Bessen DE. Population biology of the human restricted pathogen, *Streptococcus pyogenes*. *Infection, Genetics and Evolution*. 2009;9(4):581-93.
347. Pappas PG, Lionakis MS, Arendrup MC, Ostrosky-Zeichner L, Kullberg BJ. Invasive candidiasis. *Nature Reviews Disease Primers*. 2018;4(1):18026.
348. Moran GP, Coleman DC, Sullivan DJ. *Candida albicans* versus *Candida dubliniensis*: Why Is *C. albicans* More Pathogenic. *Int J Microbiol*. 2012;2012:205921.
349. Carreté L, Ksiezopolska E, Pegueroles C, Gómez-Molero E, Saus E, Iraola-Guzmán S, Loska D, Bader O, Fairhead C, Gabaldón T. Patterns of Genomic Variation in the Opportunistic Pathogen *Candida glabrata* Suggest the Existence of Mating and a Secondary Association with Humans. *Curr Biol*. 2018;28(1):15-27.e7.
350. Arsenault AB, Bliss JM. Neonatal Candidiasis: New Insights into an Old Problem at a Unique Host-Pathogen Interface. *Curr Fungal Infect Rep*. 2015;9(4):246-52.
351. Ahmadi A, Farhadifar F, Rezaii M, Zandvakili F, Seyedoshohadaei F, Zarei M, Shahgheibi S, Ramazanzadeh R, Roshani D. Group B Streptococci and *Trichomonas vaginalis* infections in pregnant women and those with spontaneous abortion at Sanandaj, Iran. *Iran J Microbiol*. 2018;10(3):166-70.
352. Ross SA, Novak Z, Ashrith G, Rivera LB, Britt WJ, Hedges S, et al. Association between Genital Tract Cytomegalovirus Infection and Bacterial Vaginosis. *The Journal of Infectious Diseases*. 2005;192(10):1727-30.
353. Trépo C, Chan HL, Lok A. Hepatitis B virus infection. *The Lancet*. 2014;384(9959):2053-63.

354. Agut H, Bonnafous P, Gautheret-Dejean A. Laboratory and clinical aspects of human herpesvirus 6 infections. *Clinical microbiology reviews*. 2015;28(2):313-35.
355. Al Shukri I, Hamilton F, Evans M, Cooper S, McKenzie G, Willocks L, Templeton K, Harvala H. Increased number of parvovirus B19 infections in southeast Scotland in 2012–2013. *Clinical Microbiology and Infection*. 2015;21(2):193-6.

Appendices

Appendix 1: Vaginal organisms targets and reasons for inclusion of the vaginal organisms in the open array qPCR panel

Target organism group	Target organisms	Reasons for inclusion (references)
Bacteria	<i>Atopobium vaginae</i>	The organism is a common vaginal bacterium being positively associated with bacterial vaginosis (BV) (304) (305) (306). A high vaginal load ($\geq 10^8$ copies/mL) of <i>A. vaginae</i> is positively associated with preterm birth (307, 308).
	<i>Bacteroides fragilis</i>	A gram-negative bacillus that commonly colonizes the gastro-intestines and causes infections in blood and preterm birth (309, 310) (311). A mutualistic relationship between Group B Streptococcus and <i>B. fragilis</i> has been reported in animal model studies (312). <i>Bacteroides fragilis</i> is a marker for BV (313).
	Bacterial vaginosis associated bacterium 2 (BVAB2)	The organism is negatively associated with vaginal GBS colonization in non-pregnant women from the United States of America (200). The BVAB2 is a predictor of BV (314, 315).
	<i>Chlamydia trachomatis</i>	The leading cause of bacterial sexually transmitted infection (STI) globally, mostly in Africa, and causes serious health sequelae such as cervicitis, pelvic inflammatory disease (316). <i>Chlamydia trachomatis</i> was found to be more common in women with preterm rupture of membranes than women with normal birth (44% vs 15%, $p < 0.001$) (317).
	<i>Enterococcus faecalis</i>	The prevalence of <i>E. faecalis</i> is significantly low in GBS colonized group and high in GBS negative women (2.5% vs 3.8%, $p=0.03$) (204). The bacteria normally colonizes the gastrointestinal tract but also a frequent cause of urinary tract infections (318).

<i>Escherichia coli</i>	The organism causes neonatal sepsis and has been positively associated with vaginal GBS colonization (194).
<i>Gardnerella vaginalis</i>	A lower prevalence of <i>Gardnerella vaginalis</i> detected in GBS colonized (2.2%) compared with non-colonized women (11%; $p<0.05$) (207). The bacteria is a good indicator of BV(315).
<i>Lactobacillus crispatus</i>	The prevalence of <i>Lactobacillus</i> species was significantly low in GBS colonized group than women not colonized with GBS (34.3% vs 53.5%; $p<0.0001$). Commonly used for grouping of community state types (CST) into CST1 (179).
<i>Lactobacillus gasseri</i>	The prevalence of <i>Lactobacillus</i> species was significantly low in GBS colonized group than women not colonized with GBS (34.3% vs 53.5%; $p<0.0001$). Commonly used for grouping of CSTII (179).
<i>Lactobacillus inners</i>	The prevalence of <i>Lactobacillus</i> species was significantly low in GBS colonized group than women not colonized with GBS (34.3% vs 53.5%; $p<0.0001$). Commonly used for grouping of CST III (179).
<i>Lactobacillus jensenii</i>	Commonly used for grouping of CST V (179)
<i>Megasphaera 1</i>	Common in the vaginal microbiome being associated with BV (319). <i>Megasphaera 1</i> was negatively correlated with vaginal GBS colonization ($r= 0.5$; $p<0.05$).
<i>Megasphaera 2</i>	Common in the vaginal microbiome being associated with BV (319).
<i>Mobiluncus curtisii</i>	Common in the vaginal microbiome being associated with BV (319). Highly specific for the diagnosis of BV
<i>Mobiluncus mulieris</i>	Highly specific but not sensitive for the diagnosis of BV (320)
<i>Neisseria gonorrhoeae</i>	One of the most prevalent STI worldwide (321).

<i>Staphylococcus aureus</i>	The prevalence of <i>Staphylococcus aureus</i> was low in GBS colonized women (11.5% vs 18.3%; p=0003) than GBS non-colonised (322).
<i>Treponema pallidum</i>	An STI detected in vaginal lesion swabs (323).
<i>Mycoplasma genitalium</i>	An STI, causes symptomatic cervicitis or urethritis (324).
<i>Mycoplasma hominis</i>	An organism associated with pre mature birth (OR 2.05; 1.78-2.37) (304).
<i>Ureaplasma urealyticum</i>	Common in women both with normal flora and BV ((325)
Herpes simplex 1	An STI that causes genital herpes (326)
Herpes simplex 2	An STI, and vaginal GBS colonization has been positively associated with HSV2 colonization (198).
<i>Enterococcus. faecium</i>	The bacteria is a common organism isolated in aerobic vaginitis (AV), a condition by which the vaginal flora contains aerobic and intestinal pathogens (327).
<i>Ureaplasma parvum</i>	Found to be enriched in women with preterm birth (328)
<i>Prevotella bivia</i>	The prevalence of <i>Prevotella bivia</i> was significantly high in GBS colonized group than GBS non-colonised (200). <i>P. bivia</i> appears to be a prominent modifier of the cervicovaginal mucosal environment (329)
<i>Veillonella parvula</i>	Positively associated with GBS and EOD (200). The genus <i>Veillonella</i> is positively associated with GBS colonization (330)
<i>Prevotella intermedia</i>	<i>Prevotella</i> species is common in the vaginal microbiome being negatively correlated with vaginal GBS colonization (200).

<i>Peptostreptococcus anaerobius</i>	The prevalence was low in GBS colonized group (204). Found to be inhibited by GBS invitro (331).
<i>Bacteroides vulgatus</i>	Common bacteria in the human gastrointestinal (332).
<i>Acinetobacter baumannii</i>	A common nosocomial pathogen and has also been isolated human gastrointestinal and vagina microbiome (333)
<i>Clostridium septicum</i>	A motile pathogen commonly isolated from the human intestines (334)
<i>Corynebacterium diphtheriae</i>	The etiologic agent of human respiratory and cutaneous diphtheria (335).
<i>Eggerthella lenta</i>	<i>E. lenta</i> is part of the normal human intestinal microbiome and vaginal microbiome. Common in women with BV (315).
<i>Enterobacter aerogenes</i>	A opportunistic nosocomial pathogen (336). Found in the flora of human intestines (337)
<i>Haemophilus influenzae</i>	A bacteria that causes life threatening infections in children and adults (338)
<i>Klebsiella pneumoniae</i>	An opportunistic pathogen causing broad spectra of diseases (339). Significantly high in GBS positive than GBS negative group (204).
<i>Listeria monocytogenes</i>	An opportunistic pathogen responsible for a foodborne listeriosis disease (340).

	<i>Proteus vulgaris</i>	The organisms colonizes the human intestinal tracts, being an opportunistic pathogen (341).
	<i>Sneathia amnii</i>	Associated with clinical conditions including bacteremia (342) (343). Common in women but more prevalent in women with BV than those without BV (329, 344)
	<i>Streptococcus pneumoniae</i>	Colonizes the mucosal surface of the upper respiratory airways in humans (345)
	<i>Streptococcus pyogens</i>	Colonizes the mucosal surface of the upper airways and the skins of humans (346).
Fungi	<i>Candida albicans</i>	A fungal species negatively associated with GBS vaginal colonization (6% vs 8.8%; p=0.001) (194). Causes invasive candidiasis (347)
	<i>Candida dubliniensis</i>	Causes neonatal sepsis and invasive candidiasis (348)
	<i>Candida glabrata</i>	Associated with Invasive candidiasis (349).
	<i>Candida krusei</i>	An opportunistic fungal pathogen that causes invasive candidiasis (350)
	<i>Candida lusitaniae</i>	An opportunistic fungal pathogen that causes invasive candidiasis (350)
	<i>Candida parapsilosis</i>	An opportunistic fungal pathogen that causes neonatal sepsis and invasive candidiasis (350)
	<i>Candida tropicalis</i>	An opportunistic fungal pathogen that causes invasive candidiasis (350)
Protozoa	<i>Trichomonas vaginalis</i>	The only pathogenic trichomonad. An STI that causes genital infections in pregnant women (351).

Virus	Cytomegalovirus (CMV)	A viral organism. Women with unhealthy vagitypes (BV) shed CMV in their lower genital tract (352).
	Hepatitis B virus	Responsible for viral infection (353).
	Human herpesvirus 6	Responsible for viral infection (354).
	Parvovirus B19	Responsible for viral infection (355).

Appendix 2

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG



HUMAN RESEARCH ETHICS COMMITTEE
(MEDICAL)

09 December 2020

Mr Takudzwa Matuvhunye
Clinical Microbiology and Infectious Diseases
Faculty of Health Sciences
Parktown
2193

Sent by email to: takudzwamat@gmail.com

Dear Mr Matuvhunye

Re: Protocol Ref No: M191171

Protocol Title: Molecular and genetic characterization of Group B Streptococcus and its association with vaginal microbiome among pregnant women from low-middle income countries

Principal Investigator: Mr Takudzwa Matuvhunye

Progress Report

This letter serves to confirm that the Chairman of the Human Research Ethics Committee (Medical) has noted the progress report for the abovementioned protocol, as detailed in your letter, dated 30 November 2020.

Thank you for keeping us informed and updated.

Yours Sincerely,

.....
Miss Mapula Ramaila
Administrative Officer
Human Research Ethics Committee (Medical)



Research Office Secretariat: Faculty of Health Sciences, Phillip Tobias Building, 3rd Floor, Office 302, Corner York Road and 29 Princess of Wales Terrace, Parktown, 2193 Private Bag 3, Wits 2050 | T+27 (0)11-717-1234/2656/2700/1252 | E: Mapula.Ramaila@wits.ac.za | Office E
HREC-Medical.ResearchOffice@wits.ac.za | Website: www.wits.ac.za/research/about-our-research/ethics-and-research-integrity/

Appendix 3: The mid-point rooted maximum likelihood phylogenetic trees of GBS colonizing isolates in pregnant women from sub-Saharan Africa and Southeast Asia. The phylogenetic trees show the genetic relationship of GBS colonizing isolates in pregnant women in each country, a) The phylogenetic tree of Bangladeshi isolates indicates the presence of serotype III associated with CC1212 and the absence of CC17 thereof. b) In India, serotype III was also associated with CC23, with absence of CC17. c) The phylogenetic tree of GBS colonising isolates in pregnant women from Ethiopia, d) Kenya e) Mali f) Mozambique g) Nigeria h) South Africa.

a): Bangladesh

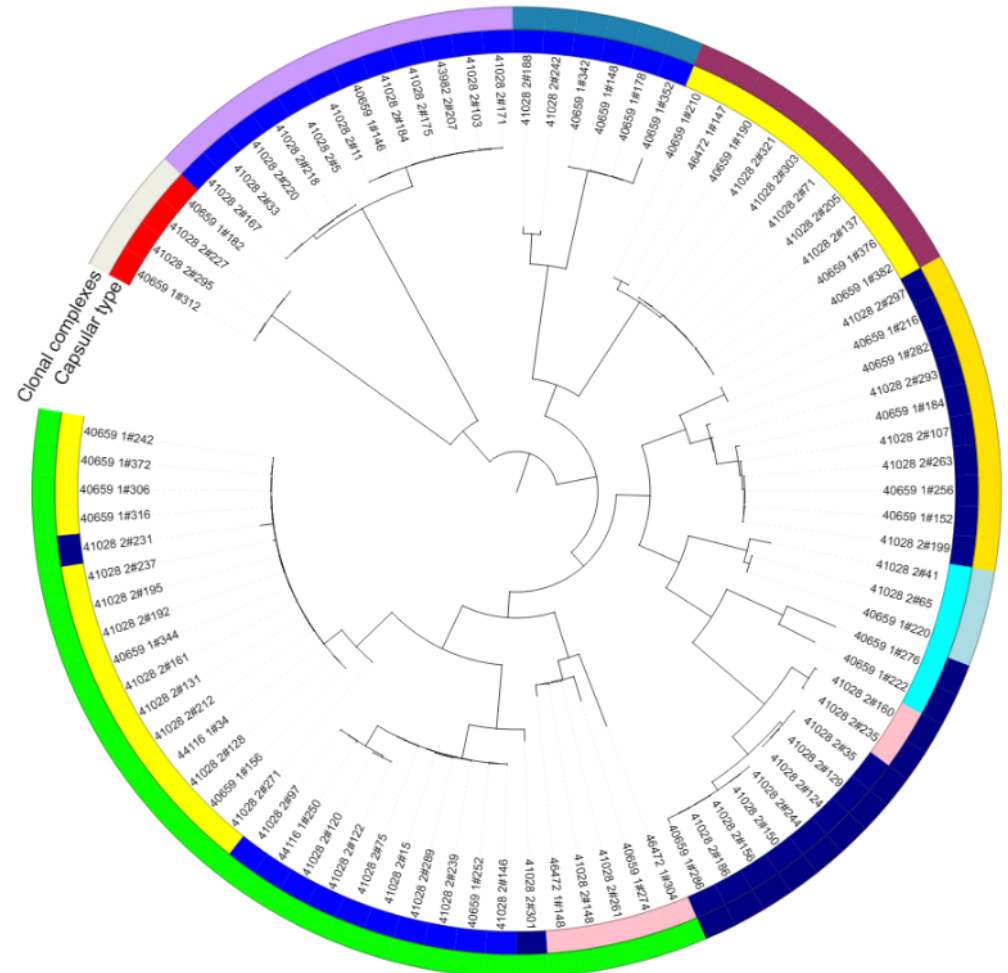
Tree scale: 0.01

Clonal complexes

- 1
- 10
- 17
- 19
- 23
- 26
- 103
- 255
- 889
- 934
- 1212

Capsular type

- Ia
- Ib
- II
- III
- IV
- V
- VI
- VII



b): India

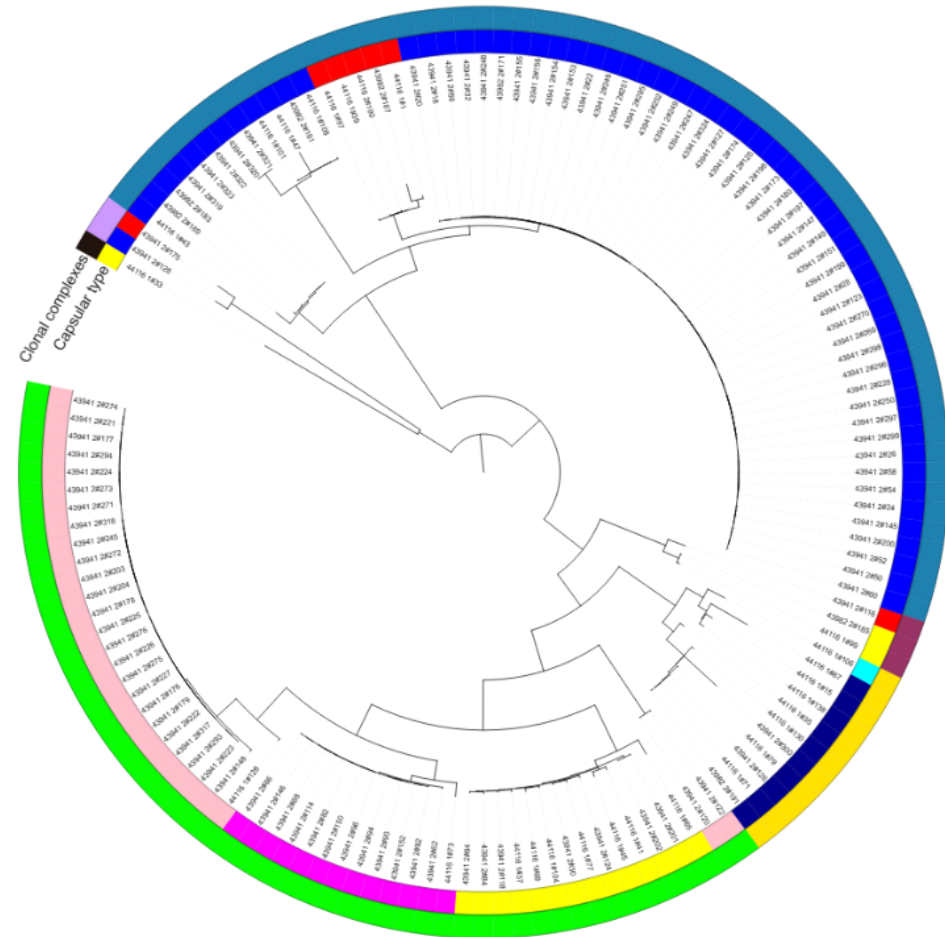
Tree scale: 0.001

Capsular type

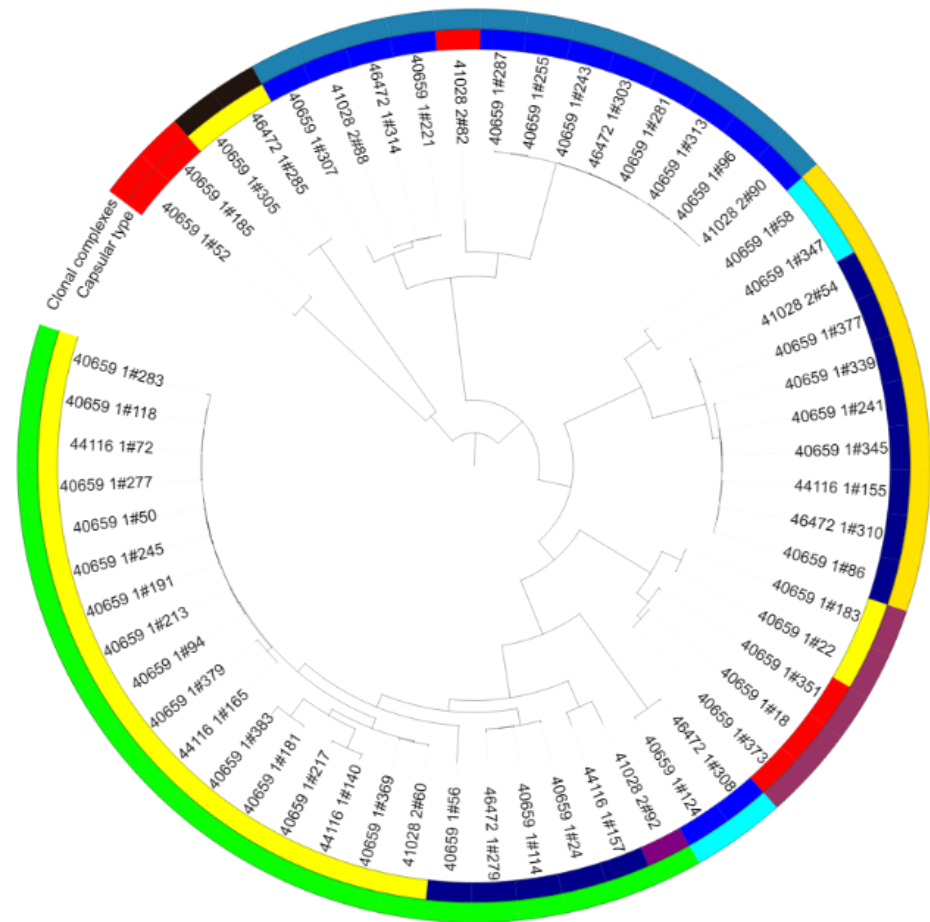
- Ia
- Ib
- II
- III
- IV
- V
- VI
- VII

Clonal complexes

- 1
- 10
- 17
- 19
- 23
- 26
- 103
- 255
- 889
- 934
- 1212



c): Ethiopia



d): Kenya

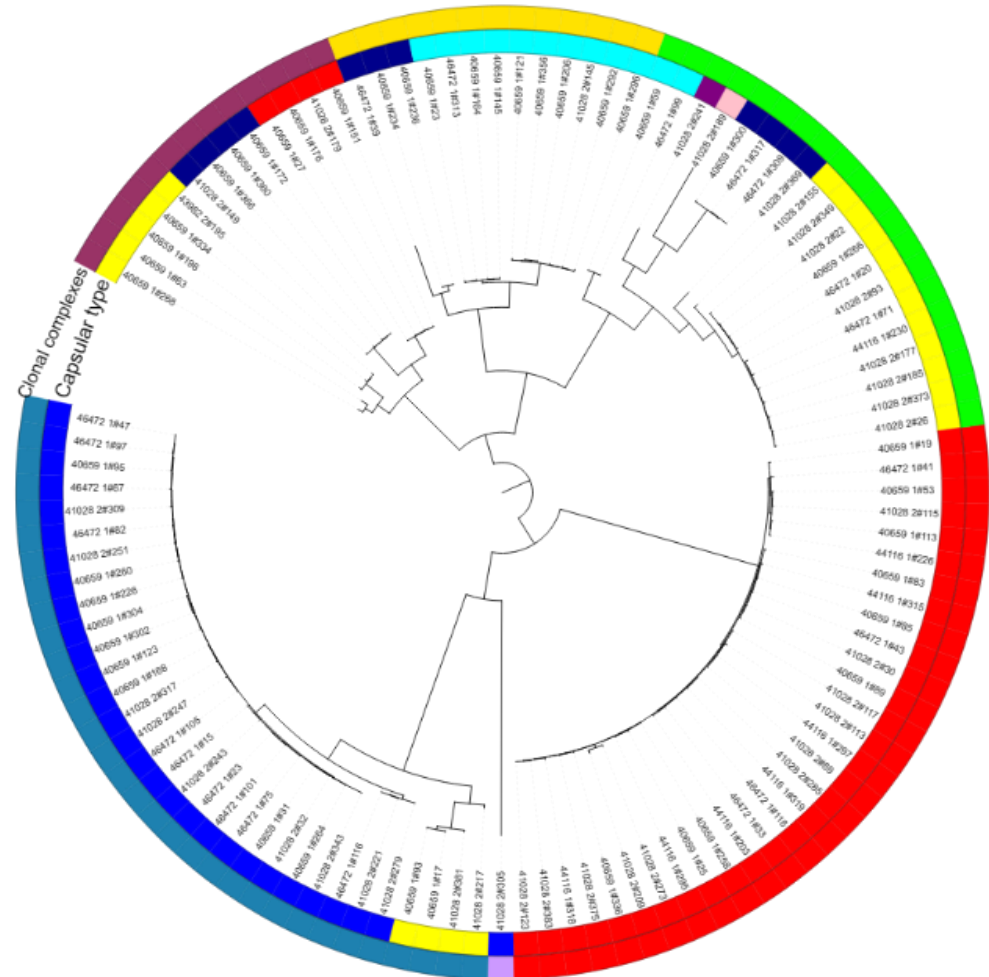
Tree scale: 0.001

Capsular type

- Ia
- Ib
- II
- III
- IV
- V
- VI
- VII

Clonal complexes

- 1
- 10
- 17
- 19
- 23
- 26
- 103
- 255
- 889
- 934
- 1212



e) : Mali

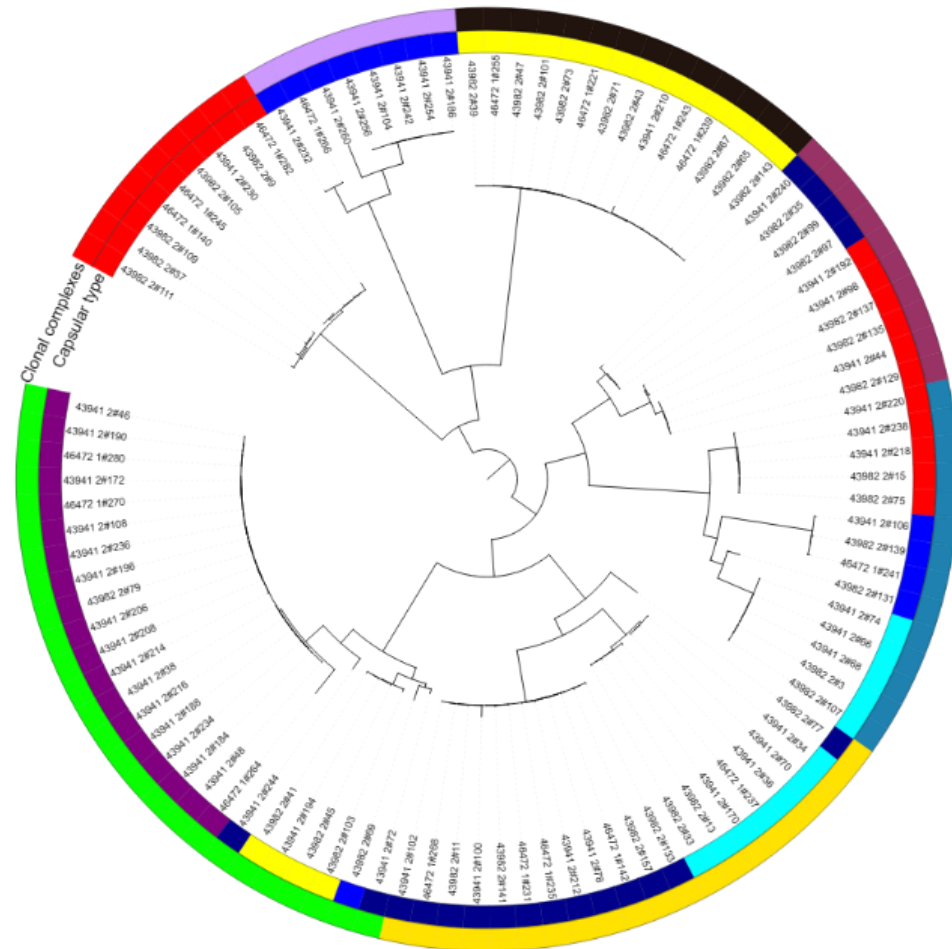
Tree scale: 0.001

Capsular type

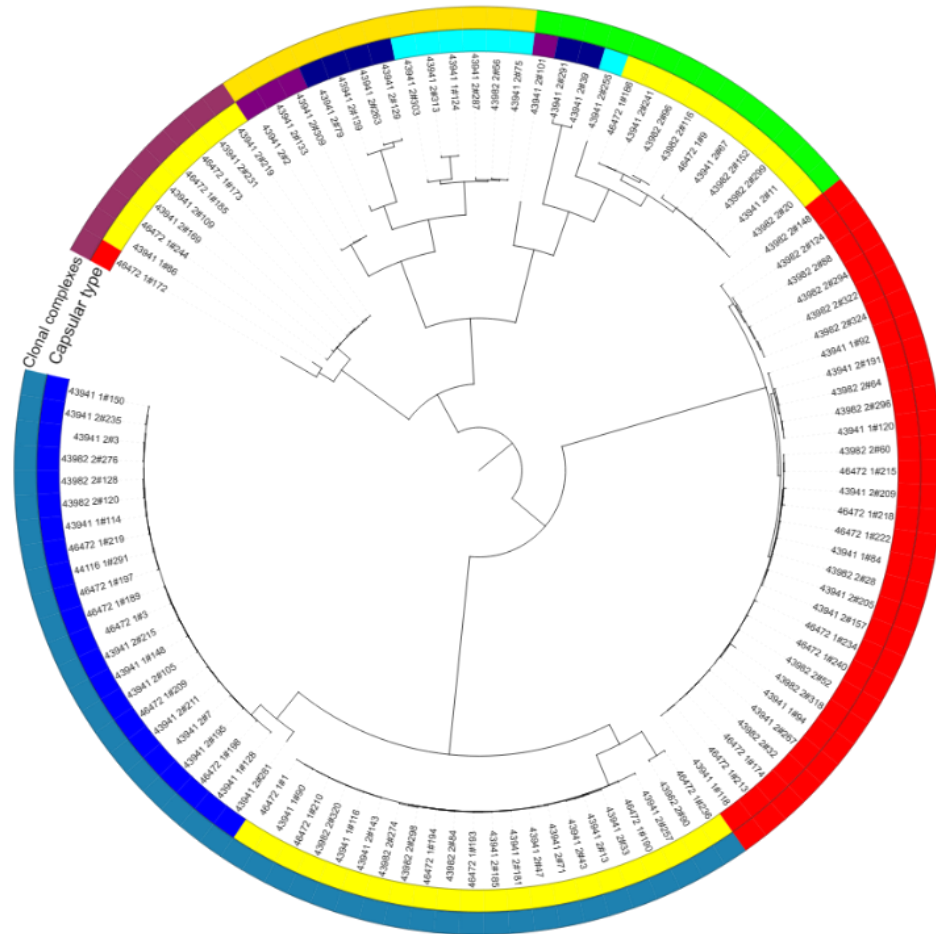
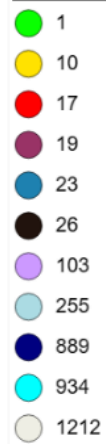
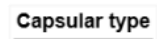
- Ia
- Ib
- II
- III
- IV
- V
- VI
- VII

Clonal complexes

- 1
- 10
- 17
- 19
- 23
- 26
- 103
- 255
- 889
- 934
- 1212



f): Mozambique



g) : Nigeria

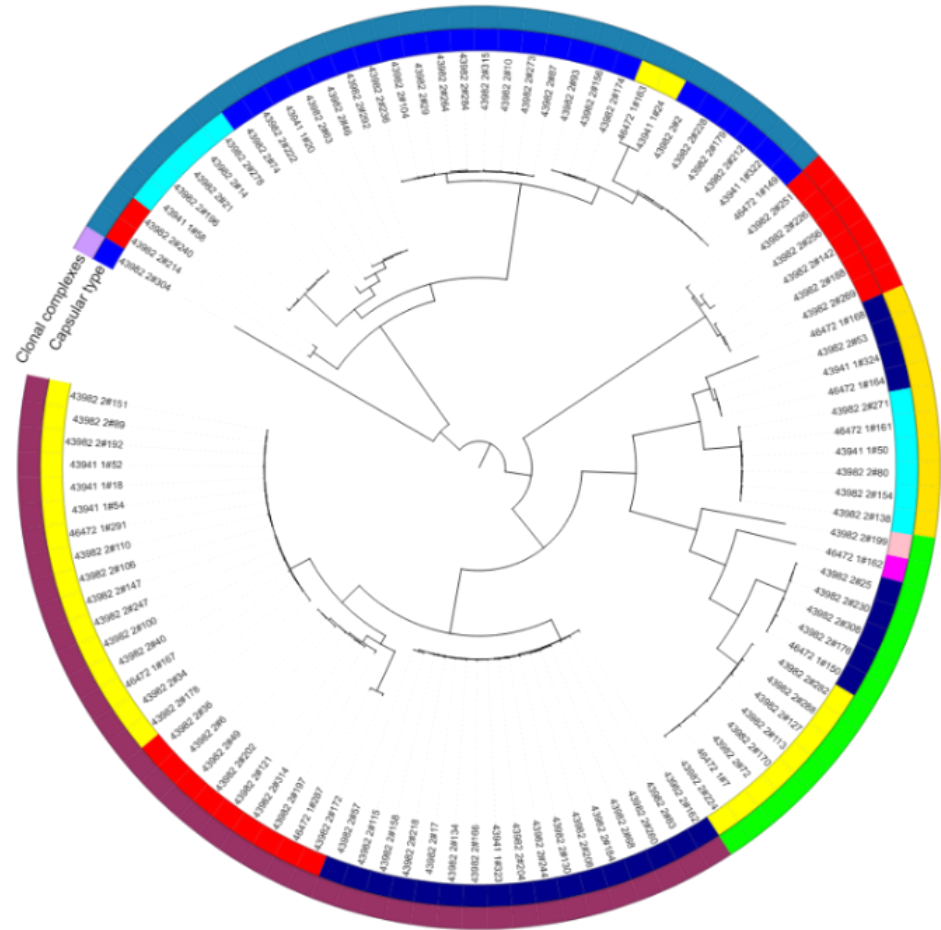
Tree scale: 0.001

Capsular type

- Ia
- Ib
- II
- III
- IV
- V
- VI
- VII

Clonal complexes

- 1
- 10
- 17
- 19
- 23
- 26
- 103
- 255
- 889
- 934
- 1212



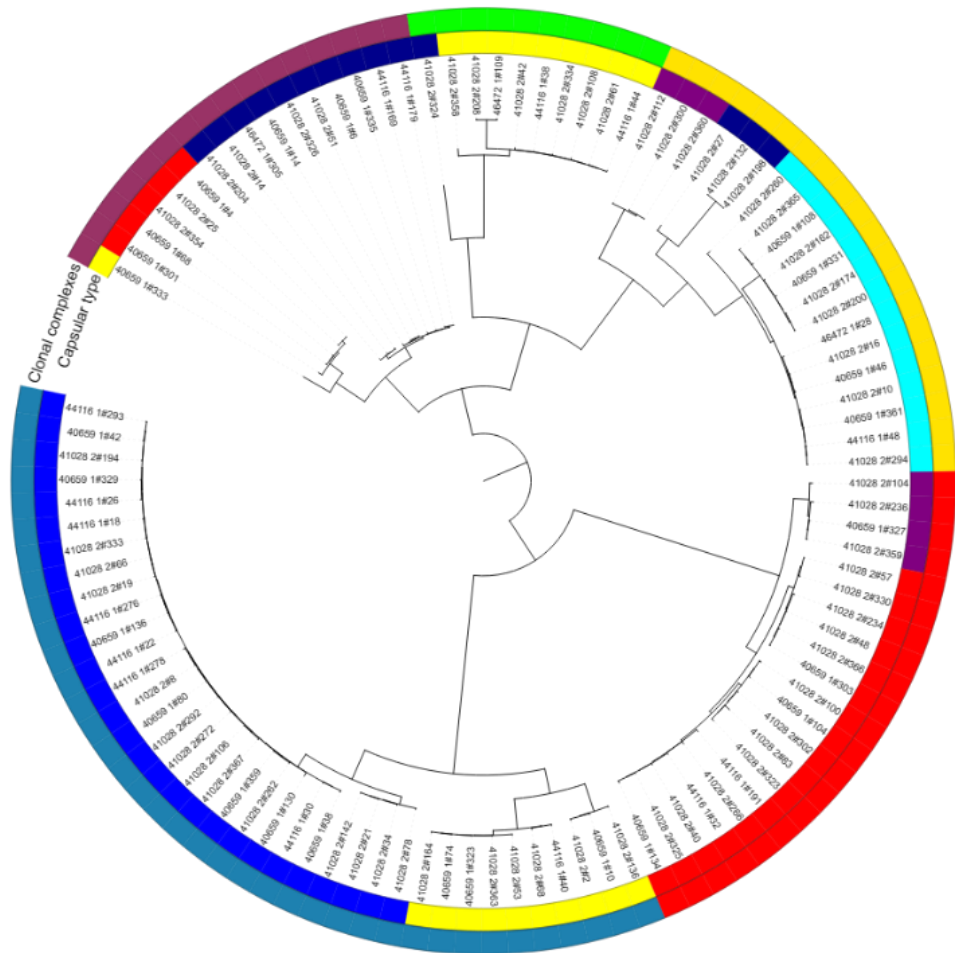
h): South Africa

Tree scale: 0.001

Capsular type

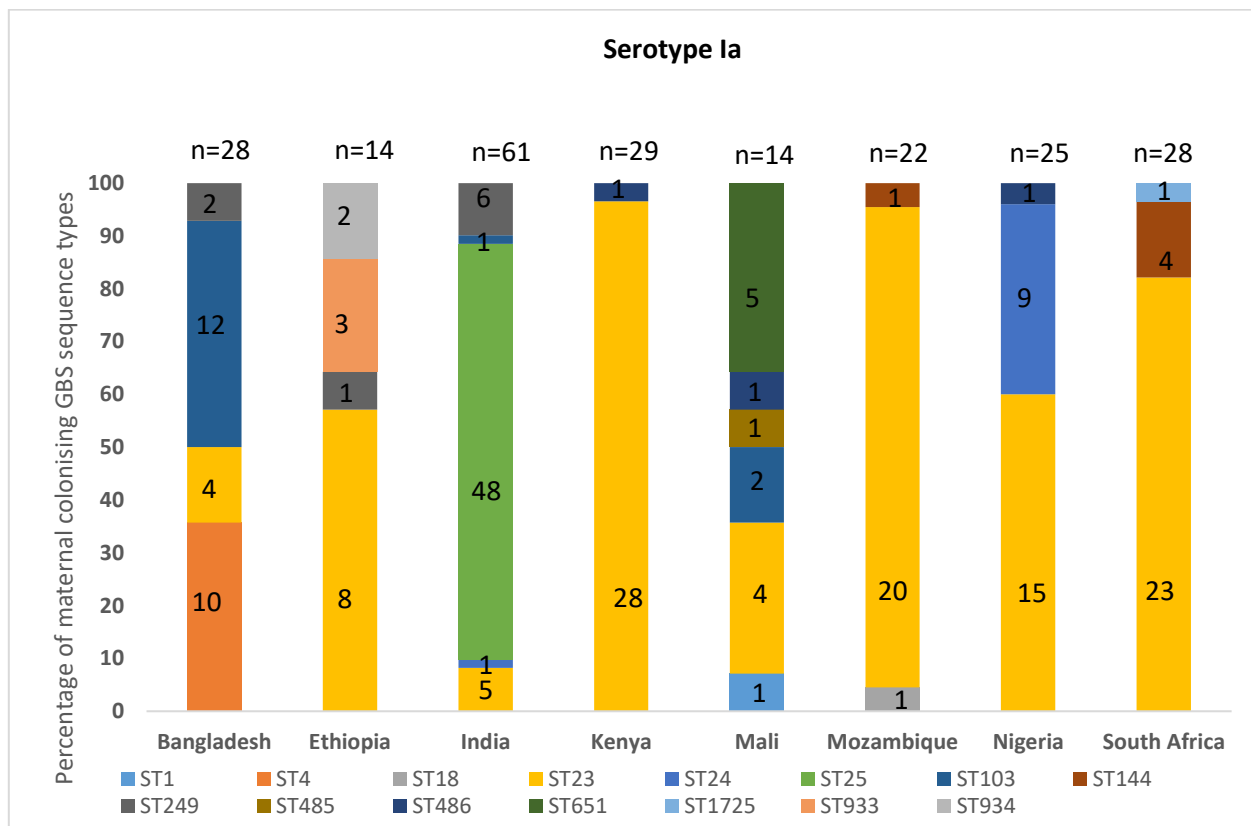


Clonal complexes

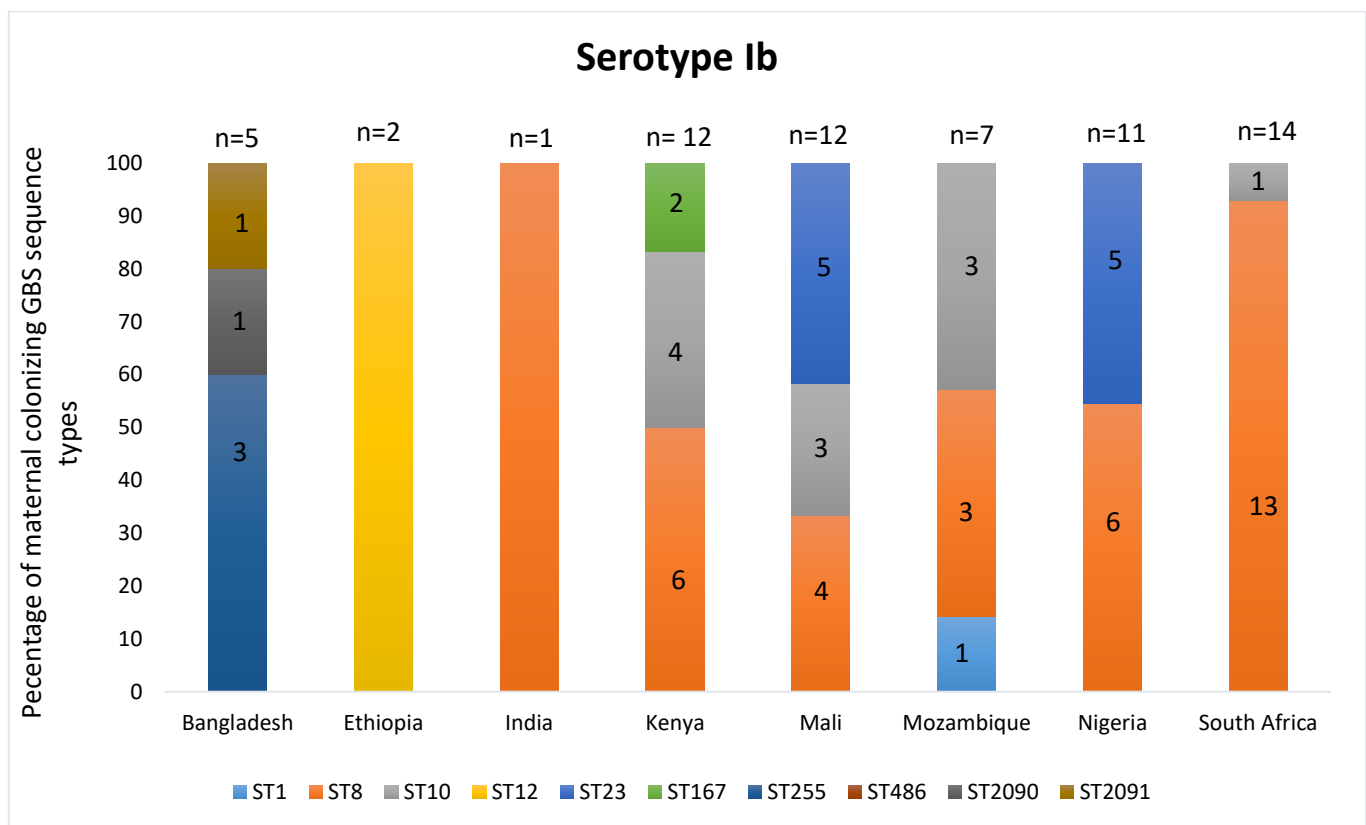


Appendix 4: Serotypes associated STs in maternal colonizing GBS isolates from sub-Saharan Africa and Southeast Asian countries.

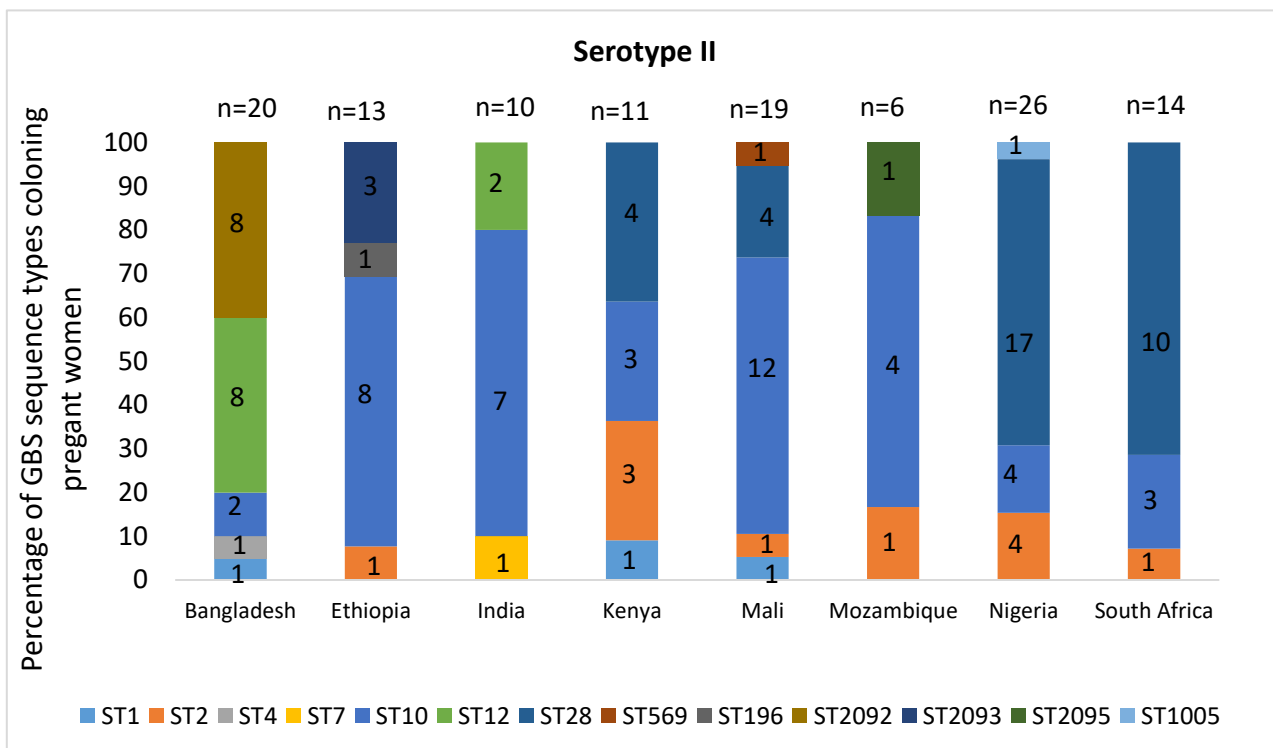
a)



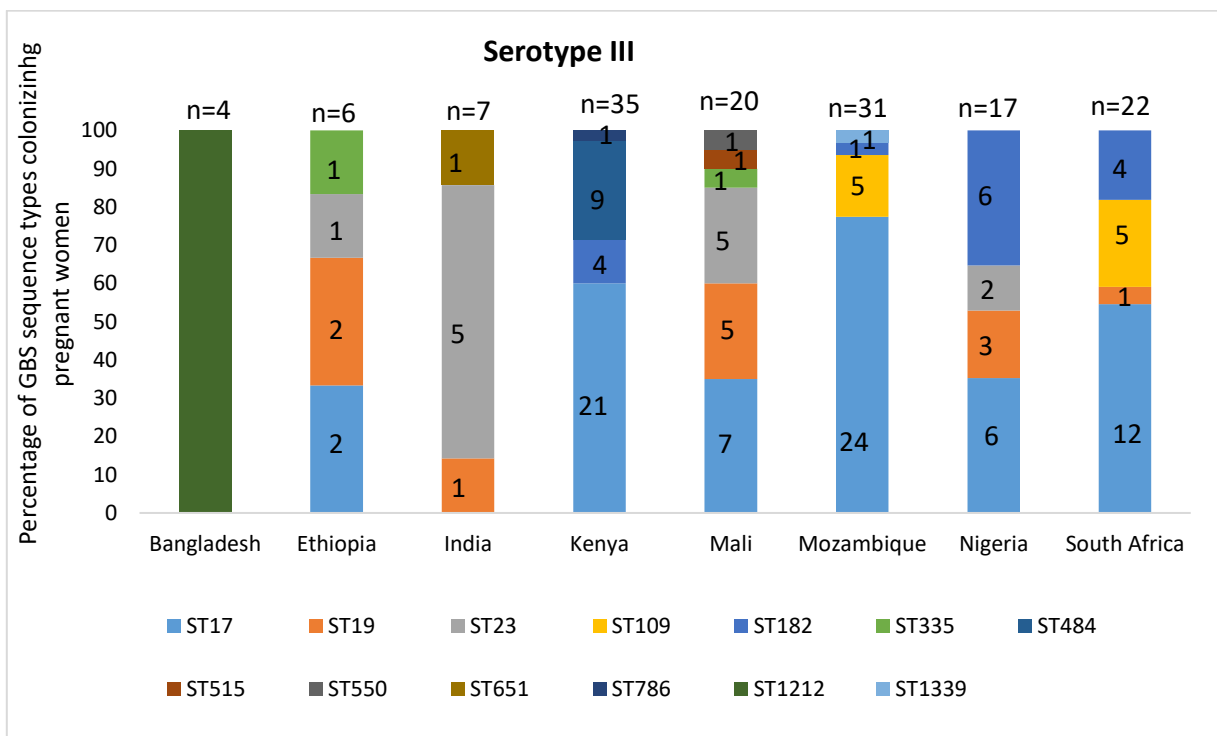
b)



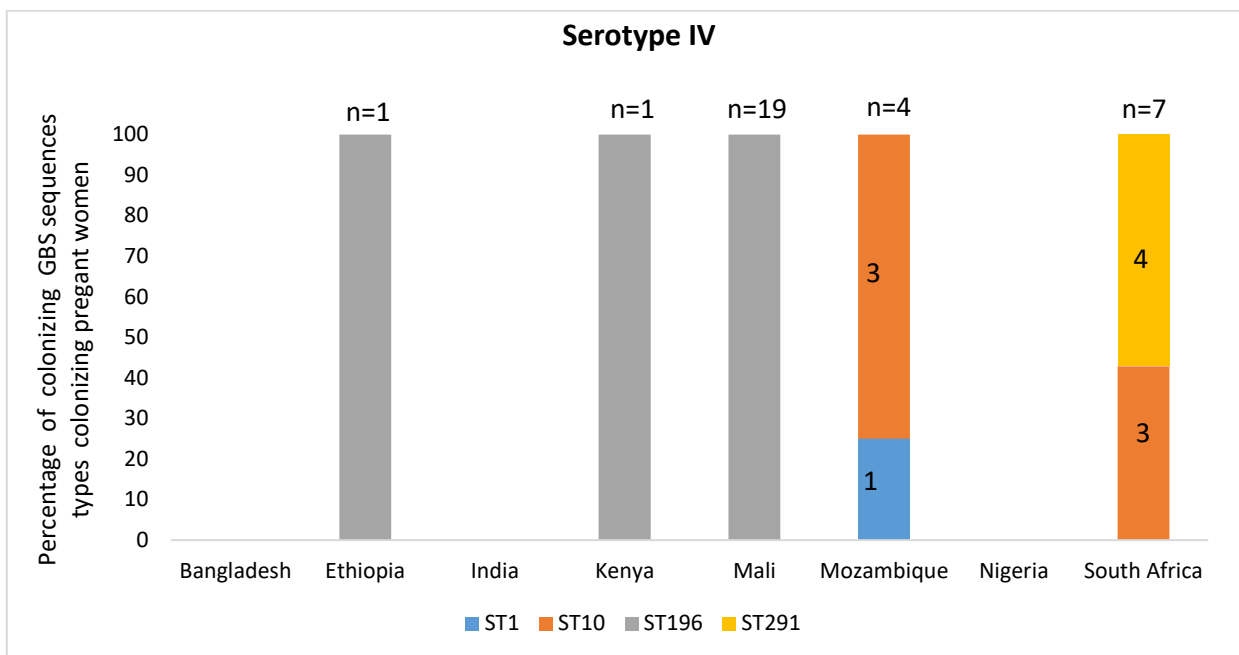
C)



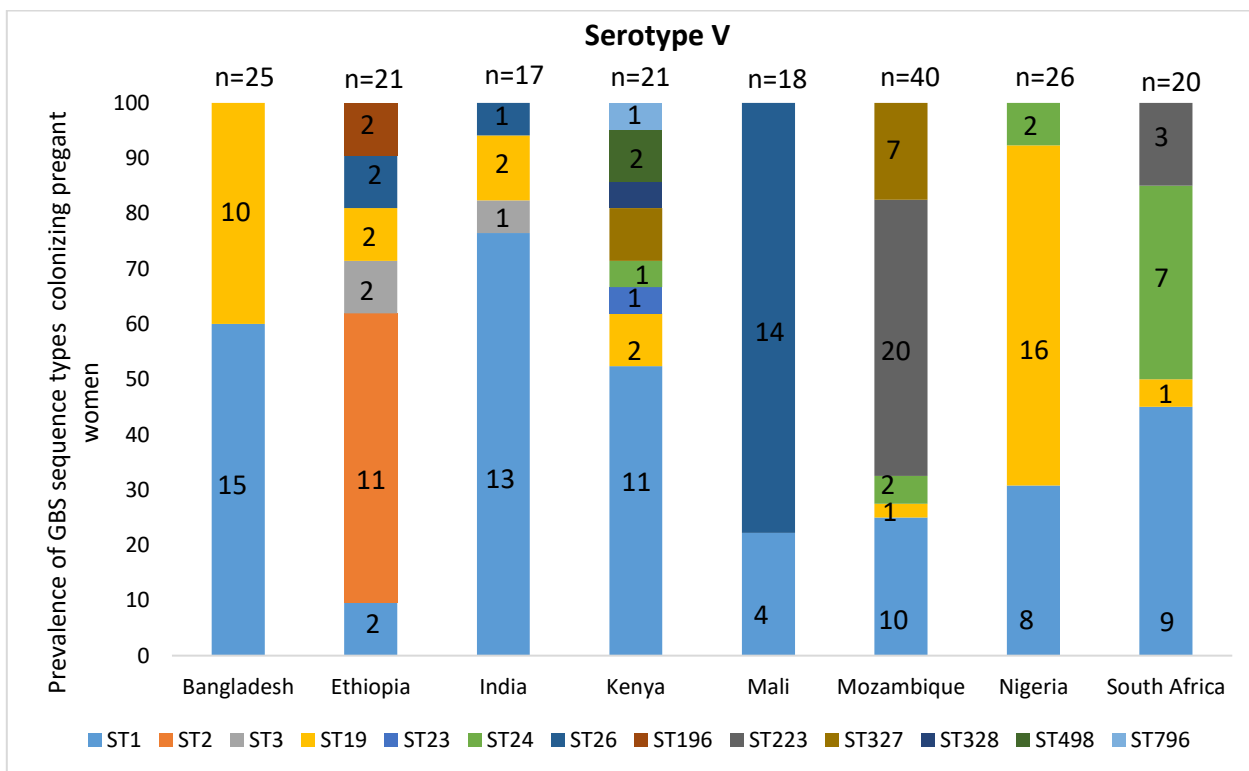
d)



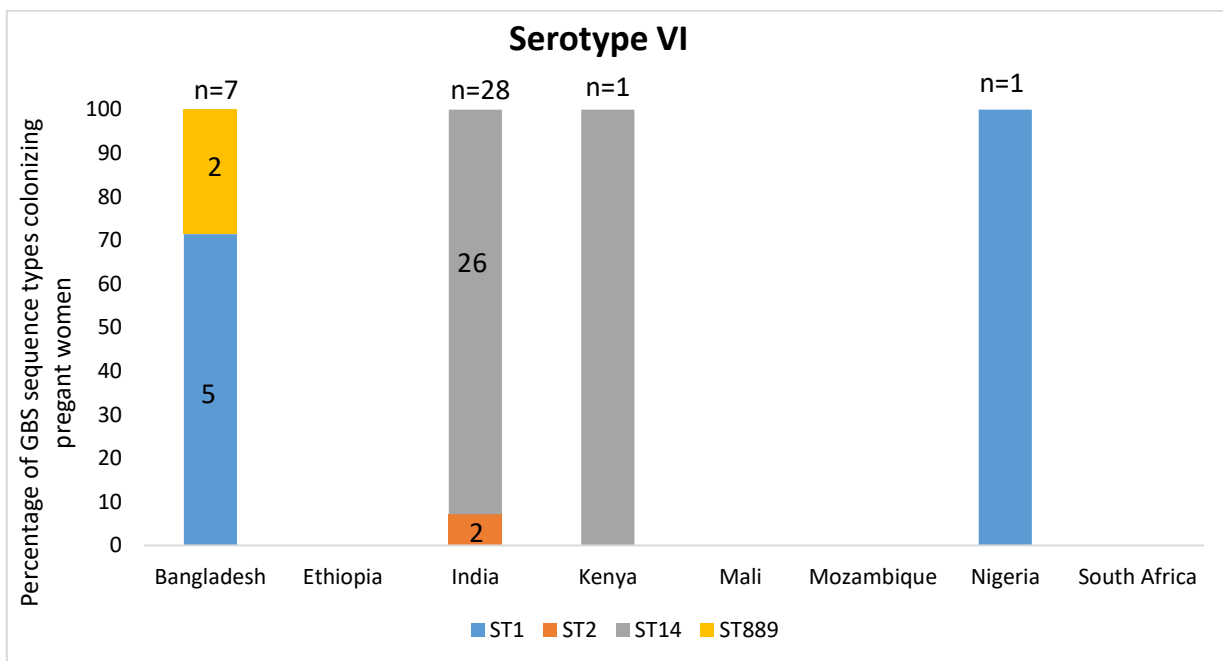
e)



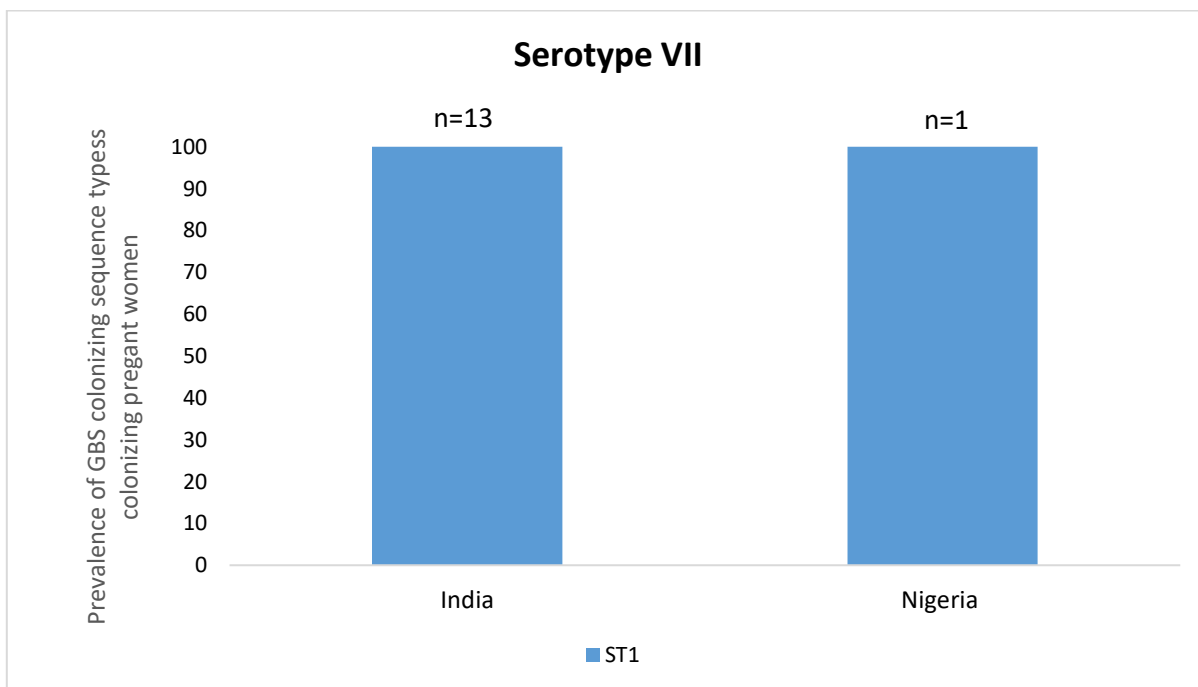
f)



g)

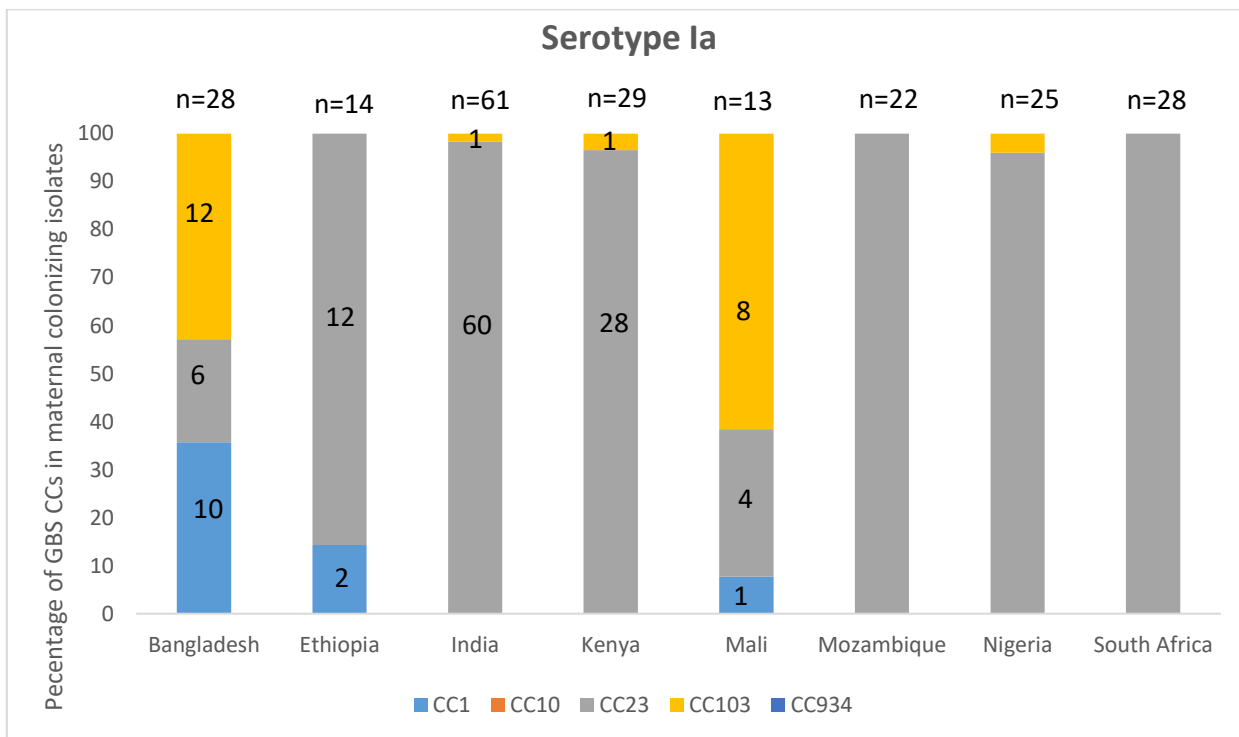


h)

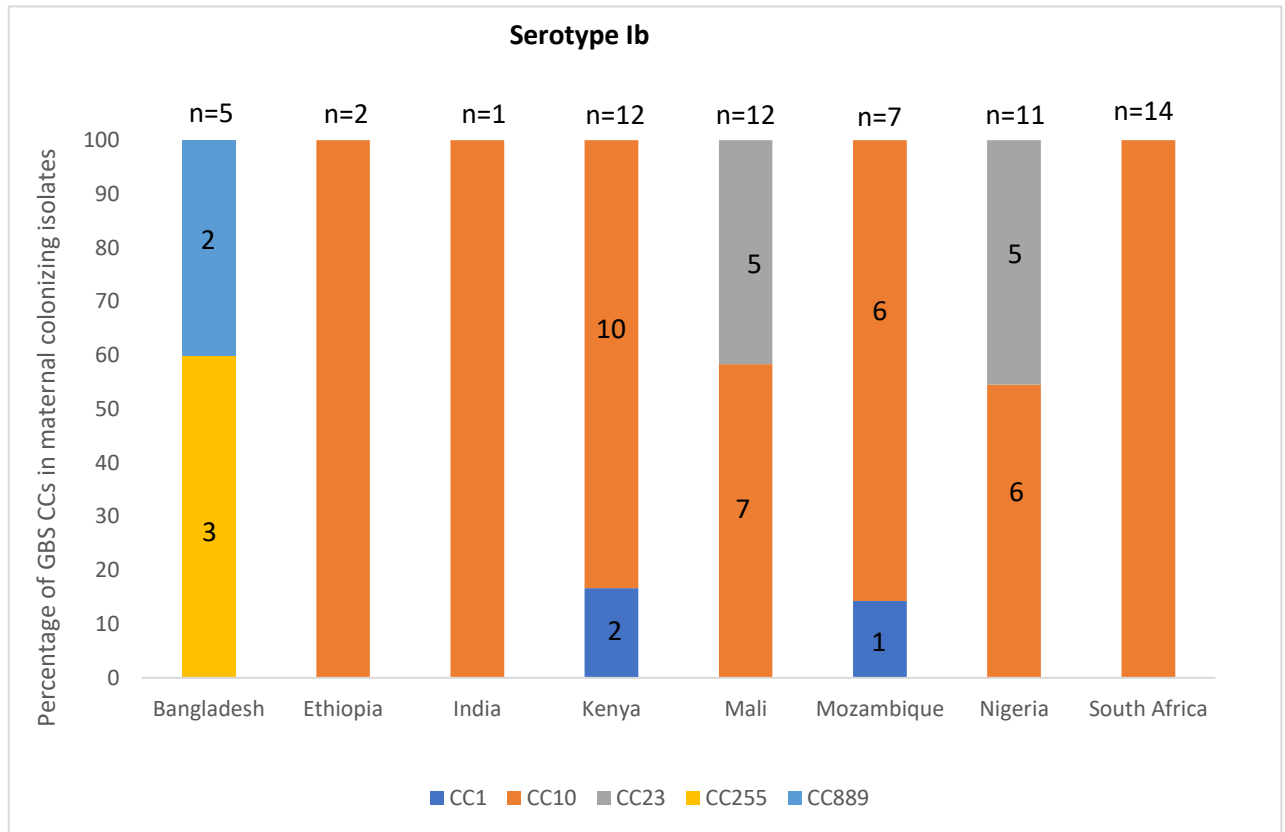


Appendix 5: The serotype associated CCs in maternal colonizing GBS isolates from sub-Saharan Africa and Southeast Asian countries.

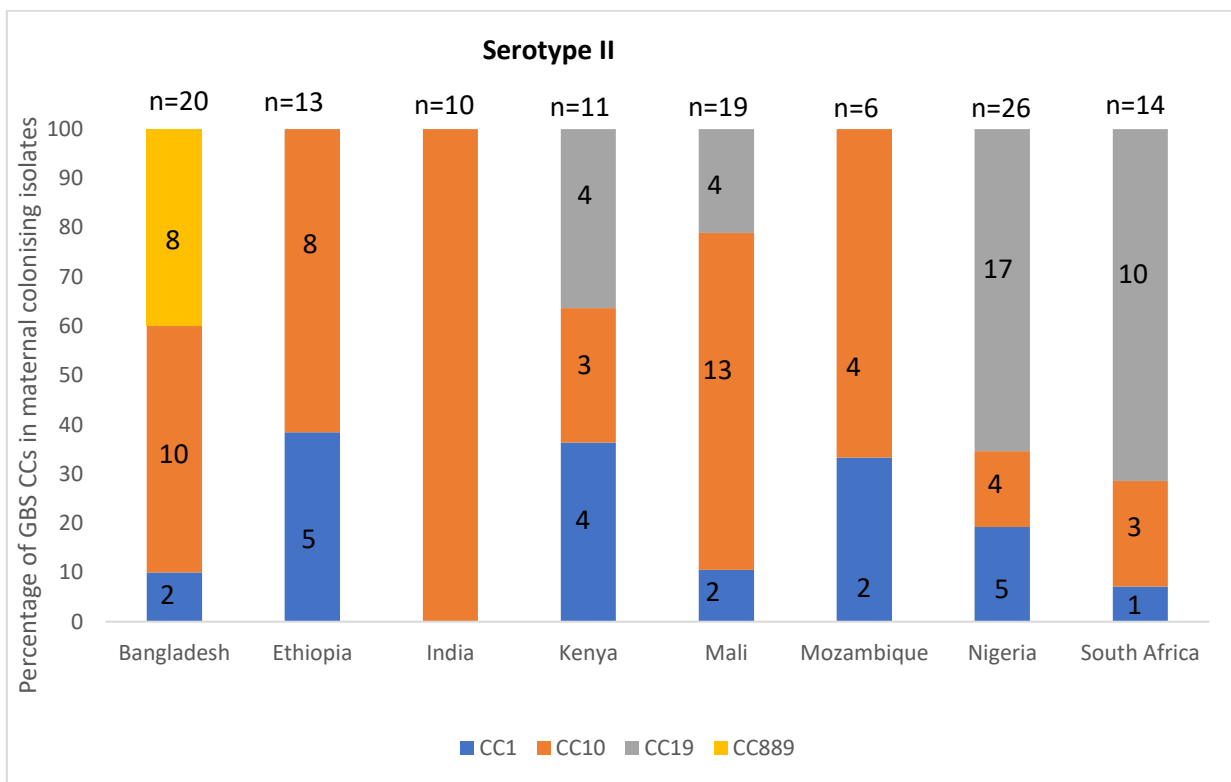
a)



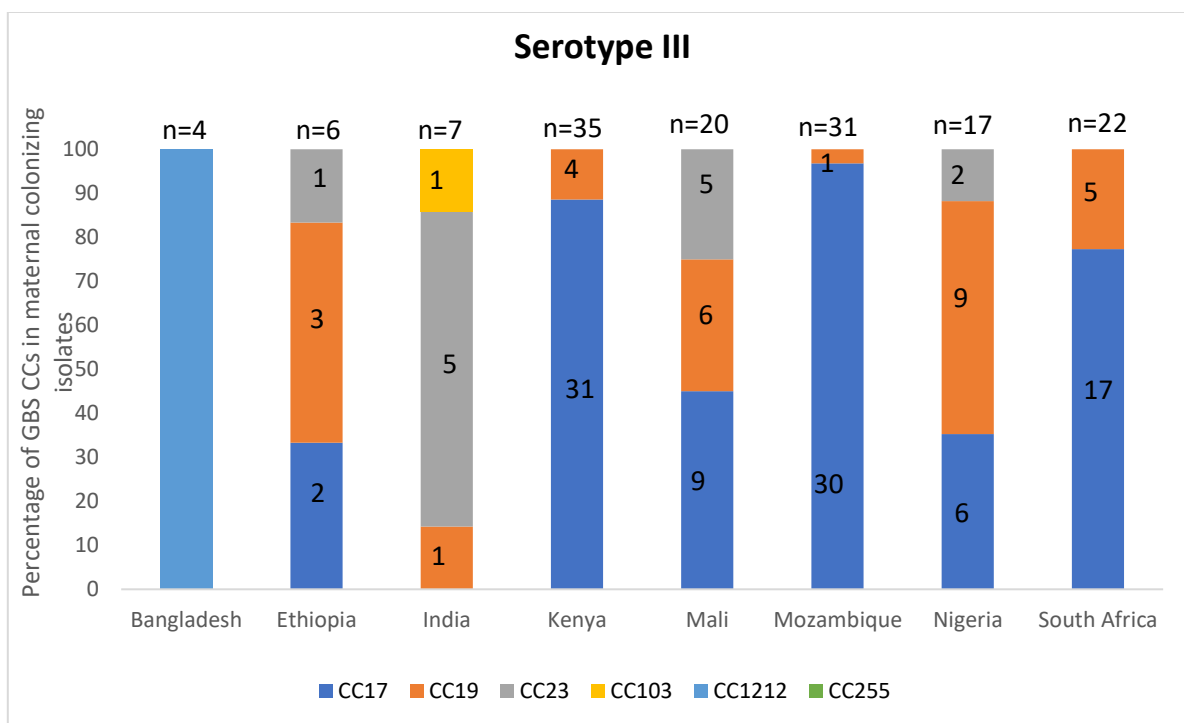
b)



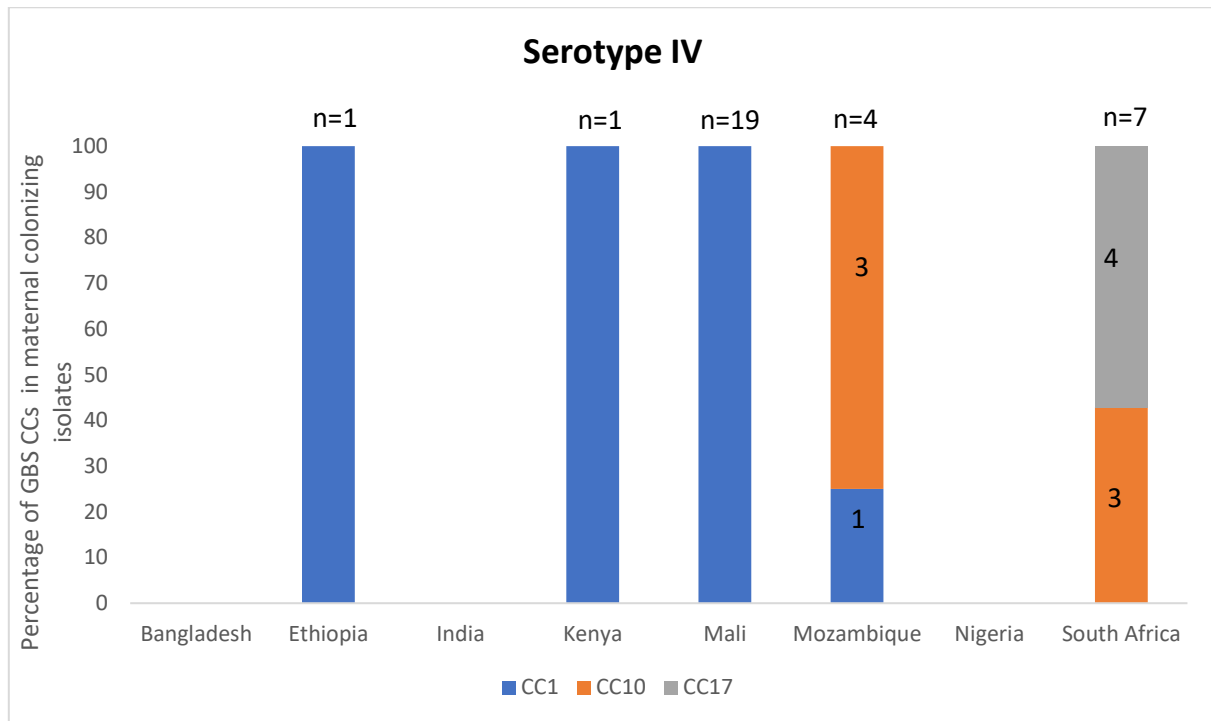
c)



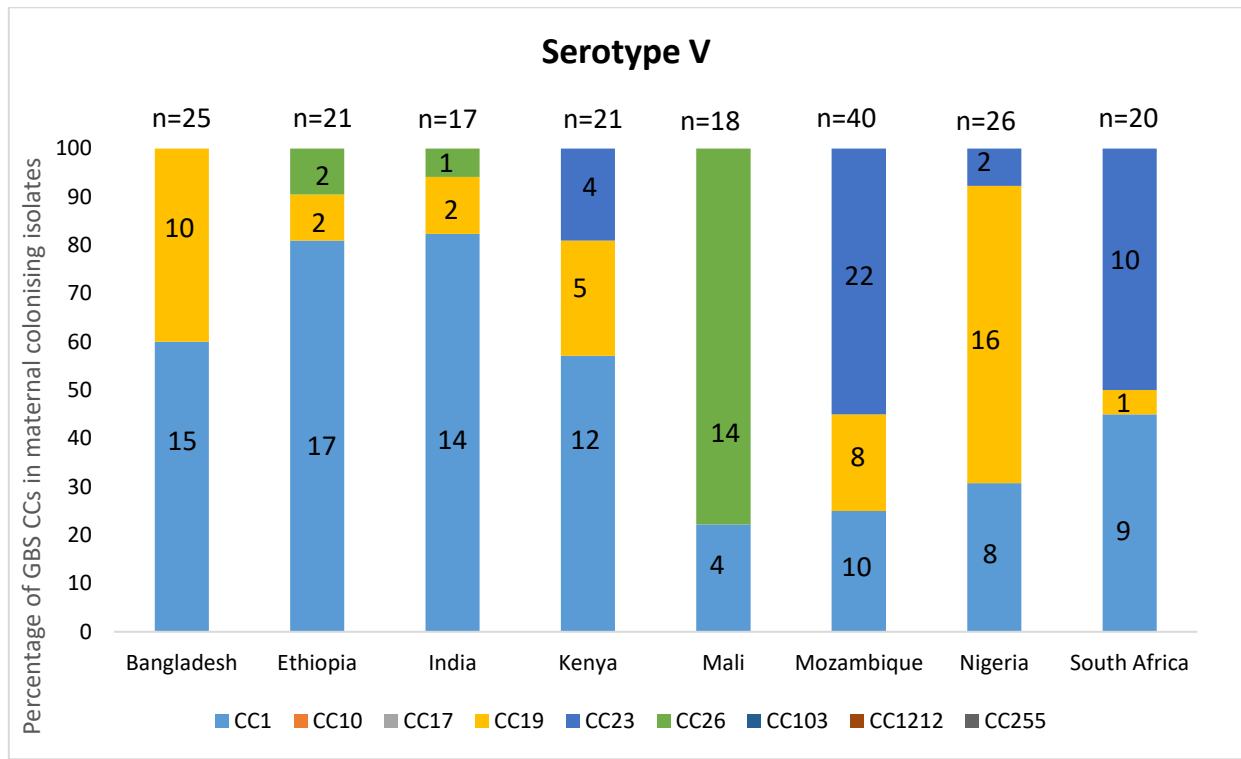
d)



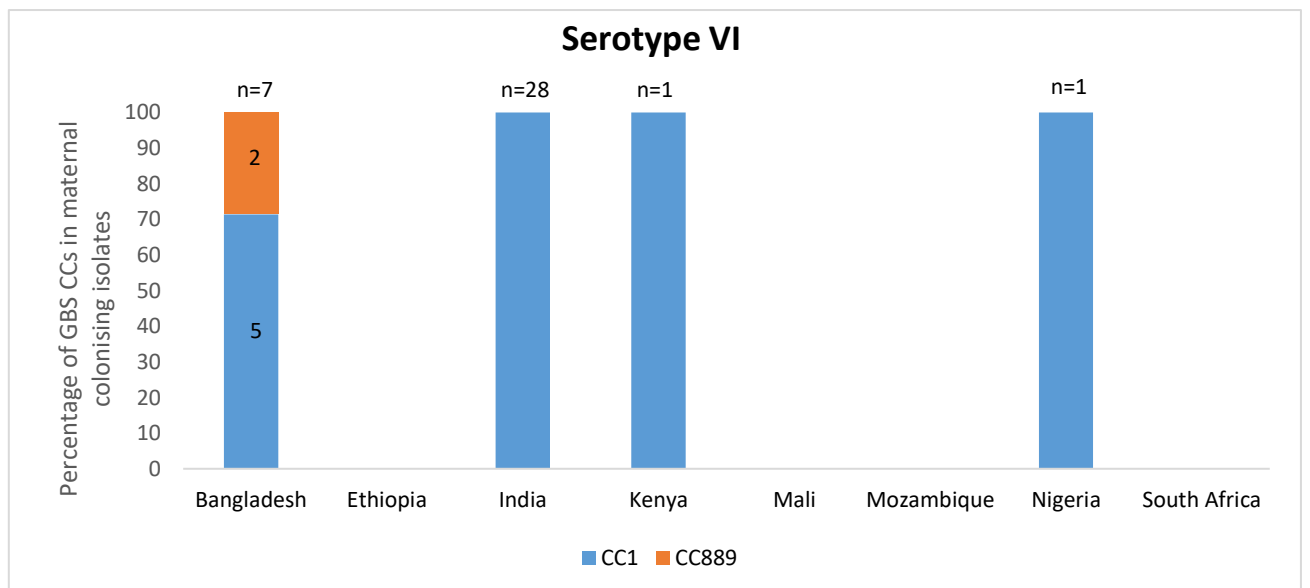
e)



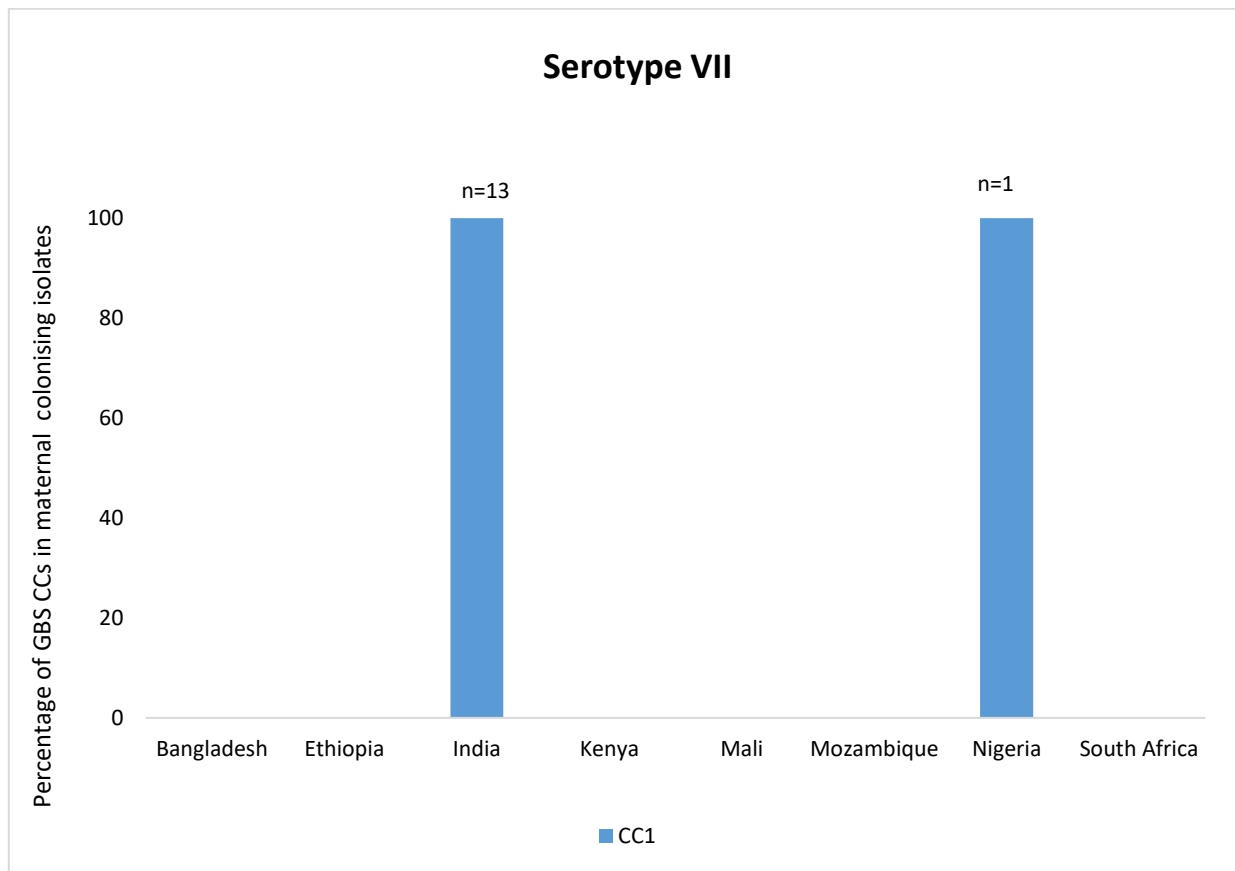
f)



g)

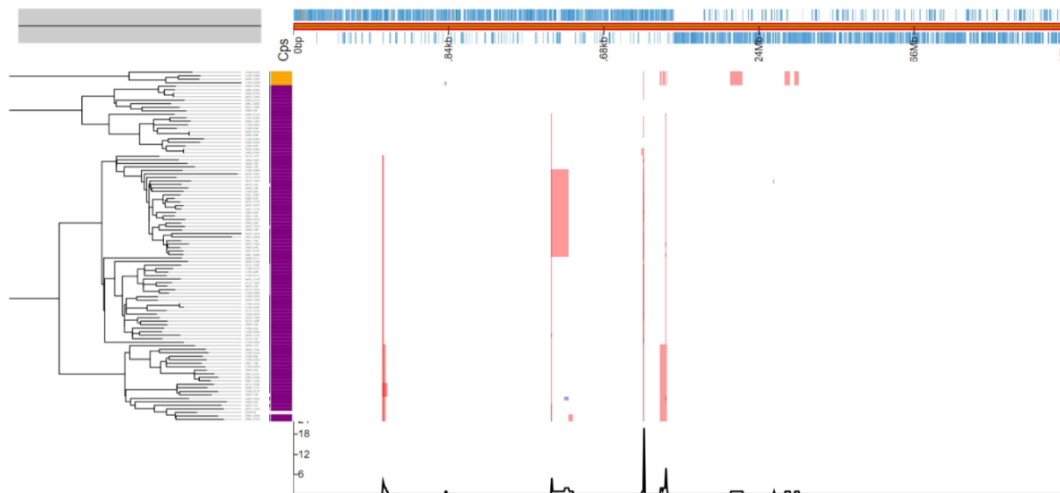


h)

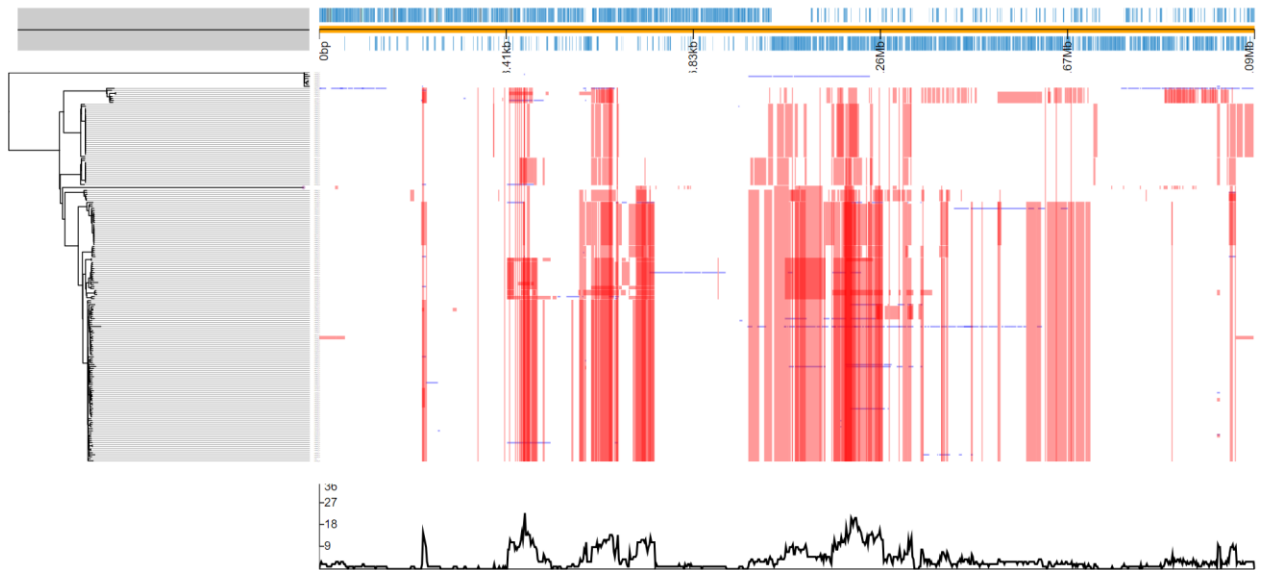


Appendix 6: Phylogenetic relationship of the major CCs and recombinational analysis of the genomes of each CC. The red block shows recombination in more than one isolates and recombination within a single isolate is represented in blue. The top blue blocks represent genomic regions within the reference genomes and the phylogenetic tree of the isolates in each of the major CCs is shown **a)** Recombination blocks within CC17 isolates. The phylogeny is annotated with capsular serotype, with serotypes III and IV indicated by purple and gold color, respectively. **b)** Recombination blocks indicated in red in each of the major CCs; CC1 c) CC10, d) CC19 and e) CC23.

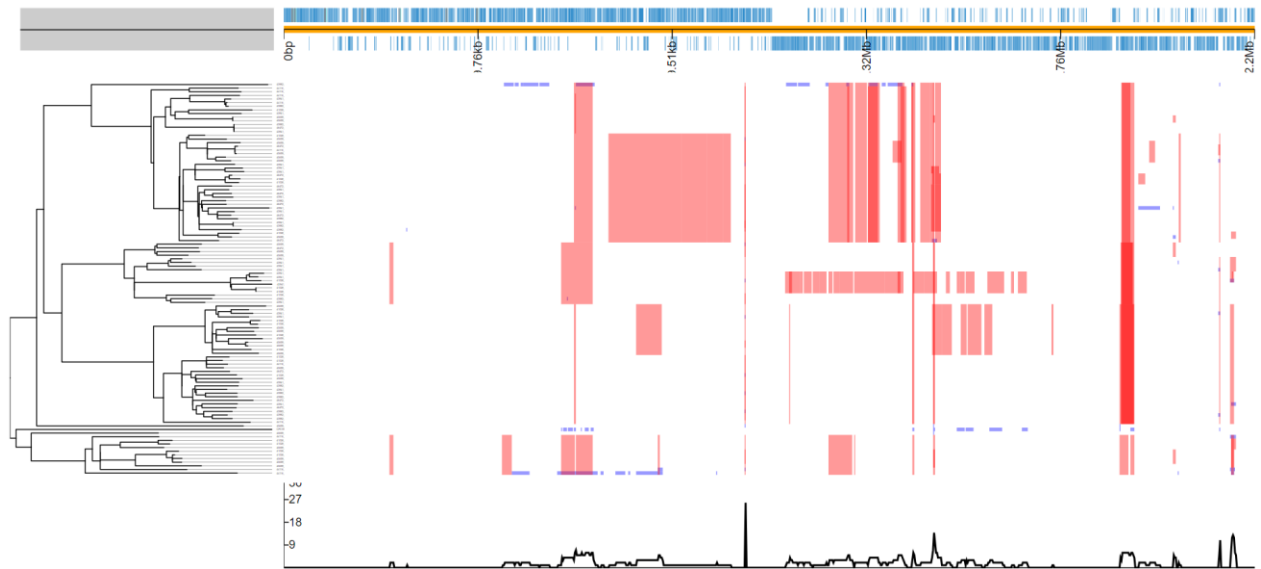
a) Recombination blocks in CC17 isolates



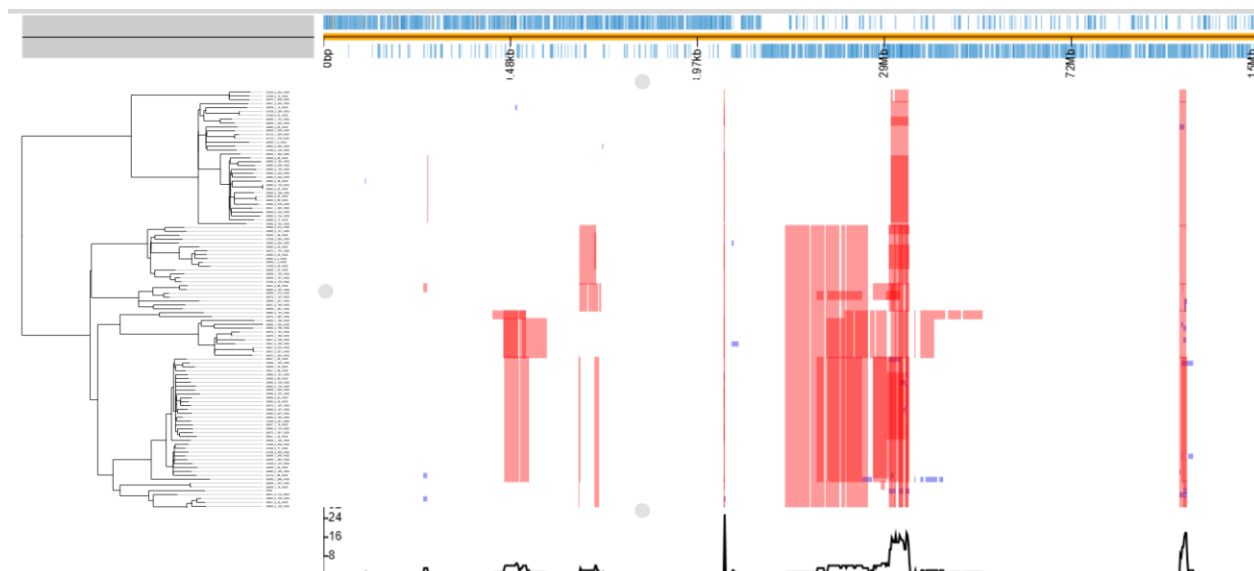
b) Recombination blocks in CC1 isolates



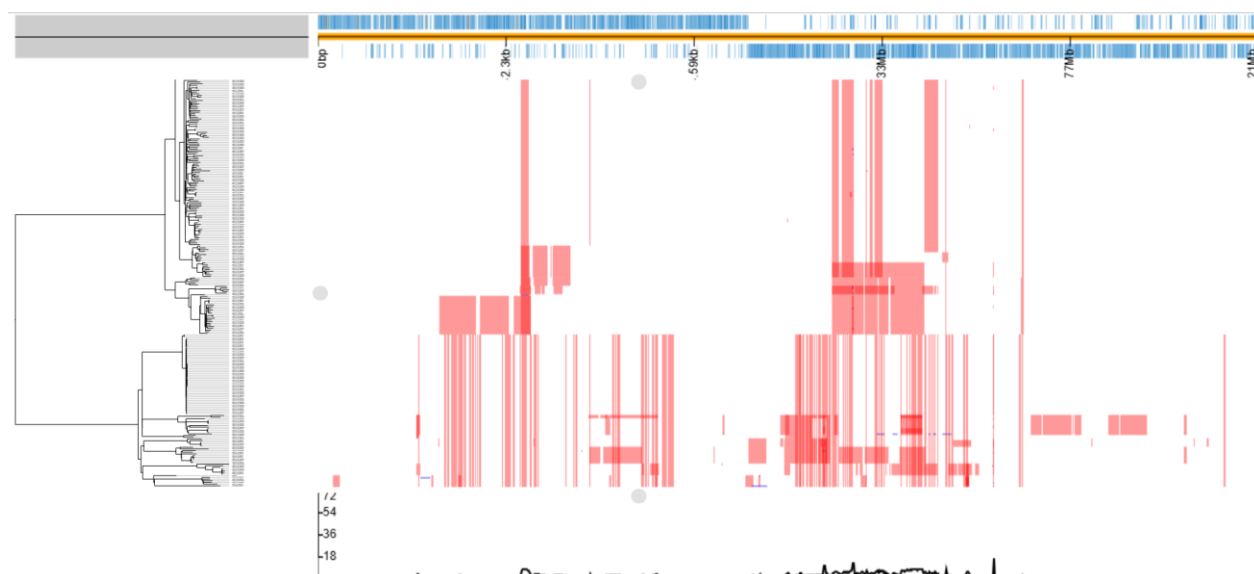
c) Recombination blocks in CC10 isolates



d) Recombination blocks in CC19 isolates



e) Recombination blocks in CC23 isolates



Appendix 7: Prevalence of antimicrobial resistance gene targets in colonizing GBS isolates in pregnant from 8 low and middle-income countries of Southeast Asia and Africa

