
# Detailed genomic and antimicrobial resistance comparison of UK *Streptococcus agalactiae* isolates from adults to those of diverse global origins

By

Uzma Basit Khan

Master of Engineering, Biomedical Engineering

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy (PhD)

Jul 2022

Division of Infection and Immunity

School of Medicine

Cardiff University

United Kingdom

# SUMMARY

Keywords: *Streptococcus agalactiae*, adults, antibiotic susceptibility, capsular switching

Invasive group B Streptococcus (GBS), a leading cause of illness and death among infants in the first week of life is also an important infectious agent able to cause invasive infections in adults. Serious life-threatening invasive GBS infections are increasingly recognized in the elderly and individuals compromised by underlying diseases such as diabetes, cirrhosis, and cancer. The significance of GBS as a cause of severe infections among adults is not widely appreciated. In adults, the modes of transmission and acquisition are less identified. Penicillin is the antibiotic of choice for treatment of GBS infection however, resistance to multiple antibiotics is increasing in Europe and worldwide in these organisms, making them increasingly difficult to treat, but also making them a potential danger as a silent resource for donating resistance genes to more aggressive pathogens. To date, there are no guidelines for the prevention of adult GBS disease; vaccines in development may hold promise. GBS emerged rapidly all over the world during the 1970s to become the leading cause of neonatal sepsis. However, several reports in recent years suggest that the incidence of GBS disease is also increasing among adults. The driving force behind this change has not been fully explained, and recent trends in disease incidence in adults have not been characterized in any systematic reviews due to inadequate data available on adult GBS disease.

The aim of the work presented in this thesis was to characterize the population structure of human source isolated invasive and non-invasive GBS from adults in United Kingdom, and to evaluate the genetic diversity of isolates recovered from invasive disease in adult patients of Brazil and from vaginal and rectal colonization in Australian pregnant women with ≤22 weeks gestations to contribute to the global epidemiology of GBS and our understanding of GBS population biology. For this a set of conventional techniques were employed including antimicrobial susceptibility testing using disc diffusion test, serotyping using software Geneious (Biomatters ltd., New Zealand) and published

primers for ten different GBS capsular types, multilocus sequence typing (MLST) and surface protein gene profiling using SRST2 v2.2. In combination, these methods allowed the identification of the main genetic lineages circulating in UK, Brazil, and Australia, providing the means for an appropriate comparison of Brazilian and Australian GBS to UK GBS population. In addition, analysis of WGS to determine GBS potential to switch capsule, antimicrobial resistance (AMR) genes associated with mobile genetic elements (MGEs), examining relatedness of the GBS strains using whole genome phylogeny and correlate serotyping, patient age group to AMR genes and pan genome wide association study (pan-GWAS) on UK, Brazil, and publicly available genomes of GBS isolates from Canada and the United States was performed. These analyses revealed a reasonable number of isolates with potential capsular switch including some cases switched from current vaccine (under trail) covered capsular type to non-vaccine covered serotypes. Further a diverse group of MGEs were identified with a capacity to disseminate the resistance phenotype, the association between strains clustered in a group based on patients age group they isolated from and the AMR genes they carry were studied. An intense pan-GWAS was performed to discover Clonal Complex (CC) specific genes that may play role in increased colonization, invasiveness, pathogenicity, and better survival of GBS in the host cell.

The thesis work started with investigating 193 clinical GBS strains isolated from adults submitted to the UK national reference laboratory (179 invasive; 13 non-invasive; 1 with no information provided) for capsule type, MLST, presence of virulence factors, antimicrobial resistance genes, phylogeny, and genetic recombination. The genetic lineages defined by MLST identified very diverse populations but consistent in terms of serotypes prevalence and clonal structure identified previously in GBS invasive disease in United Kingdom. The prevalence of serotype III in this population, regardless of age, highlighted the importance of this serotype in GBS pathogenesis as a leading cause of invasive infections in adults. Macrolide resistance is disseminated in UK by both a multiclonal mechanism resulting from the spread of resistance genes throughout most serotypes and genetic backgrounds, as well as by clonal expansion of specific lineages, such as the serotype V ST1/*alp*3. Attachment and invasion of host cells are key steps in GBS pathogenesis, strong associations were identified between

serotypes and virulence genes, such as serotype V/*alp3*, serotype II and III/*bca*+*cba,* serotype Ia/*bib*A predominantly clustered in CC1, CC8/CC10 and CC23, respectively, whereas serotype III/*rib* clustered in CC17 and CC19 demonstrating GBS strains belonging to a particular CC differ in their abilities to attach and invade to host cell types and express key virulence genes that are relevant to the disease process. A major finding includes a high number of capsular serotype-CC mismatches (14/179, 7.8%) iGBS identified with a concerning recombination of hypervirulent *hvg*A core genome expressing a non-vaccine covered serotype IV capsule. The mechanism for these genetic transfer events involved the replacement of the whole capsular locus instead of the previously proposed genetic transfer of only the serotype specific genes.

The consequent analysis of MGEs carrying multidrug resistance genes in 41/193 GBS isolates revealed a diverse group of MGES used three different insertion sites (*rum*A, *rpl*L and *rps*I) to disseminate phenotypic resistance in GBS isolated from adult patients of United Kingdom. Out of 41 isolates, only one isolate carried the macrolide resistance (*erm*T) gene was on a plasmid, while for 4 isolates fluoroquinolone resistance was mediated by double point somatic mutation in *par*C and *gyr*A; for all other isolates ARGs were acquired by MGEs including five novel MGEs identified in this study, including ICE*Sag*84 and ICE*Sag*100414 carrying *erm*A alone, ICE*Sag*662 containing *erm*B, *tet*S, *ant*(6-Ia) and *aph*(3'-III), ICE*Sag*71 carrying *erm*B and *tet*O, and ICE*Sag*139 containing *erm*A and the high gentamicin level resistance gene *aac*(6')-*aph*(2''). The Tn*916* and Tn*5801* belonging to Tn*916*/Tn*1545* family harboured majority of *tet*M genes (88%, 154/175) found in UK tetracycline resistant GBS isolates and were significantly associated to CC1, CC8/10, CC19 and CC17 and CC23 isolates, respectively suggesting these ICEs are clonally related, acquired through limited and rare insertion events and led to expansion of these lineages, also supporting the earlier interpretations [1]. In addition, the Tn*916*/Tn*1545* family ICE identified in this study were mostly integrated at two insertion sites, adjacent to two target genes (TGs) – *rum*A for Tn*916* and *gua*A for Tn*5801* demonstrating their integration preferences at hot spot regions. The high prevalence of MGEs carrying ARGs in UK adults

GBS isolates implied that in GBS, these MGEs probably act like a reservoir of ARGs, and play a central role in the dissemination of resistance genes via horizontal gene transfer.

Antibiotic susceptibility testing (AST) of GBS isolates recovered from UK (n=193) and Brazil (n=26) adults, and from vaginal and rectal sites of Australian pregnant women (n=171) revealed that all tested GBS populations were sensitive to ampicillin, vancomycin, and gentamicin, except a single UK strain that conferred high level gentamicin resistance through aminoglycoside modifying enzyme encoding gene *aac*(6')-*aph*(2"). Further all UK and Brazil GBS isolates carried five penicillin binding protein (PBPs) types with amino acid substitutions that did not appear to be associated with decreased β-lactam susceptibility suggesting penicillin as the first choice and vancomycin as the second choice of drug for GBS disease treatment as currently recommended by Centers for Disease Control and Prevention (CDC) and Royal College of Obstetricians and Gynaecologists (RCOG) guidelines [2, 3]. Accelerating resistance rates to erythromycin and clindamycin were observed in UK, which in comparison found less in Australian pregnant women colonized GBS and Brazilian GBS populations suggesting routine susceptibility testing of erythromycin and clindamycin for penicillin allergic patients to ensure effective treatment. A very few UK and Australian GBS strains were found resistant to chloramphenicol and/or levofloxacin while all Brazilian GBS were susceptible to these two antibiotics. High resistance rate to tetracycline was detected in UK, Brazilian and Australian GBS isolates and the *tet*M gene was found widespread and carried predominantly by Tn*916*/Tn*1545* family elements. Three distinctive variants of Tn*916* were observed in UK GBS strains, respectively including two Tn*916* variants carrying *tet*M as a single resistance gene with additional 9 orfs in conjugation module, while the third Tn*916* variant carried tetracycline efflux MFS transporter (*tet*L) gene near the *tet*M gene.

The pangenome wide association study (pan-GWAS) was conducted on 447 GBS genomes, including this study sequenced UK (n=193) and Brazil (n=26) GBS strains, in addition to deliberately selected serotype V and III publicly available genomes from Canada (n=134) and the United States

(n=94) to study genes specific to CC and their role in pathogenicity and invasiveness, since these two serotypes are universally found highly associated with adults and neonatal diseases, respectively. Each CC was characterized by specific genes that provides selective advantage to GBS for improved colonization, invasion, virulence, and survival within host. This analysis identified 97 CC-specific genes associated (excluding hypothetical proteins) with virulence, metabolism, and regulation of cellular mechanisms that may explain the differential virulence potential of the CCs. Among CC17 and CC23 GBS isolates, micronutrient uptake proteins (iron and manganese), two component systems, accessory secondary proteins, pilus and quorum-sensing genes were identified which were absent in less invasive lineages (CC1, CC8 and CC19). Metal resistance genes (arsenic, cadmium, and copper) and CRISPR associated genes (*cas*1/*cas*2) were confined to CC8 whereas the type IV secretory protein (VirD4) was significantly associated to CC19. Collectively this analysis underlines the lineage-specific basis of GBS niche adaptation and virulence.

In summary, in this thesis GBS shown to have an evident stable clonal structure both temporally and geographically. Interestingly, capsular switching occurred across multiple serotypes and among strains with dissimilar genomic backgrounds in high numbers demonstrated ongoing GBS diversification due to recombination and highlights the importance of ongoing surveillance of GBS and may have implications for vaccine development strategies. Regardless of increasing information on invasive disease and maternal colonization, a thorough understanding of colonization in adults and natural reservoirs of GBS is required for the appropriate management of the GBS infections.

# ABBREVIATIONS

| | |
|---|---|
| ACOG | American College of Obstetricians and Gynaecologists |
| Alp | Alpha-like protein |
| AST | Antibiotic susceptibility testing |
| bp | Base pairs |
| CDC | Centres for Disease Control and Prevention |
| CC | Clonal complex |
| CLSI | Clinical Laboratory Standards Institute |
| CPS | Capsular polysaccharide |
| DLV | Double-locus variant |
| DNA | Deoxyribonucleic acid |
| EOD | Early-onset disease |
| GBS | Group B streptococci |
| HGT | Horizontal gene transfer |
| *hvg*A | hypervirulent gene |
| IAP | Intrapartum antibiotic prophylaxis |
| ICE | Integrative and conjugative element |
| IS | Insertion sequence |
| Kbp | Kilobase pairs |
| LOD | Late-onset disease |
| LSA | Lincosamide-streptogramin A (resistance phenotype) |
| Mbp | Megabase pairs |

| | |
|---|---|
| MIC | Minimum inhibitory concentration |
| MLSB | Macrolide-lincosamide-streptogramin B (resistance phenotype) |
| MLST | Multilocus sequence typing |
| MGE | Mobile genetic element |
| PHE | Public Health England |
| NT | Non-typeable |
| ORF | Open reading frame |
| PBP | Penicillin-binding protein |
| pan-GWAS | Pangenome wide association study |
| PCR | Polymerase chain reaction |
| PFGE | Pulsed-field gel electrophoresis |
| PI | Pilus-island |
| RCOG | Royal College of Obstetricians and Gynaecologists |
| QRDR | Quinolone resistance determining region |
| RNA | Ribonucleic acid |
| SLV | Single-locus variant |
| TLV | Triple-locus variant |
| ST | Sequence type |

# ACKNOWLEDGEMENTS

Alhamdulillah, I praise and thank Allah SWT for His greatness and for giving me the strength and courage to complete this thesis.

Undertaking this PhD has been a truly life-changing experience for me, and it would not have been possible to do without the support and guidance that I received from many people. First and foremost, I offer my sincerest gratitude to my supervisor, Dr. Brad Spiller who has supported me throughout my thesis with his patience and expertise while allowing me the room to work on my own way. I attribute my PhD degree to his encouragement and efforts and without him this thesis would not have been completed. His advice on my research as well as on my writing have been priceless. May GOD bless him abundantly.

Deepest gratitude is also for my co-supervisor Dr. Elita Jauneikaite, Research Fellow in Bacterial Genomics and Epidemiology at Imperial College London, and Robert Andrews, Lead Bioinformatician at Cardiff University who have been a tremendous mentor for me and provided training on various bioinformatic tools utilized in this thesis.

My parents, Muhammad Anwer Nadeem and Tasleem Akhter deserves special mention for their inseparable support and prayers. I am much thankful to them for giving me strength to reach for the stars and chase my dreams. I am also extremely thankful to my parents in law, Abdul Farooq Khan and Rubina Ashraf for their continuous support, prayers, and words of appreciations. Words fail to express my gratitude to my husband, Abdul Basit Khan whose dedication, love, and confidence in me, has taken the load off my shoulder. My son Ibrahim for being so calm, caring, and helpful kid.

Many thanks to Cardiff University to honour me to be one of the 2016-year recipients of the International Scholarship Fund. I would like to take this opportunity to thank you for your generosity in funding my PhD and letting me pursue my dreams. Special thanks to Bill Gates Foundation for funding my bench fees.

Special thanks to Vicki Chalker, Head, respiratory and vaccine preventable bacterial reference unit at Public Health England (PHE), United Kingdom for arranging adults GBS samples collected from United Kingdom and proof-reading manuscripts, and extreme thanks to PHE for funding DNA extraction and sequencing. I would like to extend my sincere thanks to the efforts of hospital microbiology and surveillance staff across the country in referring specimens for laboratory surveillance and helping in DNA extraction and genomic sequencing.

Special thanks to Dr. Diego Andrey, Infectious Diseases Specialist at University of Geneva, Switzerland and Prof. Ana Gales, Professor, Department of Infectious Diseases at Federal University of São Paulo, Brazil for arranging adult GBS specimens collected from of Brazilian population (adults, neonates, and kid) and providing patients associated metadata. Further thanks to Dr Matthew S. Payne, Senior Research Fellow, Faculty of Health and Medical Sciences, Obstetrics and Gynaecology at the University of Western Australia for organizing colonized GBS vaginal and rectal e-swabs collected from Australian pregnant women.

Finally, I would like to thank everybody who was important to the successful realization of this thesis, as well as expressing my apology that I could not mention personally one by one.

## DEDICATION:

Dedicated to,

**My Father**,

For earning an honest living for us and for supporting and encouraging me to believe in myself.

**My Mother**,

A strong and gentle soul who taught me to trust Allah, believe in hard work and that so much could be done with little.

**And My Husband**

For your patience, love, and extreme hard work to make this dream come true.

# TABLE OF CONTENT

SUMMARY ..... i
ABBREVIATIONS ..... vi
ACKNOWLEDGEMENT ..... vii
DEDICATION ..... x
TABLE OF CONTENT ..... xi
LIST OF FIGURES ..... xvi
LIST OF TABLES ..... xix

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

**General Introduction ..... 1**

**1.1. Streptococci ..... 2**

**1.2. Beta-haemolytic streptococci ..... 3**
1.2.1. Group A streptococcus ..... 3
1.2.2 Group B streptococcus ..... 4
1.2.3 Group C and G streptococcus ..... 5

**1.3 Microbiological characterisation of GBS ..... 5**
1.3.1 Phenotyping methods ..... 6
1.3.1.1 Growth on Agar ..... 6
1.3.1.2 CAMP test ..... 7
1.3.1.3 Biochemical Profiling ..... 7
1.3.1.4 MALDI-TOF ..... 8
1.3.1.5 Serotyping ..... 9
1.3.1.6 Antimicrobial susceptibility testing (AST) ..... 9
1.3.2 Genotyping methods ..... 10
1.3.2.1 Restriction Fragment Length Polymorphism (RFLP) ..... 10
1.3.2.2 Pulsed-field gel electrophoresis (PFGE) ..... 11
1.3.2.3 Multilocus sequence typing (MLST) ..... 13
1.3.2.4 PCR-based gene profiling ..... 16
1.3.2.5 Whole genome sequencing (WGS) ..... 16

**1.4 Risk factors of GBS colonization and infections ..... 19**
1.4.1 Neonatal infections ..... 19
1.4.1.1 Early Onset Disease ..... 20
1.4.1.2 Late Onset Disease ..... 21
1.4.2 Infections in pregnancy ..... 22
1.4.3 Infections in non-pregnant adults ..... 22

**1.5 Colonization and Transmission ..... 23**

**1.6 Guidelines for Maternal GBS screening and IAP strategy in different countries ..... 24**
1.6.1 In the United States ..... 24
1.6.2 In United Kingdom ..... 26
1.6.3 In Brazil ..... 27
1.6.4 In Australia: ..... 27
1.6.5 Limitations of IAP ..... 28

**1.7 Antimicrobials effectiveness against GBS ..... 29**
1.7.1 β-lactams ..... 29
1.7.1.1 Mechanisms of β-lactam resistance in GBS ..... 30
1.7.2 Macrolides and Lincosamides ..... 30

1.7.2.1 Mechanisms of macrolide and lincosamide resistance in GBS....31
1.7.3 Aminoglycosides....32
1.7.3.1 Mechanisms of aminoglycoside resistance in GBS....33
1.7.4 Tetracycline....34
1.7.5 Quinolones....36
1.7.6 Chloramphenicol....36
**1.8 Prevention of GBS infections – an alternative to antibiotics....37**
1.8.1 CPS conjugate vaccines....38
1.8.2 Protein vaccines....39
**1.9 Virulence factors....40**
1.9.1 GBS adhesion and invasion into host cells....42
1.9.2 Proteins involved in immune evasion....45
1.9.3 Proteins involved in tissue adherence and cellular destruction....46
**1.10 CAPSULAR POLYSACHARIDE (CPS)....49**
**1.11 Global epidemiology of GBS strains....51**
1.11.1 Maternal colonization rate of GBS....51
1.11.2 Infant colonization rate of GBS....54
1.11.3 Serotype distribution of GBS in different populations....55
1.11.4 CC and ST prevalence of GBS....56
**1.12 Trends of GBS disease in neonates and adults in UK; 1999-2019....58**
**1.13 Prevalence of GBS in Brazil....63**
**1.14 Prevalence of GBS in Australia....64**
**1.15 Trends of GBS disease in neonates and adults in US; 1997-2019....65**
**1.16 Bacterial evolution....69**
1.16.1 Point mutations....70
1.16.2 Large scale insertion events....70
1.16.3 Gene Loss....70
1.16.3.1 Role of gene loss in GBS serotype diversity....71
1.16.3.2 Role of pilus genes in GBS diversity....72
1.16.4 Short repetitive sequences....72
1.16.5 Homologous recombination....73
**1.17 Mobile genetic elements (MGEs)....73**
1.17.1 Insertion sequence and Composite Transposons....73
1.17.2 Unit transposons....74
1.17.3 Integron....75
1.17.4 Plasmids....75
1.17.5 Genomic Islands....76
1.17.5.1 Integrative Conjugative Elements....77
1.17.5.2 Integrative and mobilizable elements....77
**1.18 Tn*916*/Tn*916* – like element....78**
**Chapter 2....82**
**Materials and Methods....82**
**2.1 Collection of clinical specimens....83**
2.1.1 Ethics Considerations....83
**2.2 UK isolates....83**
2.2.1 Sample collection, DNA extraction and WGS in PHE laboratories....83
2.2.2 Sample processing in Cardiff University....84
**2.3 Brazil isolates....85**
2.3.1 Sample collection at RJ Hospital, Brazil....85
2.3.2 Processing of Brazilian GBS isolates in Cardiff University....85

**2.4 Australian Isolates** ..... **86**
2.4.1 Sample collection at King Edward Memorial Hospital, Western Australia ..... 86
2.4.2 Sample processing in Cardiff University ..... 87
**2.5 GBS isolates processed at Cardiff University** ..... **87**
2.5.1 Bacterial culturing ..... 87
2.5.2 Antimicrobial screening ..... 88
2.5.2.1 Kirby-Bauer disk diffusion ..... 88
2.5.2.1.1 Australia, Brazil, and UK population isolated GBS ..... 88
2.5.2.2 D-factor test ..... 90
2.5.2.3 Use of antimicrobial gradient strips to define accurate MICs ..... 91
2.5.3 DNA extraction and genomic sequencing ..... 92
2.5.3.1 Brazil GBS isolates ..... 92
2.5.3.2 UK GBS Isolates ..... 93
2.5.4 Genomic Assembly and Annotation ..... 94
2.5.4.1 UK and Brazilian GBS isolates ..... 94
2.5.5 Capsular serotyping and Multi-locus Sequence Typing ..... 96
2.5.6 Resistome ..... 97
2.5.7 Pilus Island, *hvg*A, and Surface protein gene profiling ..... 97
2.5.8 Mobile genetic element (MGEs) analysis ..... 98
2.5.9 Phylogenetic Analysis ..... 99
2.5.10 Recombination analysis to detect potential capsular switching events ..... 99
2.5.11 Pan-genome construction and genome wide association analysis ..... 100
**Chapter 3:** ..... **101**
**Identifying large scale recombination and capsular shifting events in *Streptococcus agalactiae* causing disease in adults in the United Kingdom from January 2014 - May 2015** ..... **101**
**3.1 Summary** ..... **102**
**3.2 Introduction** ..... **103**
**3.3 Results** ..... **105**
3.3.1 Population characteristics and serotype distribution ..... 105
3.3.2 Genotypes and virulence genes prevalent in disease-causing GBS in the UK ..... 107
3.3.3 Phylogenetic relationship between GBS strains causing disease in adults ..... 114
3.3.3.1 Clonal complex 23 (CC23) ..... 114
3.3.3.2 Clonal complex 1 (CC1) ..... 115
3.3.3.3 Clonal complex 8/10 (CC8/CC10) ..... 115
3.3.3.4 Clonal complex 19 (CC19) ..... 115
3.3.3.5 Clonal complex 17 (CC17) ..... 116
3.3.4 Recombination events detected uncommon serotype within specific genotype ..... 122
**3.4 Discussion** ..... **129**
**Chapter 4:** ..... **133**
**Description of antimicrobial resistance (AMR) mechanisms and identification of mobile genetic elements carrying AMR genes in *Streptococcus agalactiae* causing bloodstream infections in adults in UK between January 2014 and May 2015** ..... **133**
**4.1 Summary** ..... **134**
**4.2 Introduction** ..... **136**
**4.3 Results:** ..... **141**
4.3.1 Phenotypic and genotypic AMR profiling: ..... 141
4.3.2 SNP investigation for mechanism of fluoroquinolone resistance using ClustalW: ..... 143

4.3.3 Transpeptidase amino acid substitutions in PBPs ..... 151
4.3.4 Mobile genetic elements (MGEs) carrying MDR genes in GBS strains ..... 153
4.3.4.1 ICE carrying *tet*M gene ..... 162
4.3.4.2 MGEs carrying ARGs in GBS belonging to different CC ..... 166
4.3.4.2.1 CC1 ..... 166
4.3.4.2.2 CC8/CC10 ..... 170
4.3.4.2.3 CC17 ..... 171
4.3.4.2.4 CC19 ..... 173
4.3.4.2.5 CC23 ..... 180

**4.4 Discussion ..... 181**

**Chapter 5: ..... 186**

***Streptococcus agalactiae* in Brazil and Australia: serotype distribution and antimicrobial susceptibility patterns ..... 186**

**5.1. Summary ..... 187**

**5.2. Introduction: ..... 188**

**5.3. Results ..... 191**
5.3.1. Brazilian GBS isolates: ..... 191
5.3.1.1 Serotype and ST distribution in Brazilian GBS isolates ..... 196
5.3.1.2 Genotypic Antimicrobial resistance (AMR) profiling ..... 196
5.3.1.3 Phenotypic AMR profiling ..... 196
5.3.2 Australian GBS isolates: ..... 197
5.3.2.1 Serotype distribution: ..... 197
5.3.2.2 Phenotypic AMR profiling ..... 200

**5.4 Discussion ..... 200**

**Chapter 6: ..... 205**

**Population structure, virulence factors, resistance determinants and pan-genome wide association study (pan-GWAS) of adult isolated *Streptococcus agalactiae* from Brazil, Canada, UK, and United States ..... 205**

**6.1. Summary ..... 206**

**6.2. Introduction: ..... 208**

**6.3. Results: ..... 211**
6.3.1. Geographic distribution of GBS genomes used in this study: ..... 211
6.3.2 Clonal complex assignment and core genome phylogeny ..... 220
6.3.2.1 CC1: ..... 223
6.3.2.1.1 Surface protein, Pilus Island genes and ARGs: ..... 223
6.3.2.1.2 Phylogenetic analysis: ..... 224
6.3.2.2 CC17: ..... 226
6.3.2.2.1 Surface protein, Pilus Island genes and ARGs: ..... 226
6.3.2.2.2 Phylogenetic analysis: ..... 226
6.3.2.3 CC8/CC10: ..... 229
6.3.2.3.1 Surface protein, Pilus Island genes and ARGs: ..... 229
6.3.2.3.2 Phylogenetic analysis: ..... 229
6.3.2.4 CC19: ..... 231
6.3.2.4.1 Surface protein, Pilus Island genes and ARGs: ..... 231
6.3.2.4.2 Phylogenetic analysis: ..... 231
6.3.2.5 CC23: ..... 234
6.3.2.5.1 Surface protein, Pilus Island genes and ARGs: ..... 234
6.3.2.5.2 Phylogenetic analysis: ..... 234
6.3.3 Pan-genome and pan-GWAS ..... 239

**6.4 Discussion ..... 254**

**Chapter 7:** ..... **261**

**General Discussion** ..... **261**

**7.1. General Discussion** ..... **262**

7.1.1. GBS Epidemiology: ..... 265

7.1.2. Serotype distribution: ..... 266

7.1.3. GBS virulence and genetic lineages. ..... 268

7.1.4. Antibiotic resistance and mobile genetic elements ..... 270

7.1.5. Capsular switching and pan-genome analysis: ..... 273

**7.2. Limitations** ..... **276**

**7.3. Future studies** ..... **276**

**Chapter 8:** ..... **279**

**References** ..... **279**

# List of Figures:

Figure 1. 1. Overview of streptococci.. 2
Figure 1. 2. PFGE profiles generated following SmaI digestion of invasive serotype IV GBS DNA isolated from infants, Minnesota. 12
Figure 1. 3. Circular representation of comparative analysis between 19 *S. agalactiae* genomes using microarray hybridization technique.. 18
Figure 1. 4. Stages of neonatal GBS infection.. 19
Figure 1. 5. Different mechanisms used by GBS for the adherence and invasion into host cell. 44
Figure 1. 6. Schematic representation of GBS three pilus islands (PI-1, PI-2a and PI-2b). 45
Figure 1. 7. Schematic representing surface proteins expressed by GBS.. 48
Figure 1. 8. Comparison of capsule genes observed in ten GBS serotypes Ia, Ib, II-IX.. 50
Figure 1. 9. Maternal GBS colonizing serotype distribution by United Nations subregion. **Error! Bookmark not defined.**
Figure 1. 10. MLST defined GBS clonal complexes and their host origin. 57
Figure 1. 11. Age-specific rates of iGBS infection, England, Wales and Northern Ireland: 1999–2019.. 59
Figure 1. 12. GBS bacteraemia age and sex rates per 100,000 population England, Wales and Northern Ireland; 1999-2019.. 60
Figure 1. 13. Rates of GBS-EOD and GBS-LOD infection in England, Wales and Northern Ireland; 1999–2019. 61
Figure 1. 14. Distribution of GBS serotypes from sterile site isolates according to patient age, England, and Wales; 1995–2010.. 62
Figure 1. 15. Resistance rate of clindamycin, erythromycin, and tetracycline against GBS causing bacteraemia, England, Wales and Northern Ireland; 2011 to 2018.. 63
Figure 1. 16. Age-specific rates of iGBS infection, in United States; 1997–2019.. 67
Figure 1. 17. Comparison of rates of iGBS disease in infants (<1 yr) and elderly (>=65 yrs), in the United States and United Kingdom in late 1990s-2019. 67
Figure 1. 18. Rates of GBS-EOD and GBS-LOD infection per 1000 live births in United States 1997–2019. 68
Figure 1. 19. Distribution of iGBS serotypes among isolates recovered from neonates < 3 months old and from all other ages.. 69
Figure 1. 20. Schematic representation of Tn*916*.. 79
Figure 1. 21. The structure of various members of the Tn*916*/Tn*1545* family.. 79
Figure 1. 22. Schematic of ICE*Sa*2603.. 81
Figure 2. 1. Kirby-Bauer test to measure antibiotic sensitivity. 88
Figure 2. 2. Disc diffusion test for inducible clindamycin resistance.. 91

Figure 2. 3. MIC test using strips.. 92
Figure 3. 1. Summary of demographic information available for 179 iGBS isolates submitted to Public Health England between January 2014 and May 2015.. 106
Figure 3. 2. Distribution of surface proteins genes (SPGs) in iGBS isolates belong to five main CCs. 113
Figure 3. 3. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 23 isolated from January 2014 and May 2015.. 117
Figure 3. 4. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 1 isolated from January 2014 and May 2015. 118
Figure 3. 5. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 8 isolated from January 2014 and May 2015. 119
Figure 3. 6. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 19 isolated from January 2014 and May 2015.. 120
Figure 3. 7. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 17 isolated from January 2014 and May 2015.. 121
Figure 3. 8. Phylogenetic relationship between 179 iGBS isolates causing disease in adult population in the UK. 122
Figure 3. 9. Recombination leading to serotype switching for ST1 and CC17 GBS strains from adult patients isolated from January 2014 and May 2015. 124
Figure 3. 10. Phylogenetic relationship between UK ST1 isolates and previously reported recombined Portugal and Canadian ST1 non-serotype V isolates. 125
Figure 3. 11. Comparison of recombination events among selected GBS groups. 126
Figure 3. 12. SNPs distribution in conserved *cps*E-*cps*G and *neu*B-*neu*A regions of *cps*IV locus of CC1 and CC17 GBS isolates from Ireland, France and England and Wales.. 127
Figure 3. 13. Comparison of recombination among serotype V ST498 (CC23) isolates.. 128
Figure 4. 1. Modular arrangement of ICE belonging to Tn*916*/Tn*1545* family. a) Genetic maps and alignment of Tn*916* and Tn*3872*. 140
Figure 4. 2. New Tn*916*-like element identified in two CC1 isolates.. 154
Figure 4. 3. Tn*916* variant observed in CC8/10 isolate - PHEGBS0586 (II/ST652).. 154
Figure 4. 4. Phylogenetic relationship between invasive and non-iGBS (n=193) within five major clonal complexes causing disease in adult population in the UK: 161
Figure 4. 5. Comparison of ICE*Sp*1108-like element found in two CC1 isolates 168
Figure 4. 6. Novel ICE*Sag*100414 identified in CC1 isolate. 168
Figure 4. 7. Newly identified ICE*Sag*87 in CC1 isolate.. 169
Figure 4. 8. Schematic representation of the newly identified integrative and conjugative element (ICE*Sag*662) in CC1 isolate.. 169

Figure 4. 9. ICE*Sp*2905-like element (NZ_FR691055.1) identified in four CC8/10 serotype II/ST12 isolates.. 171
Figure 4. 10. Comparison of ICE*Sag*37 identified in four CC17 isolates. 172
Figure 4. 11. Comparing *S. pyogenes* ICE*Spy*009 identified in three UK GBS isolates.. 173
Figure 4. 12. Comparison of *Streptococcus suis* ICE*Ssu*YS64 with two CC19 isolates.. 175
Figure 4. 13. Novel ICE*Sag*71 identified in CC19 isolate.. 176
Figure 4. 14. Comparison of *Staphylococcus aureus* composite mobile element structure $MES_{PM1}$ to CC19 GBS isolates. 176
Figure 4. 15. Schematic, but to scale, representation of the ICE*Sag*236-like element in two CC19 UK GBS isolates. . 177
Figure 4. 16. Comparison of *Streptococcus* phage identified in CC19 isolate.. 177
Figure 4. 17. Representation of newly identified ICE*Sag*139 detected in UK GBS serotype V/ST19 PHEGBS0139. 179
Figure 4. 18. Identification of *S. agalactiae* strain GB2001 plasmid (pGB2001) carrying *erm*T in CC23 isolate.. 181
Figure 5. 1. Summary of demographic information available for 26 iGBS isolates from Southeast region of Brazil. 192
Figure 5. 2. GBS antenatal serotype distribution at vaginal and rectal sites vs overall prevalence identified. 199
Figure 6. 1. Workflow diagram of Pan–GWAS analysis performed in this study. 212
Figure 6. 2. Serotype distribution from Brazilian, Canadian, the United States and UK datasets and its prevalence among different CCs.. 222
Figure 6. 3. Phylogenetic relationship of GBS strain population from Brazil, Canada, United States and UK clustered in CC1.. 225
Figure 6. 4. Phylogenetic relationship of GBS strain population from Brazil, Canada, United States and UK clustered in CC17.. 228
Figure 6.5. Phylogenetic relationship of GBS strain population from Brazil and UK clustered in CC8/10.. 230
Figure 6. 6. Phylogenetic relationship of GBS strain population from UK clustered in CC19.. 233
Figure 6. 7. Phylogenetic relationship of GBS strain population from Brazil and UK clustered in CC23.. 236

# List of Tables:

Table 1. 1. List of medically important streptococci. ..... 3
Table 1. 2. GBS management and prevention guidelines in different countries. ..... 28
Table 1. 3. Most frequently observed antibiotic/bacteriocin resistance genes in GBS. ..... 34
Table 1. 4. Summarizing candidates for GBS vaccine. ..... 37
Table 1. 5. Key virulence factors of GBS.. ..... 41
Table 1. 6. Maternal GBS colonization prevalence results from meta-analyses with reported data and meta-analyses with adjusted data.. ..... 52
Table 1. 7. Incidence rate of iGBS disease in neonates per 1000 live births worldwide.. ..... 55
Table 2. 1. Zone Diameter and Minimal Inhibitory Concentration Interpretive Standards for *S. agalactiae* asper CLSI guidelines M100S. ..... 90
Table 2. 2. Sequence statistics for 14 UK GBS isolates sequenced by both Miseq and Nanopore technology. ..... 94
Table 2. 3. Quast v2.1 essential assembly metrics observed for 193 UK and 27 Brazilian GBS genomes that passed quality control test. ..... 95
Table 2. 4. Poor assembled genomes of UK and Brazil GBS isolates that were excluded from the analysis. ..... 95
Table 2. 5. Oligonucleotide primers used for GBS serotyping using Geneious R11 (Biomatters ltd. New Zealand).. ..... 96
Table 2. 6. Oligonucleotide primers used for the detection of virulence genes in UK and Brazil GBS genomes. ..... 98
Table 3. 1. Summary of characteristics of 179 iGBS whole genome sequences used in this study.. 1088
Table 3. 2. Sequence types and serotypes of UK adult iGBS isolates belonging to five main CC. . 1133
Table 4. 1. Summary of 193 GBS isolates used in this study… ..... 15144
Table 4. 2.a) Summarizing ResFinder v2.1 predicted antibiotic genetic variants in 193 GBS isolates and b) presenting GBS isolates harbouring multiple antibiotic resistance gene determinants. ..... 15151
Table 4. 3. Summary of Penicllin binding proteins (PBPs) found in 193 GBS UK isolates … ..... 15152
Table 4. 4. Sumamry of common target genes identified in ARGs associated MGEs. ..... 1622
Table 4. 5. Characteristics of clones deriving from Tn*916*/ Tn*916*-like element and Tn*5801* insertions.. ..... 1644
Table 5. 1. Description, genotypic and phenotypic antimicrobial resistance profiles of iGBS isolates (n=26) recovered from clinical blood specimens of Brazillian aduts between July 2010 to November 2017. ..... 1943

Table 5. 2. Summary of phenotypic antibiotic resistance patterns of GBS of three geographic locations. ..... 1944
Table 6. 1. Description of 228 publicly available GBS genomes used in this study.. ..... 2133
Table 6. 2. Clonal complexes identified in 417/447 invasive GBS dataset from different countries. ..... 22020
Table 6. 3. Prevalence of antibiotic resistance genes, virulence genes, and pilus island genes in five major CCs. . ..... 22037
Table 6. 4. Functional class and pathways associated with each CC as per KEGG database results. ..... 2399
Table 6. 5. Genes specific to CC1 as identified by Scoary v1.6.16. ..... 2443
Table 6. 6. Genes specific to CC19 as identified by Scoary v1.6.16. ..... 2444
Table 6. 7. Genes specific to CC8/CC10 as identified by Scoary v1.6.16 ..... 2455
Table 6. 8. Genes specific to CC17 as identified by Scoary v1.6.16 ..... 2499
Table 6. 9. Genes specific to CC23 as identified by Scoary v1.6.16. ..... 2522

# Chapter 1:

# General Introduction

## 1.1. Streptococci

The genus Streptococcus comprised of Gram-positive bacteria - *Streptococci* that occur in pairs or chains and cause diverse human diseases [4]. Streptococci are catalase negative and can be classified based on their haemolytic effect on blood agar, antigenic composition, growth characteristics, genetic characteristics, and biochemical reactions [1]. By examining their appearance on blood agar, three distinctive patterns emerge, which are termed as alpha (α, partial); beta (β, complete), and gamma (γ, none) haemolysis [5, 6]. Using a serological system developed by Rebecca Lancefield in 1933, the beta-haemolytic streptococci species can be further classified by the cell wall carbohydrate [5]. Beta-haemolytic streptococci also termed as 'pyogenic'(relating to the production of pus) [7, 8] and are clinically important bacteria associated with human diseases (**Figure 1.1**). Characterised by the major surface polysaccharide antigen type (or Lancefield group) expressed, beta-haemolytic streptococci are classified into group A (*Streptococcus pyogenes,* GAS), group B (*Streptococcus agalactiae,* GBS), group C (multiple zoonotic species plus the human species *Streptococcus dysgalactiae subsp. equisimilis,* GCS); group G (human and animal species *Streptococcus dysgalactiae subsp. equislimilis* and *Streptococcus canis,* GGS) and other (*Streptococcus pneumoniae*) [7, 8]. The non-pyogenic streptococci are clinically less important and subdivided into species including *S. mitis*, *S. sanguinis, S. anginosus*, *S. salivarius, S. mutans,* and *S. bovis* (**Figure 1.1**) [7, 8]. There are many medically important streptococci that cause blood and invasive infection (**Table 1.1**) [9]. In particular GAS causing invasive infection, endocarditis, scarlet fever and GBS as the biggest cause of death in babies less than 4 weeks and *Streptococcus pneumoniae*, a major cause of mortality in the elderly [9].

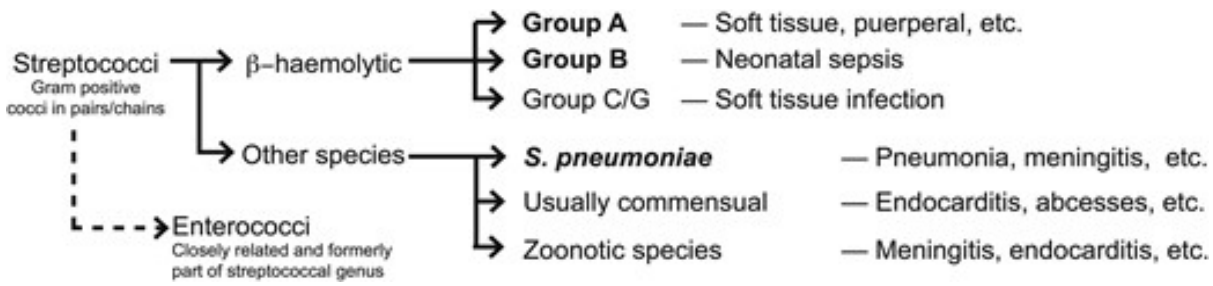



**Figure 1. 1. Overview of streptococci.** Species of greatest public health importance are highlighted in bold font. Adapted with permission from [4].

**Table 1. 1. List of medically important streptococci.**

| Type species | Lancefield group | Normal host | Significant human disease |
|---|---|---|---|
| *S. pyogenes* | A | Humans | Scarlet fever, bacteraemia, pneumonia, necrotizing fasciitis, myonecrosis and Streptococcal Toxic Shock Syndrome (STSS). |
| *S. agalactiae* | B | Cattle, Humans | Neonatal meningitis and sepsis and infections in adults |
| *S. dysglactiae subsp. equisimilis* | C | Wide human and animal distribution | Endocarditis, bacteraemia, pneumonia, meningitis mild upper respiratory infection |
| *S. bovis* | D | Human and animal intestinal tracts, dairy products | Bacteraemia, endocarditis and bone and joint infections |
| *S. anginosus* | F, G | Humans, animals | Subcutaneous or organ abscesses, endocarditis, and upper respiratory infection |
| *S. anguis* | H | Humans | Endocarditis |
| *S. salivarius* | K | Humans | Endocarditis |
| None | O | Humans | Endocarditis |
| *S. suis* | R | Swine | Meningitis |
| *S. mitis, S. mutans* | None identified | Humans | Endocarditis |
| Anaerobic or microaerophilic | None identified | Wide human and animal distribution | Brain and pulmonary abscesses, gynaecologic infections |
| *S. pneumoniae* | None identified | Humans | Lobar pneumonia and others |

## 1.2. Beta-haemolytic streptococci

### 1.2.1. Group A streptococcus

*Streptococcus pyogenes* also known as GAS, is a major human-specific bacterial pathogen that causes a wide array of clinical manifestations ranging from mild localized infections to life-threatening invasive infections [10]. These include tonsilitis, pharyngitis, scarlet fever, impetigo, erysipelas, cellulitis and pneumonia [11]. Ineffective treatment of GAS infections can result in the postinfectious sequelae acute rheumatic fever and post-streptococcal glomerulonephritis in childhood and adolescence [11]. While all invasive GAS (iGAS) infections are serious and potentially fatal, two manifestations are of particular note: necrotizing fasciitis and peripartum sepsis; both are associated with high morbidity

and mortality [12, 13]. GAS-associated peripartum sepsis was the UK's leading cause of direct maternal death during 2006–2008 [14]. Moreover, all iGAS infections can be complicated by STSS, a condition characterized by hypotension often accompanied by fever or rash with rapid progression to shock and multi-organ failure [15]. Although STSS appears in ~10% iGAS infections, early detection of STSS is extremely important, because >25% patients will die within 24 hrs of presentation [15]. Generally, the occurrence of iGAS infections appear to have been increasing worldwide since the 1980s [16]. iGAS is a notifiable disease in the UK with an estimated average of 3.5 cases recorded per 100,000 population between 2015 and 2019 [8]. Higher rates of iGAS were recorded in males for all ages, except in the 15-44 years age group relative to females [8]. The highest rates of iGAS were in the elderly (75+ years), followed by those under 1 year old [8]. Infections caused by *S. pyogenes* are highly contagious [11]. Transmission can occur through airborne droplets, hand contact with nasal discharge or with objects or surfaces contaminated with bacteria, skin contact with contaminated lesions, or contaminated food sources [11].

### 1.2.2 Group B streptococcus

*Streptococcus agalactiae* is a Gram-positive bacterium, initially identified as a veterinary pathogen and important cause of bovine mastitis in the 1800s under the label of *Streptococcus mastiditis* [5]. In 1933, after Rebecca Lancefield studied serological differentiation of streptococci, it was first recognised as GBS [5]. In 1938, GBS was recognized as an important human pathogen causing pleural infections [17], later in 1960s it was first identified as a main cause of invasive infections (bacteraemia and meningitis) in newborns [18]. In the 1970s, GBS emerged as a leading cause of neonatal morbidity and mortality in the United States [19], and since the 1990s it has been increasingly accepted as a cause of invasive infections (sepsis in particular) in all adults, and not just associated with disease in pregnant women [20]. Notwithstanding its importance as a human pathogen, GBS is an opportunistic pathobiont (this is a term commonplace in GBS research and is defined as a bacteria that can exist as a commensal, but can be elevated to pathogenic potential under certain conditions such as pregnancy or immunosuppression), constituting a part of the intestinal and vaginal microbiota [21]. In the UK,

between 2015-2019, a 31% increase in the rates of invasive GBS (iGBS) was recorded, and overall highest numbers of iGBS cases were reported in those aged less than one year compared to adults aged more than 15 years [8]. On average, rates were higher in adult males than adult females, except for cases in the 15-44 years age group [8].

### 1.2.3 Group C and G streptococcus

GCG and GGG include several species of bacteria, with the main human pathogen being *Streptococcus dysgalactiae ssp. equisimilis* [1]. These bacteria usually cause throat, skin and soft tissue infections but can also cause invasive infections including bacteraemia, endocarditis (inflammation of the lining of the heart) and bone and joint infections [1]. Toxic shock can follow both localised and invasive infections [11]. Group C and G streptococci are least important in terms of serious infections; however, some *Streptococcus dysgalactiae ssp. equisimilis* harbour GAS group A antigen and can cause life-threatening invasive infections such as STSS [22]. Recent epidemiological studies reveal that invasive SDSE infections have been increasing in Asia, Europe, and the United States [22]. An increased number of cases of iGCS infections (83%) were also recorded in the UK in past four years (2015 to 2019). Similar to GBS, iGCS infections were reported more in males than females and particularly in the 75 years and over age group [8]. Group C and G streptococcal infections are spread by person-to-person contact, such as sneezing, coughing, or touching an open wound and may be treated with antibiotics [11].

## 1.3 Microbiological characterisation of GBS

In microbiology laboratories, identification of microorganisms is one of the fundamental tasks [23]. This further allows to understand their clinical evolution, their enteropathogenic implications and applying an effective antimicrobial therapy [24]. In past, diverse phenotypic and genotypic methods were used for the characterization and identification of bacteria [25]. However, in the last decades, genotypic methods proved as a better alternative for bacteria identification and for epidemiological research of infectious diseases [25].

### 1.3.1 Phenotyping methods

For the typing of S*treptococci* strains typing, different phenotypic typing methods are used in laboratories [23, 26]. These methods are based on colony morphotype, enzymograms, resistance to various chemicals, and more improved techniques of typing, for example PCR fingerprinting [27]. Morphotyping is the simplest method to type S*treptococci* strains based on evaluating colonies appearance on blood agar. Factors like colour, surface, edge, gram staining result and haemolysis type are evaluated [28], while serotyping based on agglutination is the one of the earliest methods established for S*treptococci* phenotyping [29]. These conventional typing methods play an important role in identifying outbreak isolates; however, they are insufficient for deriving evolutionary relationships between strains [1].

#### 1.3.1.1 Growth on Agar

GBS belongs is a beta-haemolytic, catalase negative, Gram-positive coccus able to grow in chains [30, 31]. Blood agar, selective blood agar (such as neomycin-nalidixic acid agar [NNA] or colistin-nalidixic acid agar [CNA]), Granada agar, or chromogenic agar can be used to grow GBS [32]. Small white microbial colonies surrounded by a narrow zone of β-haemolysis are observed on blood agar, for chromogenic agar, colored colonies and for Granada agar, the development of brick, orange, or red colonies are indicative of GBS [30]. GBS-like colonies that develop on chromogenic agar and typical GBS colonies that are not pigmented on Granada agar should always be confirmed as GBS using additional tests (e.g., the latex agglutination test or the CAMP test) to avoid false-positive results [32]. Other laboratory methods such as biochemical profiling, matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry [33] and PCRs are also used in laboratories for the identification of GBS.

#### 1.3.1.2 CAMP test

The CAMP test is named as an acronym of the initial authors describing this test (Christie, Atkins and Munch-Petersen) in 1994 [34]. This test is used in some laboratories to confirm bacteria with increased staphylococcus beta hemolysin activity test, which has long been considered as a key presumptive identification of GBS [34, 35]. The organism under test is inoculated in a fine streak on the surface of bovine or sheep-blood agar; a second streak – perpendicular to the first but separated from it by a few millimetres, is made with a culture of *Staphylococcus aureus* strains producing beta-hemolysin [34]. The plate is then incubated aerobically at 37° for about 12 hours [34]. A test is considered positive if a typical arrowhead or flame-shaped area of extensive haemolysis is seen on the plate– due to synergy between a streptococcal CAMP factor and staphylococcal beta- hemolysin at the intersection of bacterial species growth [34, 35].

#### 1.3.1.3 Biochemical Profiling

For several years, conventional phenotypic tests have remained keystones for the identification of bacterial microorganisms in clinical microbiology [32]. Countless test systems with the ability to integrate multiple conventional phenotypic tests into a single-step procedure have been developed [32]. After the interpretation of results, the biochemical reactions included in the kit system generate a biotype number that is matched against the profile indexes from a database to identify the bacterial species [32]. Examples of biochemical profiling kits for the identification of streptococci are the API Rapid Strep identification system (bioMérieux) and the RapID STR system (Remel) [36]. However, the ability of these systems to identify microorganisms depends on the accuracy of their databases of profile indexes and the inclusion of all relevant microbial species [36]. Although these systems can be used to identify beta-hemolytic streptococci, they are not 100% specific. While the accuracy to identify GBS is high, more reliable results may be achieved by using a combination of simpler phenotypic methods, reserving the use of these kits for the identification of other streptococcal species [32].

#### 1.3.1.4 MALDI-TOF

Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS) is a rapid, high-throughput, time and cost-effective method for microorganism identification which provides highly reproducible data based on the analysis of highly abundant proteins (mainly ribosomal proteins) [37]. Spectra of MALDI–TOF MS can be distinctive for separate genus, species, and strain-specific peaks and these peak patterns have been proposed as taxonomic markers to recognize putative new species [38]. GBS genetic diversity was discovered in parallel with variations of the protein patterns evaluated by MALDI-TOF MS [39], which can even provide sub-species level discrimination of some GBS strains [40]. A recent study demonstrated that the highly conserved ribosomal subunit proteins (rsp) could serve as ideal biomarkers for strain level typing of GBS [41]. In particular, mass variations indexing of 28 rsp using MALDI-TOF MS further allowed classification of GBS isolates into the major phylogenetic lineages, identification of obligate animal associated lineages and detection of hypervirulent Clonal Complex (CC17) strains [41].

MALDI-TOF MS has been adopted as a speedy and accurate microbial identification method [42], however, MALDI-TOF MS is restricted today mostly to the identification of pure colonies grown on a solid medium, because generation of a satisfactory mass spectrum requires that an adequate amount of sample be deposited onto a target for analysis [43]. The need to study pure cultures of microorganisms originates from the inability of current software programs to analyse the spectra acquired from mixed cultures [43]. Because of that, MALDI-TOF MS is not yet able to identify microorganisms directly from most clinical specimens, such as swabs and wound specimens [32]. This precludes its use for the direct identification of GBS in vaginal and rectal swabs [32]. Also, MALDI-TOF MS cannot be used for all streptococcus species identification but can be used for some and to group level identification [32].

### 1.3.1.5 Serotyping

The gold standard method for phenotypic assessment of GBS is serotyping, which is based on the expression of unique CPS capsules at the bacterial surface that react with serotype specific antibodies [44]. Ten different serotypes have been identified to date: initially Ia, Ib, II-VIII [45] with the emergence of serotype IX in 2007 [46]. Further, occasionally GBS strains are determined to be non-typeable (NT); meaning they carry undetectable capsular polysaccharide, either non-capsulated or express polysaccharide variants that fail to react with any sera, hence are unable to be characterised by any common methodologies used in laboratories [47]. The latex agglutination method is conventionally performed for GBS serotyping, but an increased number of non-typeable strains or inaccurate classifications are occurring [48]. In such situations molecular approaches based on amplification of the genes encoding GBS capsular components are used to classify strains not detected by conventional methods [49]. Multiplex PCR is the molecular approaches that permits the detection of the ten GBS serotypes in a single reaction. It is mostly amenable to routine laboratory use because of its low cost, ease to implement and increased reliability compared to conventional serotyping [50, 51].

### 1.3.1.6 Antimicrobial susceptibility testing (AST)

GBS strains can be discriminated based on AST, which depends on their stability, diversity and prevalence of detectable antibiotic resistance profiles [23]. Mostly in clinical settings, automated systems are used to determine antibiotic susceptibility patterns, which is very useful in guiding clinically appropriate therapy [52].

The Kirby-Bauer disk diffusion method is the most frequently used method in research studies and diagnostic laboratories [53]. Small antibiotic disks impregnated with known antibiotic concentrations are placed onto agar plates after a set concentration of each strain are thoroughly distributed on the surface. Following an incubation period of 24 to 48 hrs, the diameter of inhibition around the disk is

measured with a ruler or calliper, which is then compared to the standard values containing internationally agreed values. The diameter will determine whether the bacteria is resistant, intermediately resistant, or susceptible to each specific antibiotic [54]. Another method is to use antibiotic gradient strip, which contain an increasing gradient of an antibiotic and marked with defined antibiotic concentrations along its length [55]. This strip is overlaid on an agar plate following inoculation with a standardised, defined bacterial concentration (as above for the Kirby-Bauer method) and after the incubation period, an ellipse shaped inhibition zone is created and where the growth intersects the E-test, the minimum inhibitory concentration of antibiotic is read. These values are compared to internationally agreed tables to determine if the strain is deemed to be susceptible or resistant to that antibiotic [54]. However, some antibiotic interactions are complex, such as combinations of β-lactams with β-lactam-inhibitors (i.e. co-amoxiclav), or sulfamethoxazole/trimethoprim, or when efflux pumps require macrolides to be present to induce resistance to lincosamides (i.e. $iMLS_B$ identified by the D-factor induction test) [56]. International and independent organizations like the Clinical and Laboratory Standards Institute (CLSI) [54] and the European Committee on Antimicrobial Susceptibility Testing (EUCAST) define breakpoints [57] and guidelines for testing and standardise protocols to be used and publish the interpretive standardised tables for all these phenotypic antibiotic resistance tests.

### 1.3.2 Genotyping methods

To assess genomic variation of bacterial isolates in relation to their overall somatic structure, presence or absence of plasmids, mobile genetic elements or precise gene and intergenic nucleotide sequence regions, different genotypic methods are used [58].

#### 1.3.2.1 Restriction Fragment Length Polymorphism (RFLP)

This method relies on digestion of DNA using one or more endonucleases. Separation of resulting fragments of variable length is performed using conventional electrophoresis and the pattern obtained reflects the distribution and frequency of endonuclease recognition sites [59]. These fragments can be

further hybridized with probes after transfer to a solid membrane support, for further subtyping of isolates, a method referred to as Southern Blot [60]. Recognition of DNA fragment that contain the sequence complementary to the probe is used to generate "fingerprints", and probe selection is critically important in this method for isolate typing [59].

RFLP was used in a GBS population study to investigate serotype III lineages [61]; results discriminated three separate phylogenetic lineages, with one strongly associated with neonatal isolates, suggesting an increased virulence of this lineage relative to the other lineages [61]. Southern blot and RFLP are also useful for subtyping and classifying non-typeable GBS strains using insertion sequences [62].

#### 1.3.2.2 Pulsed-field gel electrophoresis (PFGE)

In PFGE, when chromosomal DNA is exposed to endonucleases, macro-restriction fragments are generated (**Figure 1.2**) [59]. Greater discrimination in separating these bands requires a modified electrophoresis method and different equipment. Agarose gels are run once agarose plugs containing the digested chromosomal DNA are loaded into the wells of the gel, causing the restriction fragments to resolve into discrete bands. The electrophoresis apparatus switches orientation of electric field systematically across the gel. Patterns of DNA fragments are then matched to standardised markers to study their similarity or differences [63].

In GBS molecular epidemiologic studies, PFGE potential was first conveyed when matched to conventional electrophoresis [64]. Reproducible and simplified profiles showed greater sensitivity and discriminatory power to evaluate genetic similarity and diversity amongst strains relative to conserved or different serotypes by PFGE [65]. However interpreting PFGE results manually or to deal with larger PFGE profile datasets is challenging [66]. Suggested interpretative criteria, to ease translation of PFGE results into useful information for epidemiological study, were proposed as follows: Isolates are classified as closely related strains if they carried one point mutation or an insertion/deletion (one genetic event) and produce one or two bands in PFGE profile; strains with two genetic events

(independent) normally generate 4 to 6 band differences by PFGE and are considered as possibly related strains; however PFGE profiles are considered unrelated if strains differ by 3 or more independent genetic events and produce 7 or more differences in banding patterns [67].

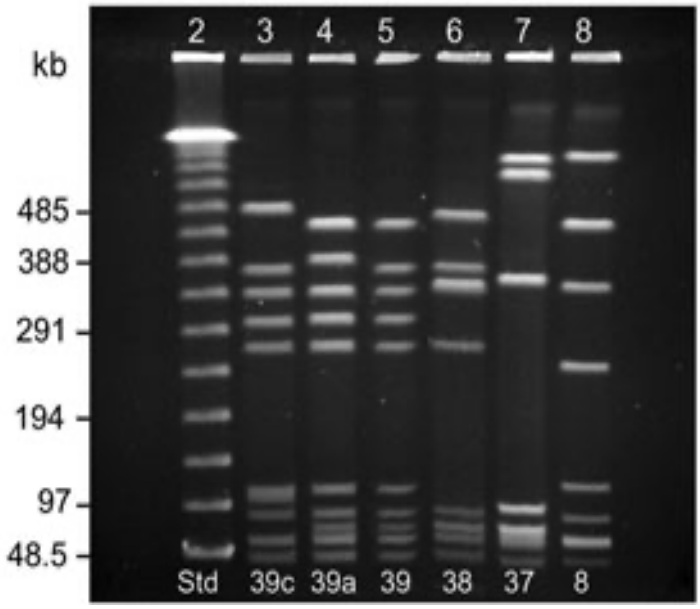



**Figure 1. 2. PFGE profiles generated following SmaI digestion of invasive serotype IV GBS DNA isolated from infants, Minnesota.** Lane 2, λ molecular size standard; lanes 3 and 4, serotype IV/C-α GBS isolates from early-onset disease; lanes 5–7, prototypes of PFGE profile groups 39 (IV/C-α), 38 (IV/C-α), and 37 (IV/C-α and BPS); lane 8, internal standard 89-022 (Ib/C-α and C-β). Adapted with permission from [68].

Many sophisticated software programs have simplified comparison of larger numbers of isolates by normalizing banding patterns over multiple gels and then construct genograms, which can be further utilized to assess phylogenetic relationships between strains [69]. PFGE is used extensively for studying large GBS populations and evaluating GBS dynamics within restricted cohorts [70]. It has been revealed to have great potential to address questions about GBS transmission and acquisition, after identifying that neonates usually carry the same GBS isolate as their mothers and that the sexual partners usually carry the same GBS clones, likely transmitted via sexual contact [70].

#### 1.3.2.3 Multilocus sequence typing (MLST)

In 1998, Multilocus sequencing typing (MLST) was first described by Maiden and colleagues which later became the most widely used and effective molecular method for the classification of bacterial isolates at the strain-level [71]. MLST uses numerical classification system to index the sequence variation, found in a set of housekeeping genes [71]. A distinct number is then assigned to the unique allele and the profile of a given isolate is specified by the numbers representing the allelic composition. Each combination of alleles is assigned a unique sequence type (ST), and STs, in turn, can be grouped into higher levels of classification [71, 72].

The MLST approach has limited ability to resolve closely related strains due to the absolute sequence diversity present in the loci and distributes across the population to be classified [72]. Further, for MLST method to work robustly, the targets and the PCR binding sites must be sufficiently conserved which also limits the resolution of classical MLST [72]. MLST schemes have now been developed for >100 genera and species, and continuously updated databases are maintained by the University of Oxford, United Kingdom (pubmlst.org), Imperial College London, London, United Kingdom (http://www.mlst.net/), and others [72]. In many cases, classical MLST method works by segmenting 400 to 500 nucleotide of usually (seven) housekeeping genes, comprised of <0.2% of bacterial genomes [72]. Availability of numerous assembled bacterial genomes and recent advances in WGS led to the emergence of new approaches to track isolates epidemiologically at strain and clone level, including the ability to use MLST schemes on a genome-wide scale [73, 74].

Interpretation of clonal relationships between strains using WGS provide more details for epidemiological investigations, than classical MLST analysis [75]. Perhaps, the most definitive approach among many available methods is single nucleotide polymorphism (SNP) based phylogenetic analysis also known as Direct alignment based phylogenetic analysis approach, which frequently works on underlying statistical assumptions of independent individual substitution events [71, 72]. This

analysis has been used to construct transmission maps, to study the population structure and recombination leading to capsular switch in a variety of epidemiologic investigations [73]. An earlier study [76] assessed 200 serotype V GBS ST1 strains isolated from the bloodstream of non-pregnant adults in the United States and Canada between 1992 and 2013 to detect phenotypic diversity and phylogenetic relationships among ST1 GBS strains using SNP based approach. The result showed that the emergence of serotype V GBS causing invasive disease in non-pregnant adults is primarily driven by ST1 strains that are highly similar at the whole-genome level but possess significant phenotypic diversity because of small genetic changes rather than extensive recombination [76]. Another study from Canada used SNP based alignment approach and Bratnextgen algorithm to investigate 111 ST1 GBS isolates collected by the Toronto Invasive Bacterial Diseases Network during 2009–2015 and reported various cases of recombination events leading to capsular switching among ST1 GBS strains [77]. More recently [78], 555 GBS strains causing invasive disease among non-pregnant adults in Portugal between 2009 and 2015 were assessed and a core SNP alignment and recombination analysis were performed with the Gubbins software. The result indicated that the Ib/CC1 lineage probably resulted from acquisition of the type Ib cps locus in a single recombination event by a representative of the V/CC1 macrolide-resistant lineage further, the small number of SNPs between the genomes of the Ib/ST1 isolates from Portugal and Canada and the larger number of SNPs relative to the genomes of serotype V isolates from the same locations, suggests a common origin of the Ib/ST1 isolates [78].

However, a more complex rearrangements over a contiguously exchanged or interrupted region or multiple SNPs could occur because of mobile genetic elements (MGEs) insertions, single recombination events, or other horizontal genetic transfer events [75]. As a result, genome-wide direct-alignment methods may result in incorrect tree topologies and increased computational expenses due to such rearrangements [75]. Instead, MLST-derived approaches applied to WGS provides computationally simplified methods for handling these problems, as the unit of analysis is the allele as opposed to the individual SNP [71, 72]. These include i) ribosomal MLST (rMLST), ii) whole-genome MLST (wgMLST), and iii) core genome MLST (cgMLST) methods. The rMLST approach analyses 52 to 53 genes encoding ribosomal protein subunits (*rps*) which almost all bacteria harbours, conferring

the advantage of universality to this method [71]. The wgMLST strategies, in contrast, analyse all loci in a given isolate, including intergenic regions if they are defined [71]. Statistical difficulties arise in wgMLST approaches when comparing loci shared only by subsets, and not by all the members of the group to be examined [71]. The cgMLST methods address this problem by comparing only genes common to a set of bacterial isolates [71]. These methods are now mature and have been applied successfully to the analysis of several bacterial outbreaks [79-81].

Previously, MLST has been regarded as the gold standard typing tool for the characterization of bacterial isolates using the sequences of internal fragments of (usually) seven housekeeping genes [71]. Nevertheless, recombination in the housekeeping genes has been reported in several bacterial pathogens, including *E. coli*, *S. enterica*, *S. aureus* and GBS [82-84]. The MLST-based phylogeny is not entirely free from the noise created by recombination, and scientists have turned to WGS to minimize this interference [85]. In a recent study [85], phylogenetic inconsistency in CPS genes and pilus islands in 1016 GBS genomes (deposited in NCBI database) of isolates collected from 28 countries in Asia, Europe, South America, Africa, Australia, and North America were assessed to explore their relationship with invasiveness using the combination of cgMLST and the Bayesian clustering method (BAPS). The comparison of the results of the genome-wide BAPS and the seven gene MLST found that the two-typing systems were compatible to each other, suggesting that the seven genes used for MLST in GBS had undergone little recombination and were good candidates for within-species typing [85].

However, compared to conventional PFGE technique, MLST plays an important role in bacterial strain characterization and is more precise, unified, and unambiguous [71]. In recent years the use of MLST over PFGE has drastically increased for characterizing bacterial populations due to its efficacy to deduce degree of similarity and to develop evolutionary relationships between strains [86]. eBURST is an algorithm that implements models of clonal diversification and relationship by dividing MLST datasets into clusters of related isolates [87].

This algorithm further predicts the emerging clonal complexes (CCs), which centers around a founding genotype and makes closely related MLSTs easier to visualise random genetic drift or a fitness advantage. The resulting genotype is greater in number, including strains arising by point mutation or genetic recombination (diversification), and gives rise to the final clonal complex. In the sense of MLST, no change is observed in the descendants of founder allelic profile, but with time variants among any one of seven alleles may arise [87]. If the genotype allelic profiles differ at only one of the seven MLST loci from the founder, they are called single locus variants (SLVs). If difference at two MLST loci occur, they are called double locus variants (DLVs) and so on [87].

Since eBURST is locally optimized and can violate the proposed rules for linking STS within the clonal group, a further modified algorithm was proposed named goeBURST, which applies tiebreak rules before linking STs and provide greater reliability by presenting a graphical visualization of hypothetical pattern of descendants [88]. This software is available online at (http://goeBURST.phyloviz.net).

#### 1.3.2.4 PCR-based gene profiling

This is the universal typing method that utilizes polymerase chain reaction (PCR) technique and has numerous applications in typing different bacterial populations [23]. This gene profiling unveils major advantages over other techniques in term of simplicity, high reproducibility, wide equipment and reagent availability and quick output time [23]. Several GBS strains have been genotyped using PCR based gene profiling methods and include, among others, molecular serotyping [89, 90]; sub-typing within distinct serotypes [91]; detection of mobile genetic elements [92]; surface protein gene profiling; and determining antibiotic resistance genes [93].

#### 1.3.2.5 Whole genome sequencing (WGS)

In recent years, significant numbers of complete genomic sequences of pathogenic organisms have been produced following the development of less expensive and more efficient methods called whole genome

sequencing [94]. Currently 17,331 GBS WGS are available at European Nucleotide Archive representing 54 different countries (https://www.ebi.ac.uk/ena, last accessed February 2022). WGS provides the possibility to determine the genetic differences within bacterial species and understand how genetic variation leads to evolution of virulence mechanisms [95].

In late 2002, the first full genome sequences of two human GBS strains (NEM316 and 2603V/R) were deposited in a public database [96]. The genomic size of GBS was about 2.2 Mbp on average and carried over 2100 proteins [96]. The study revealed a set of core genes, possessed by GBS, that remain conserved across the species and match roughly 80% of each genome individually [97]. Genomic analysis showed about 55% genes in the NEM316 strain had an orthologous variant in *S. pyogenes* [96], while among the remaining NEM316 genomic regions, half were clustered within 14 genomic islands defined by known and putative virulence genes [96]. Similarly, another comparative genomic analysis of *S. agalactiae*, *S. pneumoniae* and *S. pyogenes* strains revealed genes specific to the streptococci and to *S. agalactiae* **(Figure 1.3**) [98]. On the contrary, various known and unknown putative virulent genomic regions were identified in GBS strains that differed from other streptococcal species [96]. These regions encode virulence genes and were associated with mobile genetic elements, representing these regions as pathogenic islands. The virulence genes were probably associated with GBS during adaptation to specific niches in different hosts (human and animal) and were assumed to play a critical role in GBS colonization and/or disease [98]. Similarly, gaining multiple virulence traits from different species is another factor that is linked with MGE including transposons, bacteriophages and insertion sequences that carry a high number of genes [98].

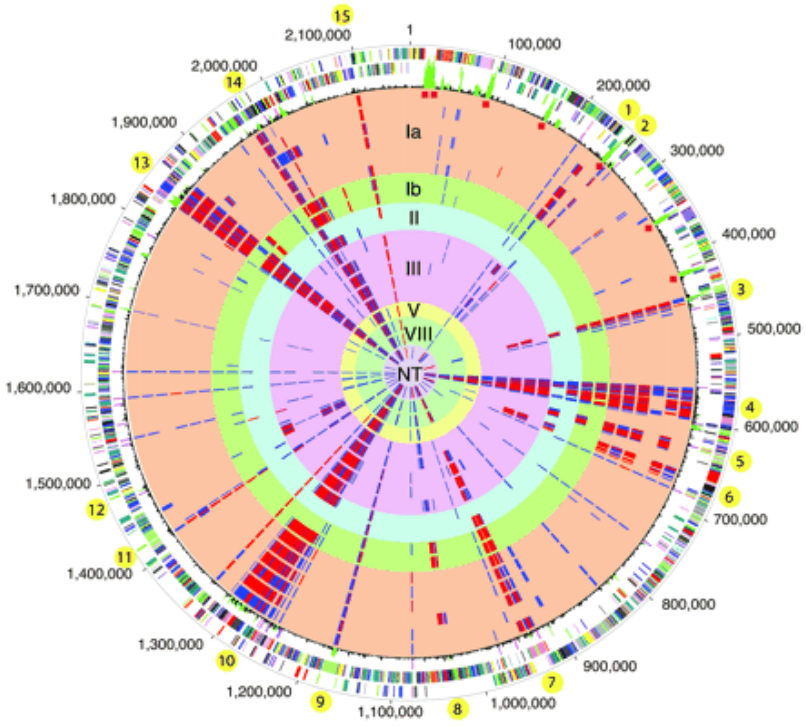



**Figure 1. 3. Circular representation of comparative analysis between 19 *S. agalactiae* genomes using microarray hybridization technique.** Circular representation of comparative analysis between 19 *S. agalactiae* genomes using microarray hybridization technique**.** Outer circle is showing coding regions on the plus strand and different colour coded as per their role categories; The second circle shows the coding regions on the minus strand. The third circle is black and shows atypical nucleotide composition curve; green, most atypical regions; magenta, insertion elements; red diamonds indicate rRNAs while circles 4–22 are showing comparative genome hybridizations of strain 2603 V/R with 19 *S. agalactiae* strains. Adapted with permission from [98].

Another noteworthy application of whole genome sequencing is the comparative analysis of multiple genomes that determines the pan-genome of each organisms [97, 99]. The pan-genome represents the entire set of genes within a species, consisting of a core genome that contains sequences shared between all individuals of the species and the 'dispensable' genome [97]. The idea of a pan-genome was first conceived for bacterial species in 2005, when the genomes of six strains of *Streptococcus agalactiae* were sequenced, revealing a core genome containing 80% of *S. agalactiae* genes [97]. The greater the number of genes found in variable regions of MGE or extrachromosomal elements, the greater the hypothesis of lateral gene transfer between species and is supported by resultant acquisition of strain specific traits [99]. Whereas the presence of genes in some strains/species and not in others

could be the result of loss over time in individual strains/species just as much as due to gain by horizontal transfer [99].

## 1.4 Risk factors of GBS colonization and infections

### 1.4.1 Neonatal infections

GBS is a leading cause of bacterial infections in newborns, such as sepsis, pneumonia, and meningitis which may lead to stillbirth, early mortality, or delayed mortality in infants. While post-infection sequelae for the survivors can lead to life-long disabilities including vision or hearing loss or cerebral palsy [100]. In newborns, GBS infections are classified into two types: early onset disease (EOD) and late onset disease (LOD) [101]. GBS-EOD is the term used for GBS infections within the first week of life (i.e. 0 to 6 days) whereas GBS LOD is used to refer to infection that develop after the first week of life, but prior to three months of age (i.e. between 7 and 90 days) [101].

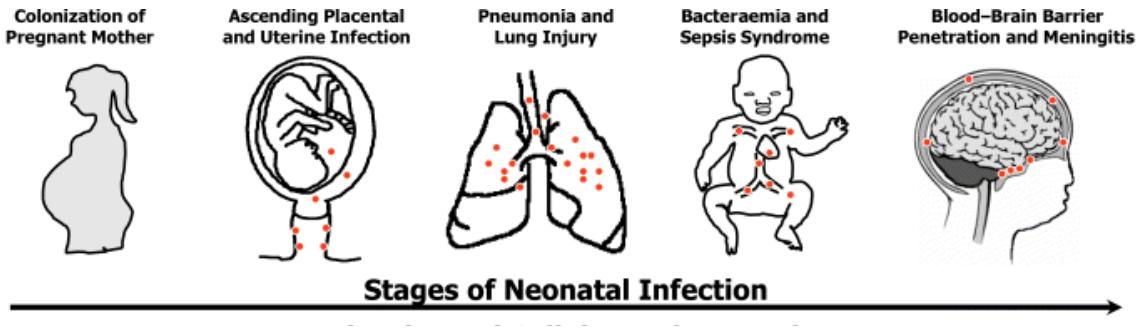



**Figure 1. 4. Stages of neonatal GBS infection.** Vaginal colonization of GBS in the pregnant woman including adherence to vaginal epithelial cells and resistance to mucosal immune defences is the first step in the pathogenesis of GBS neonatal infection. GBS may ascend into the amniotic cavity to gain access to the foetus. Bacterial proliferation allows GBS to colonize the skin or mucous membranes of the foetus or to enter the foetal lung through aspiration of infected amniotic fluid. After birth, GBS must successfully replicate within the alveoli of the neonate, adhere to respiratory epithelium, and avoid clearance by pulmonary macrophages. The invasion by GBS of the pulmonary epithelial and endothelial cells may allow GBS to enter the bloodstream causing septicaemia. This bloodstream dissemination may lead to meningitis and osteomyelitis. Adapted with permission from [102].

#### 1.4.1.1 Early Onset Disease

While the clinical syndromes, risk factors and mode of transmission of EOD and LOD can differ, the most common clinical syndromes associated with GBS-EOD are sepsis (80-85%), pneumonia (10-15%), and less frequently meningitis (5-10%) [103, 104]. In the majority of EOD and LOD cases, respiratory signs are the first clinical findings [104]. The main Risk factor for GBS-EOD is maternal carriage potentially proportional to the amount of bacterial colonization [105]. If the mother is colonized with a heavy GBS load, she is more likely to deliver a symptomatically infected infant and if the infant is heavily colonized they are more likely to have invasive disease [105]. GBS-EOD is vertically transmitted and has different stages to initiate infection (**Figure 1.4**). Infection can either occur by ascending infection passing through the cervix (including with premature rupture of membranes [PROM]) and infects the foetus *in utero* or infection can occur during delivery as the neonate passes through the mother's vagina [21]. However infants who acquire GBS infection during the birthing process are thought to become colonised only, without developing GBS-EOD [21]. About half of neonates born to a GBS-colonized mother are thought to be colonised, but only 1 to 2% of these newborns will develop EOD if no preventative measures are taken [106]. In the US, the estimated incidence rate of GBS-EOD in the early 1990s was between 0.7-3.7/1000 live births [107] while in Europe, reported incidence varied from country to country: 0.2–4 per 1000 live births in Eastern Europe [108], 0.3–2 in Western Europe [108], 0.76–2 in Scandinavia [108], and 0.57–2 in southern Europe [108]. After the adoption of antenatal screening and intrapartum antibiotic prophylaxis (IAP) in which antibiotics are administered to women with known GBS status or identified risk factors to prevent GBS transmission from mother to newborn [109], the reported incidence rate of GBS-EOD is reduced significantly in US (0.73 per 1000 live births in 1990 to 0.26 in 2010) [107], in New Zealand (from 0.5 per 1000 live births in 1998 to 0.23 in 2011) [110, 111], in Spain (2.4 per 1000 live births in 1996 to 0.26 in 2008) [112], in Vienna (5.4 per 1000 live births in 1994 to 1.1 in 1997) [113], in France (1.8 per 1000 live births in the early 1990s to 0.26 cases in 2010) (www.invs.fr), while in Canada in early 1990s' the GBS-EOD rates declined from (3 per 1000 live births to 0.5) but there was a 2-fold increase observed in Alberta (one of the thirteen provinces and territories of Canada) EOD cases from 2003 to 2013 (from 0.15 cases per 1000 live births in 2003 to 0.34 in 2013) even with implementation of the

preventive measures; it is not clear why Alberta experienced this increase [114]. In contrast surveillance data show an increasing trend of GBS-EOD in the Netherlands and the UK where risk-based prevention approach is implemented for more than two decades [115, 116]. The incidence rate of GBS-EOD was (0.31 per 1000 live births in 1999 to 0.55 in 2019) in UK [8] and (0·20 per 1000 livebirths in 1987 to 0·32 in 2011) in Netherlands [116]. In Denmark, where risk based strategy is in use, the rate of GBS-EOD remain steady during the last two decades between 2005 to 2018 with 0.18 per 1000 live births recorded in 2005 and 0.17 in 2018 [117].

Other risks factors include labour onset before 37 gestation weeks – (Prematurity will have an increased risk of infection and GBS may be an underlying cause of prematurity) [118]. Also, preterm neonates may be more immunocompromised and therefore more at risk for sepsis and invasive infection. Increased risks are also associated with longer duration of membrane rupture, very low birth weight, chorioamnionitis (inflammation of the chorion and amnion), intrapartum fever (>38 ºC or 100.4 ºF), young maternal age or HIV infection [119]. If a GBS-colonised woman undergoes preterm labour, she is at higher risk of passing GBS to her baby. Nearly a quarter (22%) of all preterm neonates born to women colonised with GBS were reported to have EOD in the UK in 2015 compared to 1/2000 of babies born at term [120]. The mortality rate from other infection also increases from 2-3% for term babies to 20-30% for those born before 37 weeks [120]. For this reason, the Royal College of Obstetricians and Gynaecologists (RCOG) guidelines now recommend all women who go into preterm labour, regardless of whether their waters have broken, receive intravenous antibiotics during labour to prevent onset of the GBS infection [3].

#### 1.4.1.2 Late Onset Disease

In 2015, more than 50% of reported cases of neonatal GBS disease in the UK occurred during the late-onset period [121, 122]. Although some neonatal infection cases suggest a maternal source of acquired GBS infection, arising from infection during their passage through the birth canal, the pathogenesis of LOD is poorly understood [21]. Another suggested source for LOD is the ingestion of contaminated

mother's breast milk [123]. Although the capsule type in 50% of mothers of GBS LOD infected infants were the same GBS isolated from their baby, it was disparate in the remaining infants and the source of infection was unclear [21]. It is generally presumed that GBS LOD are not vertically transmitted from mothers, but horizontally transmitted by hospital and community sources [124]. Unlike GBS-EOD, risk factors for GBS LOD have not been well identified, however prematurity is considered as the major risk factor for LOD [120]. GBS LOD sequelae include meningitis and bacteraemia [125]. Blood cultures from febrile newborns are required for definitive LOD diagnosis [126]. Long-term consequences of GBS meningitis in surviving neonates are hearing loss, blindness and mental impairment [100].

### 1.4.2 Infections in pregnancy

Many risk factors are associated with asymptotic GBS colonization among pregnant women including maternal age and education, PROM, frequency of pregnancy and sexual intercourse, number of sex partners and smoking [127-129]. Diverse clinical presentations including urinary tract, intra-amniotic or wound infections, endometritis and less frequently meningitis and puerperal sepsis are observed in pregnant women with GBS infections [130, 131].

### 1.4.3 Infections in non-pregnant adults

Over the last two decades, GBS has been increasingly recognised as a pathogen causing invasive disease in non-pregnant adults [132]. Adults of advanced age (i.e. >65 years of age), especially those with underlying diseases including diabetes mellitus, neurologic impairment, cancer and liver disease are at primary risk for GBS infection [21]. Clinical presentation of GBS disease in adults are varied: bacteraemia, soft tissue and skin infections, urinary tract infections and osteoarticular infection are more frequently observed [133]. Less frequently observed are meningitis and endocarditis, but these latter invasive infections are associated with higher rates of morbidity and mortality [134, 135].

In a recent report, association of cystic fibrosis with possible emergence of GBS as a respiratory pathogen was proposed [136] where GBS was noticed in respiratory secretions of 30/185 (16%) of cystic fibrosis patients. Pathogenesis of adult GBS infections are complex due to diverse clinical presentations and poor outcomes in adults [137]. In previous studies, diabetic patients aged between 20 to 64 years were found to have an increased risk (11 to 30 fold) of iGBS [21] and correlations to advanced patient age and black race with GBS infection and colonisation in non-pregnant adults was also observed [132].

## 1.5 Colonization and Transmission

The natural reservoir for GBS is the gastrointestinal tract; also considered as the source of vagino-rectal colonization. Vaginal or rectal GBS colonization is observed in approximately 10 to 30% of pregnant women [129]; however, this includes transient, chronic or intermittent GBS modes of colonization [129]. A requirement for GBS-EOD is the GBS carriage in the mother's gastrointestinal or genital tracts and preventive measures should be based on studying the different prevalence of GBS colonization in mothers of different demographics [138].

GBS colonization rates among non-pregnant adults is less well studied, however healthy young and old adults have reported rates of vagino-rectal GBS colonization between 20 to 34%, similar to those observed in pregnant adults [139-141]. During pregnancy, transmission of GBS is likely to happen between sexual partners; however, other modes of transmission may also exist [141]. Limited GBS species transmission from livestock to human is suggested in numerous studies [142-144]. But GBS is a major pathogen of males in 40-50 y age group consuming raw fish in some geographies including United States and China [145, 146].

## 1.6 Guidelines for Maternal GBS screening and IAP strategy in different countries

The transmission of GBS from mother to baby can be reduced with the use of IAP [147]. There is, however, uncertainty about the most effective way of identifying women who should receive IAP [147]. One approach is to offer IAP to women with one or more risk factors and another approach is to screen all pregnant women using a bacteriological test at 35–37 weeks and offer IAP to women who are found to be colonised with GBS [147]. It is also possible to combine these two methods to form other strategies [147]. In the sections below, preventive strategies implemented in US, Australia, Brazil, and UK to reduce burden of GBS disease in pregnant women and neonates (**Table 1.2**) are described in detail, from where the subset of GBS were also used in this thesis to study serotypes distribution, potential capsular switching events, mobile elements carrying antibiotic resistance determinants and association of genes to a particular lineage.

### 1.6.1 In the United States

In the mid-1960s IAP was suggested as a preventive measure of GBS neonatal infections, followed by the proposal of selective GBS carriage testing in pregnant women a few years later [148]. The consensus guidelines for the prevention of GBS infections in neonates were first published by American College of Obstetricians and Gynaecologists (ACOG) in 1996, and was quickly followed by CDC and American Academy of Paediatrics (AAP) which identified IAP candidates by a risk-based strategy [149]. After publication of the 1996 guidelines, a 65% decrease in EOD incidence rates was observed in the United States [150]. In 2002, revision of the GBS prevention guidelines took place replacing risk based strategy to the universal culture-based GBS screening for all pregnant women at 35 to 37 weeks of gestation and IAP administration to GBS colonised mothers [151]. This resulted in a further 50% reduction in EOD incidence rate in the United States, for the period 2003-2005 relative to 1999-2001 [151, 152]. In 2010, the guidelines were once again revised by CDC and continued to recommend universal antenatal culture-based approach to identify women with positive GBS and offer IAP to those found positive to prevent GBS-EOD [153]. Significant changes were made in the 2010 revision, dictating the

implementation of IAP for women in preterm labour and those with PROM, the selection of antibiotics for IAP, and the use of nucleic acid amplification tests (NAATs) for the identification of maternal GBS colonization. Following universal maternal antenatal screening and IAP implementation, drastic decline in the rate of GBS-EOD cases were seen from 1.8/1000 live births to 0.23/1000 live birth in 1993 and 2015, respectively [154].

In partnership with numerous professional societies, the CDC has published recommendations for GBS screening since 1996 [155, 156]. In 2019, the stewardship of these guidelines was transferred to three professional organizations. ACOG and AAP are now accountable for curation of the guidelines for prophylaxis and treatment of GBS infection in pregnant women and newborns, and the American Society for Microbiology (ASM) is responsible for maintaining and updating guidelines for standard laboratory practices related to detection and identification of GBS [155, 156]. Most recently in 2017, the ASM published a new guideline for the appropriate implementation of GBS-EOD screening and IAP which replaces the earlier 2010 guidelines published by CDC [2]. These guidelines states that:

i. Antenatal screening is recommended for all pregnant women at 36 0/7 to 37 6/7 weeks of gestation unless IAP is already recommended due to identified risk factors.
ii. All women with positive vagino-rectal cultures at 35-37 weeks of gestation should be given appropriate IAP unless a prelabour caesarean birth is performed in case of intact membranes.
iii. IAP is not needed for women who present in labour with unknown GBS colonization status if they have a negative intrapartum GBS nucleic acid amplification testing (NAAT) result and have no clinical risk factors.
iv. When labour starts, if the result of the prenatal GBS culture is unknown, IAP should be given for women who have following risk factors for GBS-EOD including preterm birth, PROM, rupture of membrane for more than 18 hours at term or have intrapartum fever (100.4ºF or higher). Broad spectrum antibiotic therapy should be provided for women with suspected intraamniotic infection.
v. If the GBS status of a woman in labour at term is unknown and the woman does not have any of the risk factors but has a history of GBS colonization in previous pregnancy, then IAP is reasonable to offer based on the woman's history of colonization.

vi. Penicillin remains the first choice of antibiotic to be administered intravenously for IAP with ampicillin as an appropriate alternative. For penicillin allergic women, clindamycin is the recommended alternative only if the clindamycin susceptibility has already been tested.

vii. In case of clindamycin resistance and high-risk penicillin allergic women, vancomycin remains the only microbiologically and pharmacokinetically recommended option for IAP. The dosage for vancomycin should be 20mg/kg and given intravenously every 8 hours with a maximum of 2g/single dose.

### 1.6.2 In United Kingdom

The RCOG published its Green-top guideline in November 2003 to provide guidance for obstetricians, midwives, and neonatologists on the prevention of neonatal GBS-EOD [157]. In 2017, an audit was conducted, intended to assess practice in UK obstetric units on the prevention of GBS-EOD in neonates relative to the proposals of the first edition of the relevant RCOG 2003 Green-top guideline [157]. More recently in 2017, these guidelines were updated to prevent early-onset neonatal infection before birth in UK [3] which states that:

i. Risk based routine screening for antenatal GBS carriage is recommended.

ii. If the women show any of the known risk factors for GBS including incidental carriage, premature birth < 37 weeks, prolonged rupture of membranes (18 hours or more), intrapartum fever > 38ºC then the clinicians should discuss the use of IAP. If the women appear to have two or more risk factors, then the argument for prophylaxis becomes stronger.

iii. Women with positive vaginal or urine cultures for GBS should be offered IAP.

iv. IAP should be offered to women who have had a previous baby diagnosed with neonatal GBS disease. However, there is no good indication to endorse the administration of IAP to women with positive GBS carriage in a previous pregnancy or to women undergoing planned caesarean section.

v. IAP is not recommended for women with PROM unless they are in active labour.

vi. Benzylpenicillin (penicillin G) is the first choice of therapeutic for IAP and should be administered intravenously immediately after the onset of labour and at least 2 hours before delivery. The recommended dose is 3g followed by 1.5g every 4 hours during labour. Cephalosporin should be used (e.g., Cefuroxime 1.5g loading dose followed by 750 mg every 8 hours) unless she has had a severe allergic reaction (swelling of the skin or throat, difficulty

breathing, and/ or fainting/low blood pressure), in which case, vancomycin (1g every 12 hours) should be used.

### 1.6.3 In Brazil

IAP was implemented for the first time in Brazil in January 2014, based on the most recent CDC guidelines described in 2010 [159]. Previously, the Brazilian Society for Paediatrics [160, 161] had recommend a culture-based policy to select pregnant women to be offered IAP; however, compliance to such guidelines had been very poor (around 20%) and as a result not much is known about GBS carriage in Brazil [147].

### 1.6.4 In Australia:

In the late 1970s, a GBS control strategy was formulated in Australia, based on unexpectedly large number of EOS for GBS reported in one city [162]. As a result of a review into local GBS-EOD rates in a large metropolitan Melbourne hospital, the subsequent policy recommended screening all pregnant women for GBS, with provision of IAP to those at risk [162]. This review influenced GBS management throughout all Australia. However, a unified national Australian GBS policy has never been implemented and individual territories are allowed to recommend different approaches for selecting women for IAP [163]. In the state of Queensland, for example, a risk-based approach is implemented [163] while the state of New South Wales recommendations vary from universal screening to a risk-based approach depending on locality [164]. The most recent guidelines from the Royal Australian and New Zealand college of Obstetricians and Gynaecologists (RANZCOG) in 2016 [165] also recommend either approach. A recent Australian systematic review concluded that the odds of GBS-EOD from infants infected with GBS were significantly lower with universal screening compared with the risk-based screening approach (OR 0.45, 95% CI, 0.37–0.53) [166]. However, the authors qualified their conclusions based on disparity of the quality of the studies critiqued, with many being quite poorly designed [23]. The 2016 guidelines by RANZCOG for Maternal GBS screening and management in Australia [165] stated:

i. Universal culture-based screening, using combined low vaginal, plus or minus anorectal, swab at 35 - 37 weeks of gestation, or 3-5 weeks prior to anticipated delivery in high-risk pregnancy such as poorly controlled diabetes, multiple pregnancy, OR a clinical risk factor-based approach, are both acceptable strategies for reducing EOS caused by GBS.
ii. If the GBS status is unknown at the time of labour onset, then treatment according to the clinical risk factors is appropriate.
iii. IAP with IV penicillin-G or ampicillin should be administered intravenously to all women at risk.
iv. For penicillin allergic patient, antibiotic susceptibility test should be performed. Appropriate alternatives include cefazolin, clindamycin, and vancomycin depending on the antibiotic resistance and nature of previous adverse reaction to penicillin.

**Table 1. 2. GBS management and prevention guidelines in different countries.**

| **Country** | **Universal screening** | **Risk based screening** | **IAP to mothers with unknown GBS status** | **IAP to mothers with identified risk factors @** | **Choice of antibiotic** | **For women with severe penicillin allergy** |
|---|---|---|---|---|---|---|
| US | Yes | No | No | Yes | penicillin | clindamycin* or vancomycin |
| UK | No | Yes | No | Yes | penicillin | vancomycin |
| Brazil | Yes | No | No | Yes | penicillin | clindamycin* or vancomycin |
| Australia | Yes ^ | Yes ^ | No | Yes | penicillin or ampicillin | cefazolin*, clindamycin*, or vancomycin* |

@ risk factors include (PROM, preterm birth, rupture of membrane for > 18h at term, previous history of GBS colonisation and fever > 99F)
* Only after susceptibility has been tested
^ In Australia, either approach (risk or universal screening) is recommended [165].

### 1.6.5 Limitations of IAP

For the prevention of GBS infections in neonates, screening approaches have always proved challenging [167]. However, for all women who are about to deliver and have high risk factors, especially those whose GBS colonisation status is unknown, IAP should be offered to ensure the

likelihood of GBS-EOD is minimised [167, 168]. Significant reductions in GBS-EOD mortality and morbidity have consistently been observed after widespread use of IAP and IAP has also reduced iGBS infections among pregnant girls and women in the US by 21% [150]. However, several problems remain regarding the management and prevention of GBS neonatal infections: First, IAP has no effect on the incidence of LOD rates, which have remained stable or increased in different geographical locations [152, 169, 170]. Romain *et al* studied the effect of IAP in reducing GBS disease in neonates born in France between 2001 and 2014. Among 848 GBS meningitis cases from 2001 to 2014, the incidence of EOD decreased by 63.3% and that of LOD increased by 58.1% [170]. It has been suggested that IAP might delay GBS-EOD onset instead of preventing it, hence emerging instead as an LOD infection [107]. Secondly, widespread use of antibiotics for IAP may have adverse effects: i.e. increase emergence of resistant bacterial strains to IAP antibiotics (GBS or other species) [171], can result in anaphylaxis which is still very low in UK 0.8/100,000 IAP administration in UK [172], can alter microbiota of the newborn [173] and can lead to increased neonatal infections by other pathogens in place of GBS [174].

## 1.7 Antimicrobials effectiveness against GBS

### 1.7.1 β-lactams

Treatment of different streptococcal infections is mainly based on the use of penicillin V administered orally, intramuscular benzathine penicillin, parenteral penicillin G, amoxicillin or cephalosporins (cephalexin, cefotaxime, ceftriaxone) [175]. To date, after 70 years of use, penicillin is still universally active against GAS, GBS, GCS and GGS, although since 2008, GBS strains with reduced susceptibility to this antimicrobial agent have been described in Japan, USA, UK and Canada [176-179]. The reported prevalence has been high in Japan: 2.3% between 2005–2006, increasing to 14.7% between 2012 and 2013 [180]. Among 14.7% (45/306) isolates with reduced penicillin susceptibility (PRGBS), 71.1% (31/45) were resistant to erythromycin and 95.6% (43/45) were resistant to levofloxacin, respectively (with 68.9% resistant to both antibiotics) [180]. PRGBS were recovered from a range of sample types

including respiratory samples from elderly patients in addition to a couple of invasive samples from neonates and adults [180, 181]. Supporting genomic analysis data is very limited. Kimura *et al* found that ST1 and a single *locus* variant ST458 were predominant clones for reduced penicillin susceptibility recovered between 1998 and 2003 periods in Japan, while results obtained in USA identified responsible penicillin-binding mutations in isolates [177, 182]. GBS is genetically very isolated and horizontal gene transfer in GBS are exceptionally rare due to plasmids, which are uncommon in GBS [183]. This also explains why penicillin resistance and sulphonamide resistance genes are not seen in GBS despite being the most common resistance genes found in other bacterial populations.

#### 1.7.1.1 Mechanisms of β-lactam resistance in GBS

Penicillin resistance in gram positive organisms is due to the production of altered, low-affinity target penicillin-binding proteins (PBPs) that catalyse the terminal stage of bacterial cell wall peptidoglycan synthesis [175]. In PBPs, three conserved motifs, SXXK, SXN, and KT(S)G, commonly found in transpeptidase domains form the catalytic center, and alterations within or adjacent to these motifs are associated with streptococcus reduced affinity for β-lactams [184]. Specifically, in GBS with reduced susceptibility to penicillin, the most reported substitutions are V405A and/or Q557E in PBP2X [177, 178, 185]. Moreover, an amino acid substitution (T145A) in PBP1A confers high-level resistance to cephalosporins [186]. These mutations correlate to their penicillin MIC levels [186]. Recently ceftibuten resistant, but penicillin susceptible GBS with amino acid substitutions in PBP2X were reported [187].

### 1.7.2 Macrolides and Lincosamides

Macrolides have importance as a main alternative antibiotic in treating GBS infected patients who have serious penicillin allergies [153]. However, for penicillin-allergic patients with low risk to anaphylaxis, cefazolin can be used as an alternative therapeutic [188]. Clindamycin (a lincosamide) given intravenously is also recommended as an alternative for penicillin-allergic patients at high risk of anaphylaxis [188]. Intravenous vancomycin remains the only option for intrapartum antibiotic

prophylaxis (IAP) in women with a high-risk penicillin allergy, when the GBS isolate is additionally non-susceptible to clindamycin [188]. In recent years, increasing GBS resistance to erythromycin, clindamycin (more so than resistance to fluoroquinolones, high-level resistance (HLR) to gentamicin, or decreased susceptibility to vancomycin) have been reported worldwide [78, 189-191].

#### 1.7.2.1 Mechanisms of macrolide and lincosamide resistance in GBS

Macrolides (such as erythromycin or azithromycin) bind to bacterial 50S ribosomal subunit to inhibit bacterial growth by impeding the passage of newly synthesized polypeptides through the nascent peptide exit tunnel [192]. Whereas the lincosamides (such as clindamycin) also bind to the 50S ribosomal subunit but directly interact with, and block, the ribosomal peptidyl transferase center (PTC), where amino acids are assembled into proteins by peptide bond formation [192]. GBS strain resistance to erythromycin and/or clindamycin is attributable to several mechanisms (**Table 1.3**), most frequently, acquisition of genes including *erm* (erythromycin ribosome methylase) encoding a methylase that modifies the antibiotic binding target site on the 23S rRNA conferring resistance to macrolides, lincosamides and streptogramin B (MLSB phenotype), which can be either constitutive ($cMLS_B$) or inducible ($iMLS_B$) [193]. In $iMLS_B$, inactivate methylase is produced by bacteria that, in the presence of macrolide (inducer), become active. If the inducer (erythromycin) is present, mRNA rearrangement occurs allowing translation of the methylase coding sequences. On the other hand, in $cMLS_B$, production of active methylase mRNA occurs when the inducer is absent [193, 194]. The second significant resistance mechanism is the active efflux of antibiotics via an ATP-dependent efflux pump (ABC) (**Table 1.3**), such as those encoded by *mef*/*msr* (*msr*A, *msr*B and *msr*C) genes that mediates 14-membered and 15-membered macrolide and streptogramin type B resistance [195]. Further, *lsa* (A and C) genes encode an active efflux ABC transporter-like transmembrane proteins and mediates lincosamide and streptogramin type A antibiotic resistance. The third mechanism, as exemplified by *lnu* (A, B, C D, E and F) genes, inactivate the antibiotic through direct modification (**Table 1.3**). The *lnu* gene family catalyse the adenylation of lincosamides and streptogramin type B by encoding lincosamide nucleotidyl transferases enzymes [196]. This resistance determinant was earlier found in a

Canadian GBS isolate which was susceptible to erythromycin but resistant to clindamycin, then later in a Korean GBS isolate carrying both *lnu*B and *ermB* genes on the same mobile genetic element (MGE) [197]. Furthermore, cross-resistance to lincosamides, streptogramin A and veterinary pleuromutilins ($LS_AP$; another 23S rRNA binding protein inhibiting antibiotic) is mediated by the acquisition of *lsa* genes (**Table 1.3**) [198].

For GBS, macrolide resistance is most frequently mediated by genes belonging to the *erm*B class [199, 200], and to a lesser extent to *erm*TR – a subset of *erm*A class of enzymes [201]. Specific association between *erm*B/$cMLS_B$ and *erm*TR/ $iMLS_B$ have been reported in many studies [202, 203]. Additionally, the related, plasmid carried *erm*T gene was initially identified as $iMLS_B$ in a *Streptococcus bovis* strain [204] but recent reports of macrolide resistant human GBS isolates were also found positive for ermT [205, 206]. Among M-phenotype GBS isolates (i.e., macrolide resistant) other efflux genes include mef (A and E) genes [202, 207]. A wide consensus has established an association for macrolide resistance to serotype V GBS notwithstanding variations in antibiotic resistance that occur geographically [202, 208].

### 1.7.3 Aminoglycosides

Aminoglycosides belong to complex family of antibiotic compounds, largely used for treating infections caused by gram-negative aerobic bacilli, staphylococci, and other Gram-positive pathogens [209]. However, when used against Gram-positives, aminoglycosides are recommended to be used in combination with other antibiotics such as β-lactams or vancomycin with which they employ a synergistic effect [210]. There are various types of aminoglycosides, each vary in their activity spectrum [211]. Aminoglycosides alone have little or no effect on GBS, but synergistic killing with penicillin has been shown in vitro [212]. The EUCAST database show that ECOFF (epidemiologic cut-off) values for the aminoglycoside gentamicin is too high (64 mg/L) for therapeutic use, and GBS isolates with high-level resistance (HLR; >128mg/L) have been reported for isolates in the UK, South Africa and France [213-215]. While routine screening for HLR to gentamicin is clinically pointless, surveillance

does elucidate movement of associated mobile genetic elements. It is standard clinical practices and guidelines for the management of neonates to be prophylactically treated with penicillin or ampicillin in combination with gentamicin [216]. Further suspected or proven early-onset sepsis often defaults to penicillin-aminoglycoside combination therapy, although the synergistic effect of these antibiotics is still under debate [217, 218].

#### 1.7.3.1 Mechanisms of aminoglycoside resistance in GBS

Aminoglycoside resistance occurs through several mechanisms: modification of the target ribosomal protein or 16S rRNA [219], methylation of 16S rRNA [220], modification of outer membrane's permeability to diminish inner membrane transport [221], by active efflux pump [222], active swarming [222] or enzymatic modification of the antibiotic molecule [223]. The most common aminoglycoside resistance mechanism observed in all bacteria is the enzymatic modification [224]. Three different types of antibiotic modification enzymes (AMEs) namely N-Acetyltransferases (AAC), O-Adenyltransferases (ANT) and O-Phosphotransferases (APH) are encoded by genes usually found on plasmids and transposons [224, 225]. Acetyltransferases catalyse acetyl CoA-dependent acetylation of an amino group and are encoded by the following genes: *aac*(3)-(Ia, Ib, Iia, Iib, Iic, IIIa-c, Iva, Via), *aac*(6')-(Ia-Ii), *aac*(6')-*aph*(2") and *aac*(2')-Ia [225, 226]. Adenyltransferases catalyze ATP-dependent adenylation of a hydroxyl group and are encoded by: *ant*(2")-(Ia-Ic), *ant*(3")-Ia, *ant*(4')-Ia, *ant*(4')-Iia, *ant*(6)-Ia, *ant*(6), *aad*E and *ant*(6)-Ib [225]. Whereas phosphotransferases catalyse ATP-dependent phosphorylation of a hydroxyl group and are encoded by: *aph*(2")-Ia, *aph*(3')-(Ia-Ic, Iia, III, Iva, Va-Vc, Via-Vib and VIIa), *aph*(3") -(Ia, Ib), and *aph*(6)-(Ia-Id) genes [225]. The most common aminoglycoside genes encoding AMEs in GBS are *aad*E, *aph*(3-III) and *ant*(6-Ia) (**Table 1.3**) [227].

**Table 1. 3. Most frequently observed antibiotic/bacteriocin resistance genes in GBS.**

| Gene | Antibiotic/Bacteriocin targeted | Resistance mechanism |
|---|---|---|
| *aad*E | Streptomycin | Aminoglycoside modification |
| *aph*(3'-III) | Kanamycin | Aminoglycoside modification |
| *ant*(6-Ia) | Streptomycin | Aminoglycoside modification |
| *aac*(6')-*aph*(2") | All aminoglycosides except streptomycin | Aminoglycoside modification |
| *erm*A | Macrolides, lincosamides and streptogramins B | Ribosomal target modification |
| *erm*B | Macrolides, lincosamides and streptogramins B | Ribosomal target modification |
| *erm*TR | Macrolides, lincosamides and streptogramins B | Ribosomal target modification |
| *erm*T | Macrolides, lincosamides and streptogramins B | Ribosomal target modification |
| *lnu*B | Lincosamides | Lincosamide nucleotidylation |
| *lnu*C | Lincosamides | Lincosamide nucleotidylation |
| *lsa*C | Lincosamides and streptogramins A | Ribosomal protection |
| *mef*A/*msr*D | Macrolides | Efflux pump |
| *tet*M | Tetracycline | Ribosomal target modification |
| *tet*O | Tetracycline | Ribosomal target modification |
| *tet*L | Tetracycline | Efflux pump |
| *tet*K | Tetracycline | Efflux pump |
| *tet*S | Tetracycline | Ribosomal target modification |
| *tet*W | Tetracycline | Ribosomal protection |
| *cat*Q | Chloramphenicol | Inhibition of bacterial protein biosynthesis |
| *cat(*pc194) | Chloramphenicol | Inhibition of bacterial protein biosynthesis |

### 1.7.4 Tetracycline

Resistance to tetracyclines is usually attributed to one or more of the following: the acquisition of mobile genetic elements carrying tetracycline-specific resistance genes, mutations within the ribosomal binding site, and/or chromosomal mutations leading to increased expression of intrinsic resistance mechanism [228]. This antibiotic is no longer used to treat GBS infections because resistance to tetracycline among GBS isolates is frequently found at very high rates [229]. However continuous surveillance of tetracycline specific resistance genes (*tet*) in GBS strains is of potential use, because

often mobile genetic elements that harbour *tet* genes also carry other resistance genes for macrolides and chloramphenicol [230].

While *tet*M (a ribosomal protection protein) is the most frequent tetracycline resistance gene identified among human GBS isolates, *tet*O (also a ribosomal protection protein) is more frequently observed in bovine GBS isolates. The genes *tet*K or *tet*L are less commonly observed and have a different mechanism of resistance as they encode tetracycline efflux pumps. Whereas *tet*T and *tet*W are less common ribosomal protection group genes (similar mechanism to *tet*M and *tet*O) and are reported rarely in human GBS isolates [203] (**Table 1.3**). In many macrolide and tetracycline resistant GBS strains, *tet*(O) and *erm* genes are commonly carried by the same genetic element (transposon) supporting the epidemic spread of multi-resistant strains [203, 231, 232], whereas in other tetracycline resistant strains (usually *tet*M) that are not macrolide resistant, it supports the independent dissemination of resistance [230].

Tetracycline was first used clinically in 1948 and became widely used, but the first reported case of tetracycline resistance was for *Shigella dysenteriae* isolated in 1953 from Japan [228]. In a contemporary study, Da cunha *et al* [1] performed advanced genomic analysis of the tetracycline resistant GBS populations isolated between 1953 and 2011 spanning four continents. Despite the high prevalence of tetracycline resistance globally, there is a lack of heterogeneity found for the genomic placement and sequence of the most prevalent *tet*M resistance gene. These authors concluded that vertical gene transmission of tetracycline resistance conserved within clonal clades and not horizontal gene transfer, as the primary mechanism underlying common tetracycline resistance [1]. These results strongly suggests that removal of tetracycline sensitive strains resulted in the limited diversity of GBS population seen today, whereas selection of virulent strains explains the emergence of GBS as serious infant pathogen during 1960's – 1970's [1].

### 1.7.5 Quinolones

The quinolones interfere with the bacterial DNA replication and transcription processes by targeting enzymes (DNA gyrase and topoisomerase IV) involved in bacterial DNA supercoiling and progeny circular genome separation [233]. Single point mutations in the quinolone resistance-determining region (QRDR) of in *gyr*A confers low level resistance to early generation fluoroquinolones (e.g. nalidixic acid) while mutations in both *gyr*A and QRDR *par*C results in higher MICs and resistance to later generation fluoroquinolones (e.g. levofloxacin, moxifloxacin) [234, 235]. Double point mutations in the *gyr*A leading to (e.g. Ser81Leu; Glu85Ala/Lys) amino acid substitutions, and *par*C (e.g. Ser79Phe/Ala/Tyr) amino acid substitutions commonly confer fluoroquinolone resistance in GBS [181, 236], which was first described in 2003 in Japan [235]. While, fluoroquinolones resistant GBS are less frequently reported in in Taiwan (1.3%) [237] and in France (0.7%) [214] , they are more common in Canada (12%) [238], Argentina (14.8%) [239], and China (32.4%) [240].

### 1.7.6 Chloramphenicol

Inhibition of bacterial protein synthesis occurs when chloramphenicol binds to A2451 and A2452 residues in the 23S rRNA of the 50S ribosomal subunit, preventing peptide bond formation and protein chain elongation by inhibiting the peptidyl transferase activity of the bacterial ribosome [194]. The mechanism of chloramphenicol resistance is based on enzyme synthesis (chloramphenicol acetyltransferase) encoded by *cat* genes (*cat*Q and *cat*(pc194)) which acetylates chloramphenicol and prevent its binding to ribosome [241]. In GBS, *cat* genes have been observed both on transposons and on plasmids [241]. Frequency is still relatively rare with (1%, 4/368) Italian chloramphenicol resistant isolates in one study [242] and (24.9%, 21/128) South African chloramphenicol resistant isolates in another [215]; however, limited clinical utility of chloramphenicol equates to rare evaluation for resistance.

## 1.8 Prevention of GBS infections – an alternative to antibiotics

For the prevention of GBS infections in non-pregnant adults and neonates, the most promising approach is the development of vaccines [243, 244]. Experimental and clinical evidence has shown specific antibodies against specific surface proteins of bacteria are protective [245]. However, in pregnant women (85 to 90%), lack of naturally occurring antibodies is seen at the time of delivery [246]. To confer immunity to peripartum maternal infection and passive immunity to the newborn (as IgG antibodies are actively transferred across the placenta) a maternal GBS vaccination program given before mothers conceive or while they are pregnant has been contemplated [21, 245]. Generally, it is understood that vaccination would confer protection to both pregnant women and their newborns providing better health benefits, compared to a non-vaccinated population and to those elderly adults (> 65 years) at greater risk for serious GBS diseases [247]. Experimental studies suggests that mothers immunised with capsular polysaccharide (CPS) conjugated vaccines may decrease the risk of developing GBS diseases in neonates and young infants in a serotype specific manner [248-251]. Further, surface protein-based vaccines are also under evaluation and have the potential to provide protection across the serotype spectrum. **Table 1.4** summarizes the development status of current vaccine candidates.

**Table 1. 4. Summarizing candidates for GBS vaccine.**

| Vaccine candidates | Preclinical | Phase I | Phase II | Trials in Pregnancy | Phase III |
|---|---|---|---|---|---|
| Monovalent and bivalent conjugates (TT/$CRM_{197}$- CPSS) | x | x | x | x | |
| Trivalent $CRM_{197}$-CPS* | x | x | x | x | |
| Hexavalent $CRM_{197}$-CPS | x | x | x | x | |
| N-terminal domains of the Rib and AlphaC proteins | x | x | | | |
| Pilus proteins | x | | | | |
| Other proteins | x | | | | |

* $CRM_{197}$-CPS means CPS-cross-reactive material (CRM197) conjugate vaccine.

### 1.8.1 CPS conjugate vaccines

CPS was initially identified as a potential vaccine candidate in the 1960s when Rebecca Lancefield discovered a protective role of CPS specific antibodies in prevailing and sustained GBS infection in mice [252]. Further investigations revealed that vaccination using CPS alone was well tolerated and safe. However, it had insufficient immunogenicity when used alone and demonstrated that CPS might elicit a better immune response when conjugated with a carrier protein [253].

The CPS polysaccharide, when conjugated with a protein carrier, can act as a T-cell independent antigen and can trigger both a protective and a memory B-cell response [254, 255]. $CRM_{197}$ is currently used as the main protein carrier for GBS conjugate vaccine; it utilises a non-toxic variant of diphtheria toxin which is highly immunogenic [256, 257]. Studies demonstrated that compared to unconjugated vaccines, conjugated vaccines were found to have better immunogenicity with high levels of protective antibodies [256, 257]. The monovalent CPS Ia, Ib, II, III and V vaccine were first tested in 2003 under clinical Phase I and II trials in healthy adults [257-261]. However, single serotype vaccines do not generally produce cross reactivity with other different GBS capsule serotypes. Therefore, distinct serotype vaccination offered protection only to homologous serotypes and emphasized the need for multivalent conjugative vaccine development to cover a spectrum of different GBS serotypes [245, 262]. Thus multivalent mixture vaccines were formulated, a trivalent (Ia, Ib and III) $CRM_{197}$ conjugate vaccine underwent a phase I/II clinical trial (registered as NCT01193920 at the ClinicalTrials.gov database) in pregnant women and results published in 2014 showed higher levels of CPS-specific antibodies in infants at birth with no safety concerns [262]. More recently, in 2017 [263, 264], clinical trials of a hexavalent vaccine (Ia, Ib, II, III, IV, and V) (NCT03170609) in non-pregnant healthy adults were started. This vaccine included serotype IV due to emerging reports of iGBS disease in adults caused by serotype IV. The overall aim was to provide a good population coverage since these six GBS serotypes account for (93-98)% of neonatal, pregnant women and adults iGBS disease reported globally [265]. To also verify immunogenicity and clinical safety of this hexavalent vaccine in pregnant women and their infants [266]. In 2019, phase I and II clinical trial (NCT03765073) were started in healthy

non-pregnant women, pregnant women and their infants to assess its immunogenicity profile, safety and tolerability [266]. The estimated completion date of this study is March 2022.

The limitation associated with the use of capsular polysaccharide conjugated vaccines is the potential for serotype replacement and capsular switching [267]. In 2008, post pneumococcal vaccine implementation, capsular switching was reported in pneumococcal population as a result of removal of most of the antibiotic resistance types of *S. pneumoniae* that led a global decrease in AMR-associated or more virulent strains [267]. A similar response of capsular switching could be possible in the future following implementation of GBS vaccination [247]. Temporal changes in serotype distributions and the contribution of capsular switching to such changes have been reported [78, 268]. Currently, it is unclear to what extent vaccine-induced selective pressures drive capsular switching, it is assumed that vaccines could provide the selective pressure for virulent genotypes to switch capsules and escape vaccine coverage [268, 269].

### 1.8.2 Protein vaccines

To formulate a non-CPS based GBS vaccine, other surface proteins, namely the α- and β-antigens of protein C [270], *scp*B (C5a peptidase*)* [271] and *rib* [272] were taken into consideration, which showed elevated antigenicity and as virulence associated determinants, do not have the CPS-associated risk of replacement/switch. A phase I clinical trial (NCT02459262) has been conducted on a protein vaccine based on the highly immunogenic N-terminal domains of AlphaC and Rib (GBS-NN) in 240 healthy women [273, 274]. A 30-fold increase in GBS-NN specific antibodies was observed in these participants who were immunised with one or two doses of GBS-NN [273, 274]. However, no data was available to determine if this was protective against GBS infection.

GBS expresses two forms of surface anchored glycosylated serine rich proteins previously designated as (srr-1 and srr-2) repeats. Through a “dock, lock and latch” mechanism these bind to the Aα chain of fibrinogen and contribute to development of GBS meningitis and colonization of the vaginal surface

[275]. Serotype independent protection was observed in murine models after being vaccinated with the latch-peptide vaccine whose antigenic "latch" domain comprised of 13 amino acids containing both srr1 and srr2 motifs that were shown to play a significant role in GBS pathogenesis [276]. Another GBS virulence factor, C5a peptidase, was also considered as a universal protein vaccine or a putative conjugate candidate for CPS vaccines [277]. C5a peptidase has been shown to trigger systemic and mucosal immune responses in murine models, that may produce a protective immune response against against multiple GBS serotypes [276, 278].

Pili-like structures on the surface of bacteria are another potential candidate for vaccine development because of their role in pathogenesis and their common expression in almost all GBS strains [279]. In 2009, newborn mice were immunized with a combination of the three-known pilus-islands, which mediated protective immunity against all GBS strain serotypes [280]. Furthermore, the role of pili in promoting bacterial adherence to host tissue suggests that pilus-based vaccines might also elicit antibodies capable of preventing colonization [280].

## 1.9 Virulence factors

GBS is classified as a pathobiont, a term that means bacteria is normally identified as a colonising bacteria causing no pathology, but under some circumstances (which may include any change in the specific genetic or environmental condition of the host) it is identified as the infectious agent present associated with pathological sequelae [281]. GBS encodes several virulence genes that are critically important for its pathogenesis (**Table 1.5**) [282] and mediates a range of effects: facilitate adherence to host surfaces, invade across endothelial and epithelial barriers, and interfere with innate immune clearance mechanisms, to adapt host environments during infection [283].

**Table 1. 5. Key virulence factors of GBS**. Reproduced from [284].

| **Virulence factor** | **Genes** | **Molecular or cellular action** | **Proposed contribution to pathogenesis** |
|---|---|---|---|
| Polysaccharide capsule | *cps*A-L<br>*neu*A-D | Impairs complement C3 deposition and activation<br>Masks pro-inflammatory cell wall components<br>Decreases immune recognition, perhaps through molecular mimicry of host sialic acid epitopes | Blocks opsonophagocytic clearance, preventing the recognition of GBS through molecular mimicry of host cell surface glycoconjugates<br>Delays neutrophil recruitment |
| β-Haemolysin/cytolysin | *cyl*E | Forms pores in cell membranes<br>Induces inflammatory response and apoptosis<br>Triggers cytokine release" | Direct tissue injury<br>Penetration of epithelial barriers<br>Induction of sepsis syndrome<br>Phagocytic resistance<br>Impairs cardiac and liver function |
| Hyaluronate lyase | *hyl*B | Cleaves hyaluronan | Promotes spread through host tissues during infection<br>Impairment of leukocyte trafficking |
| C5a peptidase | *scp*B | Cleaves human complement C5a<br>Binds fibronectin | Inhibits polymorphonuclear leukocytes recruitment<br>Epithelial adherence and invasion by binding to extracellular matrix fibronectin |
| CAMP factor | *cfb* | Forms pores in host-cell membrane<br>Binds to glycosylphosphatidylinositol anchored protein<br>Binds to IgG, IgM | Direct tissue injury<br>Impairment of antibody function |
| C protein (α and β components) | *bca* (α)<br>*bac* (β) | Binds epithelial cells<br>Blocks intracellular killing by neutrophils<br>Non-immune binding of IgA | Epithelial cell adherence<br>Epithelial cell invasion<br>Resistance to phagocytic clearance |
| Alpha-like protein (Alp) family | *bca*, *eps*, *rib*, | Binds epithelial cells<br>Suffers antigenic variation as evasion mechanism of antibody detection | Epithelial cell adherence<br>Epithelial cell invasion |

| | | | |
|---|---|---|---|
| | *alp2*, *alp3* | | |
| Fibrinogen binding proteins A and B | *fbs*A<br>*fbs*B | Binds ECM fibrinogen through repetitive structure motifs | ECM attachment<br>Epithelial adherence<br>Promotes entry of GBS into host cell |
| Pili | PI-1PI-2a or PI-2b | Promotes resistance to antimicrobial peptides by an unknown mechanism | Promotes adherence of GBS to host cells |

### 1.9.1 GBS adhesion and invasion into host cells

To initiate and trigger GBS infection, attachment of the GBS to extracellular matrix (ECM) is the first step prior to binding to host-cell-surface proteins such as integrins [285]. GBS has the ability to adhere to and invade different cells including lung endothelial and epithelial cells, vaginal epithelial cells, gut epithelial cells and other human barrier endothelial cells [286] using various surface cell proteins namely laminin binding protein (Lmb), Streptococcal C5a peptidase group B (ScpB), and fibrinogen-binding protein (FbsA) that can bind to host ECM (**Figure 1.5**) [287-289]. Other surface proteins that promote GBS adherence to host cells are serine-rich repeat domain protein (Srr-1) [290], pilus island proteins (PI-1, PI-2a and PI-2b) [291] and hypervirulent protein encoded by *hvgA* [292]. In particular Lmb binds to human laminin and ScpB acts as an adhesin to bind to fibronectin [288, 293]. Both genes encoding Lmb and ScpB are present adjacent to each other on a composite transposon of 16 kb and the flanking Lmb-ScpB region carries insertion sequences that are responsible for the horizontal transfer of the element [294]. All human GBS strains harbour the Lmb-ScpB element [295, 296] and share 98% identity to the respective genes of *Streptococcus pyogenes* [294].

FbsA is a surface localized protein typically composed of 16 amino acid repeats and plays a specific role in binding fibrinogen [289]. Nearly all GBS strains are FbsA positive but can differ in the number and composition of repeat units it contains [289]. The capacity of GBS to bind fibrinogen depends on the number of repeats of FbsA: the greater the number of number of repeats, the greater will be the

affinity of fibrinogen bound by FbsA [289]. The binding of fibrinogen to FbsA contributes to GBS virulence by inhibiting complement factor (C3b) deposition on the bacterial surface, which in turn prevents opsonophagocytic killing by the immune system [102].

Serine-rich repeat (SRR) glycoproteins belong to the diverse and large family of adhesins characterised in most gram-positive bacteria [297]. GBS expresses either Srr-1 or Srr-2, two allelic SRR proteins [298], which have highly conserved domains and both can bind to the fibrinogen A α-chain through the "dock, lock, and latch" mechanism and these interactions contribute to the GBS colonization of vaginal surfaces and in the pathogenesis of GBS meningitis [275, 298].

Pili have been reported to perform different roles in i) adherence; ii) invasion and iii) epithelial cell translocation and iv) biofilm formation, however the specific function of pili is still unclear [291, 299]. Pilus like structures in GBS were originally discovered in 2005 [279]. Pili is encoded by genes which are present at two distinct loci, at different genomic positions, and are abbreviated as pilus island PI-1 and PI-2, (later divided into two more subtypes: PI-2a and PI-2b) (**Figure 1.6**) [300]. A pilus is composed of three structural subunit proteins: PilA (Gbs1478), PilB (Gbs1477), and PilC (Gbs1474), and its assembly involves two class C sortases (SrtC3 and SrtC4) required for the polymerization of pilus structures [299]. PilB (also known as the "bona fide pilin") is the major component; while PilA, (the pilus associated adhesin), and PilC are both accessory proteins incorporated into the pilus backbone [301]. PI-1 and PI-2b are observed at a higher frequency in neonatal GBS strains than in maternal colonizing strains, and observed exclusively in CC17 strains [302]. The PI-2a genes were found highly variable, compared to PI-1 and PI-2b, using comparative sequence analysis of pilus encoding genes [280]. This is consistent with adaptation or drift of an ancestral genotype to the different lifestyles and multiple genetic backgrounds where it is now found. Since the discovery of pilus structure in GBS, extensive studies discovering their role in pathogenesis have been undertaken [303]. GBS infection models have demonstrated that the ancillary protein of PI-1 and PI-2a provide an adhesive property of GBS to human tissues [299]. whereas the backbone protein of PI-1 and PI-2a provides GBS with intra- and para-cellular translocation across cells [304]. Further, PI-2a pili also play a specific role in GBS

biofilm formation [299] while the Spb1/SAN1518 protein of PI-2b is suggested to enhance intracellular survival of GBS in macrophages [305]. The pilus components are highly immunogenic making them an appealing candidate for vaccine development [280]. In 2005, pan-genome analysis of five GBS serotypes (Ia, Ib, II, III and V) strains identified 396 core genes including pilin proteins and antigens, that were shown to induce a serotype-independent immune response against GBS infection [97, 303]. Following GBS adherence and colonisation of host cells, it secretes toxins or utilizes cell-surface virulence factors, known as invasins, to promote bacterial entry and survival within host cells (**Figure 1.5**). In addition to FbsB, a surface-anchored GBS epithelial cell adhesin that binds fibrinogen via its N-terminal domain, Lmb and ScpB play an obvious role in supporting efficient endothelial or epithelial cell invasion.

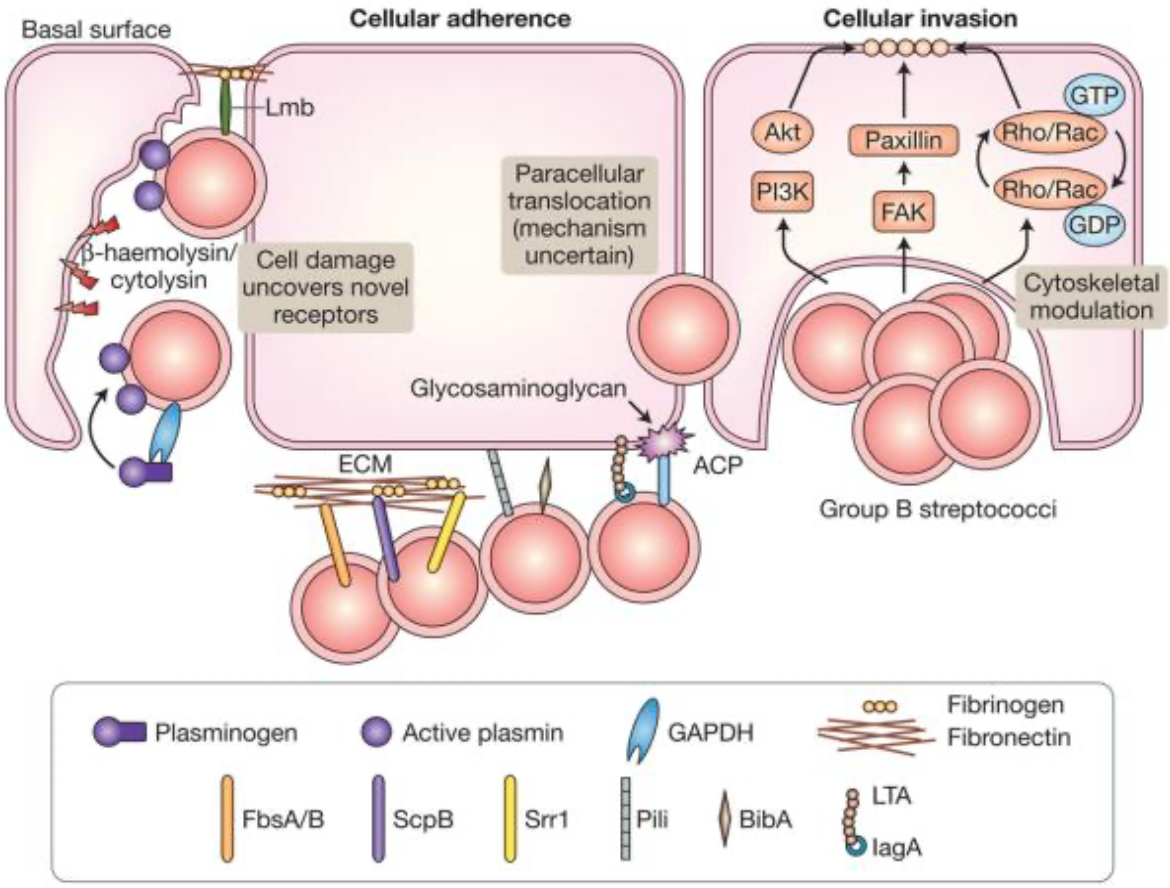



**Figure 1. 5. Different mechanisms used by GBS for the adherence and invasion into host cell**. Various cell surface proteins such as fibrinogen-binding proteins A and B (FbsA/B), C5a peptidase (ScpB), GBS immunogenic bacterial adhesin (BibA), serine-rich repeat domain protein 1 (Srr-1), pili, cell wall associated lipoteichoic acid (LTA) and surface-anchored AlphaC protein (ACP) promotes GBS binding to ECM components such as fibronectin and fibrinogen and to the host cells. Secreted

beta cytolysin/hemolysin promotes GBS to invade host cells by cleaving down host barriers to reveal basement membrane receptors such as laminin. Further to activate host plasminogen for the degradation of ECM, GBS uses GAPDH. The host PI3K/AKT- and FAK-signalling pathways and Rho family GTPases activate bacterial-dependent cytoskeletal rearrangement to enhance intracellular GBS invasion. Several other adhesins such has including FbsB, ScpB, pili, LTA and ACP, also contribute to cellular invasion. Adapted with permission from [285].

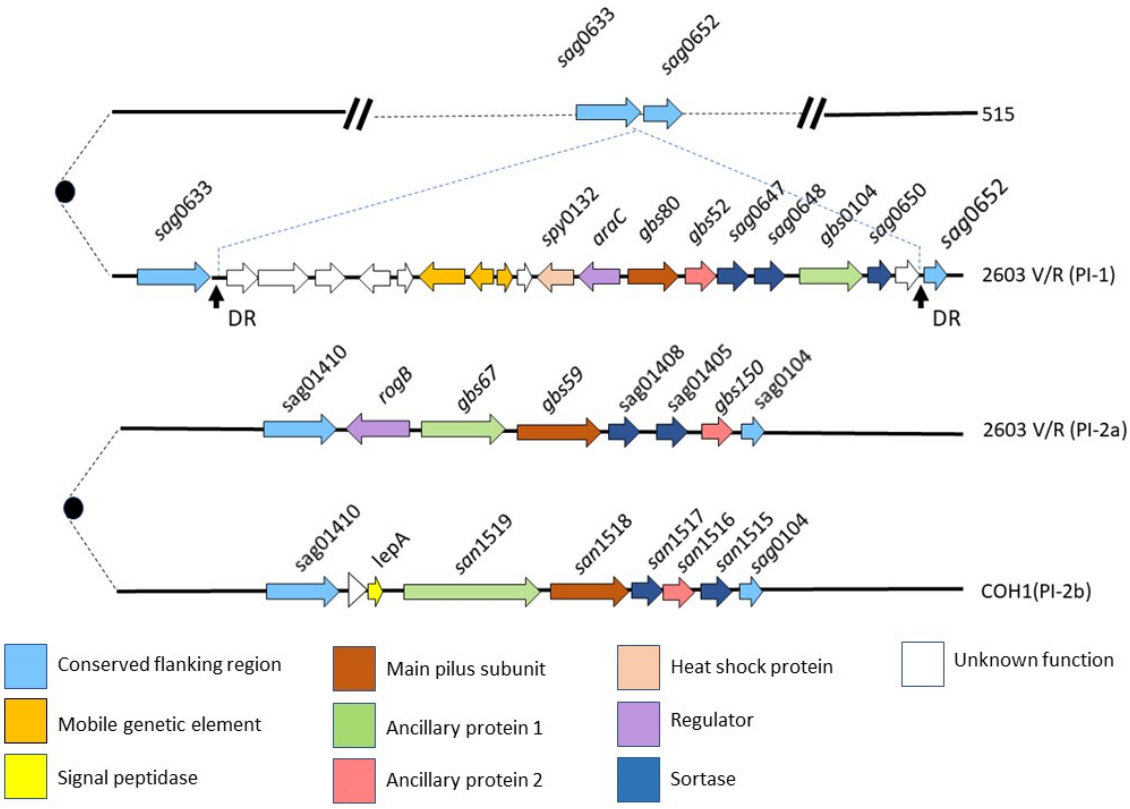



**Figure 1. 6. Schematic representation of GBS three pilus islands (PI-1, PI-2a and PI-2b):** Upper panel is representing PI-1 island in GBS reference 2603V/R strain while lower panel is representing PI-2 variants (a and b) in reference GBS COH1 strain. Reproduced from [306].

### 1.9.2 Proteins involved in immune evasion

Immune evasion is a strategy used by pathogens to continue its growth within host cell and maximise its ability to transfer into a fresh host [307]. Besides FbsA and ScpB, cell surface associated protein (CspA), GBS immunogenic bacterial adhesion (BibA), CPS are also involved in host immune evasion [308-310]. The CspA protein is closely related to ScpB protein and, as a serine protease, cleaves

fibrinogen into adhesive fibrin like fragments that coat the bacterial surface to interfere with immunological recognition [308]. GBS further expresses a cell surface adhesin known as BibA that binds to human C4-binding protein [310]. Its main function is to confer resistance by inhibiting complement activation and complement fragment deposition that would otherwise enhance neutrophil and phagocyte killing and is required for GBS survival in the blood [310]. In hypervirulent clones (almost exclusively restricted to sequence type ST17) a *bib*A locus variant was observed, known as the hypervirulent GBS adhesin (*hvg*A) [292]. Both consist of a variable core with 50 to 60% sequence identity, flanked by conserved 5' and 3' ends. Importantly, it was shown that *hvg*A was overexpressed *in vivo* and efficiently adheres to intestinal epithelial cells [292]. Further experiments showed that *hvg*A contributes to the bacterial translocation across the intestinal barrier and to GBS invasion of the blood-brain barrier (BBB) [292]. Altogether, these observations demonstrated that *hvg*A is a major virulence factor in GBS, associated with the onset of meningitis and LOD [292]. Higher prevalence of serotype III ST17 HvgA expressing protein, and antibiotic resistant GBS isolates have previously been reported in infants younger than 90 days and diagnosed with iGBS disease from China [311], Serbia [312], Canada [313], and France [269]. In addition to neonatal disease, ST17 *GBS* can cause disease in adults. For example, a recent study from Hungary [314] demonstrated an elevated rate of GBS ST17 hypervirulent clones causing invasive disease in non-pregnant adults (21.7%) than previously observed in this region. This is comparable to an earlier French study where 22% of adults with iGBS isolates were found to be serotype III ST17 [269].

### 1.9.3 Proteins involved in tissue adherence and cellular destruction

GBS has successfully gained fame as a human pathogen because of its ability to adhere to various tissues including respiratory epithelium, vaginal endothelium, and BBB endothelium [315]. GBS can be classified based on the ten polysaccharide capsules (Ia, Ib, II-IX) and set of surface proteins it expresses and certain capsule types have common association with other virulence genes [316]. Seven members of the alpha-like proteins (Alp) family, such as Cα (encoded by *bca*), Cβ (encoded by *bac*), Alp1 (*alp*1/epsilon), Alp2 (*alp*2), Alp3 (*alp*3), Alp4 (*alp*4) and R4/Rib (encoded by *rib*) have

been identified to date [317] which intercedes GBS binding to human cervical epithelial cells and interacts with host cell glycosaminoglycan (GAG) to promote bacterial internalization [318]. Over 90% of GBS clinical isolates were found to encode at least one of the Alp protein family genes [319].

The first surface protein gene identified in GBS was the C protein [320]. Subsequent characterization revealed two components i) the trypsin-resistant α protein and ii) the trypsin-sensitive β protein antigens of the C protein, both of which provoke protective immunity in animals [320]. A GBS strain can express a Cα or Cβ protein, or both [285]. Besides Cα protein and Cβ proteins, the major surface-localised proteins include the R-proteins (R1, R2 R3 and R4), the last of which (R4) has been found to be identical to the Rib protein [321] and the α-like proteins Alp2 and Alp3 may be also be variants of R1 protein [322, 323]. The epsilon protein has also been called α-like protein Alp1 [317].

Surface proteins of the Alp family expresses protective epitopes [324] and contain an N-terminal secretion signal sequence (S) and conserved domain (N); a variable number of tandemly arranged repeats of 70–80 amino acids (R); 8 to 10 repeats, and a C-terminal LPXTG cell-wall anchoring motif (**Figure 1.7**) [325]. High sequence similarity may be seen in certain domains of Alp proteins, which provides them structural basis for cross-protective immunity and for their interactions with the same host receptor.

Among different members of the Alp proteins, the most extensively studied are the Cα and Cβ proteins [326]. The Cα protein plays a critical role in invading human cervical epithelial cells through its interaction with glycosaminoglycan (GAG) [326] whereas the MLKKIE sequence motif of Cβ protein binds to the Fc region of human IgA or human factor H (FH) to protect GBS from opsonophagocytosis [327]. Several biochemical features are common between Cα and Rib proteins, but no immunological cross-reaction between Cα or Cβ proteins has been found [328, 329]. The potential invasive properties of other Alp proteins have not been studied [329].

Specific association is seen between surface protein genes and particular serotypes e.g. the *alp1* is present in the majority of serotype Ia GBS strains; *bca* and *bac* with serotypes Ib and II; *rib* with serotype III and *alp3* with serotype V [330]. Studies of surface proteins, and of the genes encoding these proteins, are important for epidemiological analysis of GBS infections [330]. Due to their critical role at different stages of infection (adherence; invasion to host cells and immune system evasion), they are also important as potential candidates for vaccine development [325]. Preclinical vaccine investigations of the Cα, Alp3, and Rib proteins have been conducted, but the use of Alp proteins as universal vaccines has been limited due to the heterogeneity of the Alp sequence [270, 317].

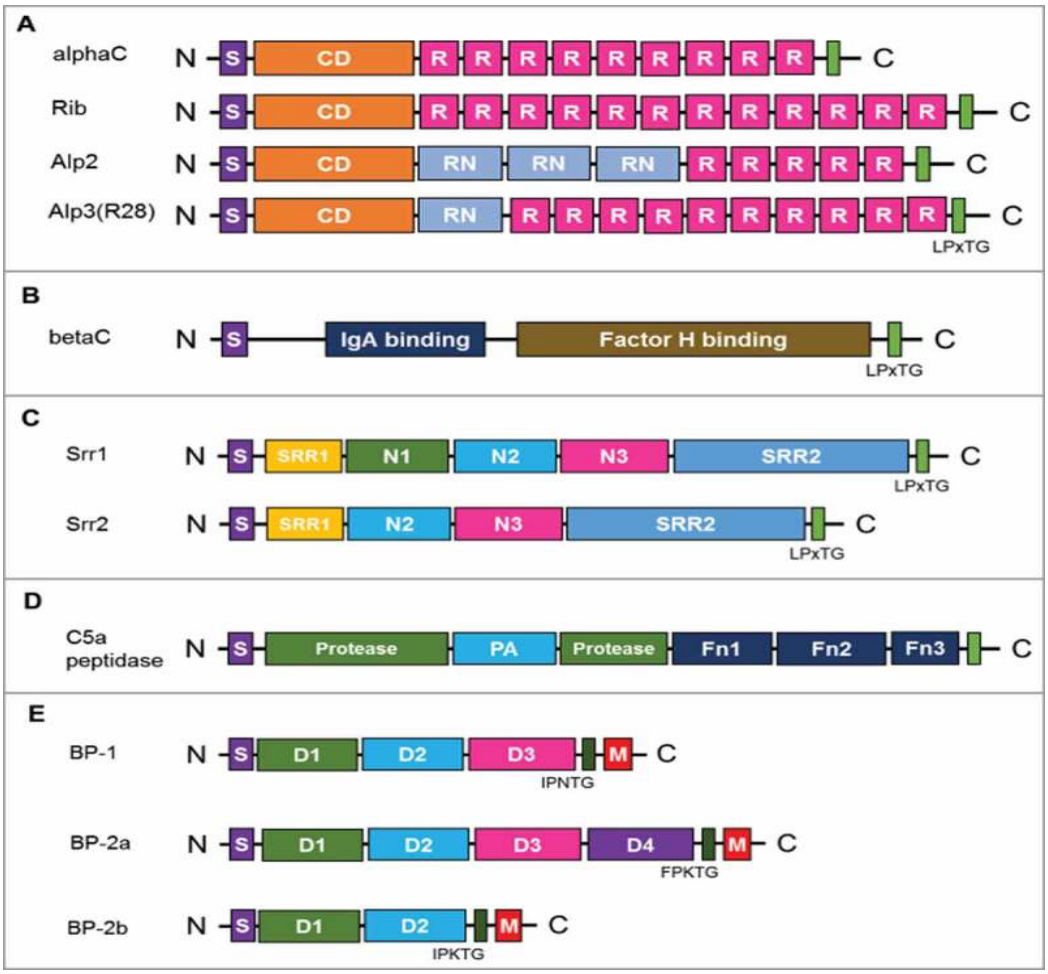



**Figure 1. 7. Schematic representing surface proteins expressed by GBS.** (A) showing the Alp protein family where (S) represent Signal peptide, (CD) - conserved domain, (R) - repeating domain, green block - LPxTG cell-wall anchoring domain. (B) represents Cβ protein, with the blue block showing IgA binding domain and brown showing the factor H binding domain. (C) shows the Serine-rich repeat protein family containing (N1, N2, and N3) - Ig-like domains and (SRR-1 and SRR-2) - serine-rich repeat glycosylation domains. (D) shows C5a peptidase containing - Protease activity domains, (PA) – a protease-associated domain, and (Fn1, Fn2, Fn3) - fibronectin type III domains. (E)

shows Pilin proteins containing (D1-D3) - Ig-like domains, (IPNTG, FPKTG, IPKTG) - pilin sorting motifs, and (M) - a membrane-spanning hydrophobic domains. Adapted with permission from [329].

## 1.10 CAPSULAR POLYSACHARIDE (CPS)

The majority of GBS isolates infecting humans are encapsulated, which play a critical role in pathogenicity of GBS by conferring resistance to phagocytosis and opsonophagocytic killing and also playing a role in inhibiting clearance by the complement system [102]. GBS CPS are predominantly composed of repeating units containing four elements i) glucose ii) galactose iii) N-acetylglucosamine and iv) sialic acid moieties on the side chain of all serotypes [331]. Serotypes VI and VIII are an exception to this generalisation of sugar composition by lacking the N-acetylglucosamine, while serotype VIII also uses ramnose residues in addition the other four moieties outlined above [332]. The polysaccharide capsule is encoded in a gene cluster consisting of a central variable group of genes that encode specific glycosyltransferases and polymerases that separate the CPS composition of different serotypes, which are flanked by glycosyltransferase and polymerase genes conserved across all 10 of the serotypes [333]. The conserved genes upstream of this "serotype-specifying region" encode enzymes that synthesize and activate sialic acid, while the genes downstream are hypothesized to function in export of the polysaccharide capsule [332]. There are ten distinct serotypes of GBS CPS based on structure, chemical composition that are recognised through distinct serological typing reagents (named Ia, Ib and II-IX) [332]; these are directly related to the different gene composition of the *cps* ("serotype-specifying region") gene clusters (**Figure 1.8**) that are both structurally and genetically closely related.

The role of the CPS as a virulence element was first studied in 1987 using a rat model of neonatal GBS infection which showed that a non-encapsulated GBS mutant had significantly reduced virulence compared to the encapsulated GBS parent strain [334]. Loss of capsular sialic acid was also associated with loss of virulence in a mutant strain in a neonatal rat model of lethal GBS infection [245]. However, while the amino acid sequences of carbohydrate modifying enzymes in different serotypes perform similar functions, their varying carbohydrate specificity mediates significant heterogeneity [90, 333].

Evidence of inter- and intra-species homologous recombination events leading to horizontal transfer of capsular genes has been examined in detail using capsular biosynthesis clusters, contributing to GBS serotypes diversity [335]. Adaptive immune responses in the host population may drive changes at the capsular locus, leading to serotype variation and preservation of specific serotype which may be necessary for virulent clone [333, 335].

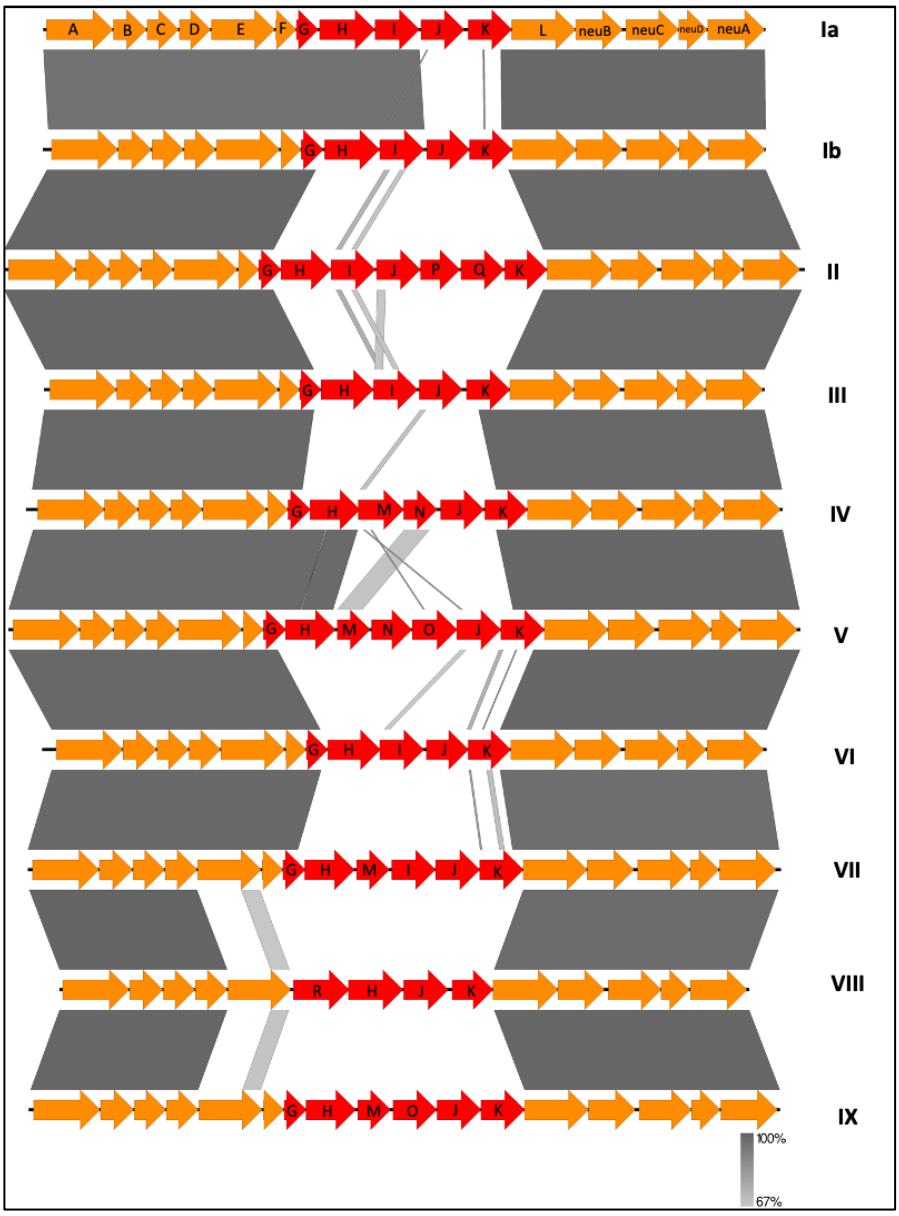



**Figure 1. 8. Comparison of capsule genes observed in ten GBS serotypes Ia, Ib, II-IX**. EasyFig v 2.2.2 [336] was used to determine nucleotide identity between capsular genes. Annotation was performed using Artemis [337]. Red colour arrows represent the variable region *cps*G-K that have significant diversity for each serotype, whereas orange arrows represent conserved region *cps*A-E,L and *neu*A-D. Reproduced from [45].

## 1.11 Global epidemiology of GBS strains

### 1.11.1 Maternal colonization rate of GBS

Colonizing GBS strains may possibly act as reservoirs of serotype diversity, antibiotic resistance, virulence factors and are also the major source of disease-causing isolates [338]. Therefore, for comprehensive understanding of GBS disease, categorizing the colonizing GBS population structure is important [338]. According to two recent reviews on maternal colonization rate of GBS and its serotype distribution worldwide [122, 339], estimated maternal GBS-colonization was 18%, worldwide ranging from (10-30)% in U.S, (6.5-36)% in Europe, (7.1-16)% in Asia, (9.1-25.3)% in the Middle East, and (11.9-31.6)**%** in Africa, and (20-24)% in Australia as reported in different studies [21, 153, 340-343]. These maternal GBS colonization prevalence rates were supported by a recent systematic review [147] estimating maternal GBS colonization prevalence for countries, subregions (South America, Central America, Caribbean, Western Asia, Southern Asia, South-Eastern Asia, Eastern Asia, Oceania) and regions (Latin America, Asia, Africa, developed Oceania) to apply to estimates of live births in 195 countries for 2015, using latest United Nations data. Analysis showed that the highest prevalence was observed in the Caribbean (34%) and lowest in Melanesia (2%). Similar prevalence rates were noted in Europe, North America, and Australia (15%–21%), with a slightly higher prevalence in Southern Africa (25%), and relatively lower prevalence in Western Africa (14%), Central America (10%), and South, South-Eastern, and Eastern Asia (9%–12%) [147]. The geographical distribution of the most prevalent colonizing serotypes and rates identified [147] are shown in (**Table 1.6) and** (**Figure 1.9)**, respectively.

**Table 1. 6. Maternal GBS colonization prevalence results from meta-analyses with reported data and meta-analyses with adjusted data.** Adapted with permission from [147].

| Region/ Subregions | No. of Countries | No. of pregnant women tested | Reported prevalence, % | 95% Confidence Interval | Prevalence from studies with recommended | 95% Confidence Interval | Adjusted prevalence[b], % | 95% Confidence Interval |
|---|---|---|---|---|---|---|---|---|
| Developed regions | 29 | 144604 | 18.4 | 17.0–19.8 | 21 | 19.6–22.3 | 19.2 | 17.7–20.7 |
| Australia and New Zealand | 2 | 2369 | 23.3 | 18.8–27.8 | 23.3 | 18.8–27.8 | 23.3 | 18.8–27.8 |
| North America | 2 | 27462 | 22 | 19.2–24.8 | 23 | 20.9–25.1 | 23.2 | 21.1–25.3 |
| Northern Europe | 7 | 6702 | 20.6 | 16.6–24.7 | 24.1 | 21.9–26.4 | 22.2 | 19.1–25.4 |
| Eastern Europe | 7 | 15737 | 20.8 | 17.3–24.4 | 22.9 | 18.7–27.2 | 23 | 19.2–26.8 |
| Southern Europe | 5 | 42870 | 15.4 | 12.2–18.7 | 16.7 | 14.7–18.6 | 17.6 | 14.5–20.8 |
| Western Europe | 6 | 49464 | 15.2 | 13.1–17.3 | 18.3 | 16.0–20.7 | 19.5 | 13.9–25.1 |
| Americas | 13 | 20507 | 18.3 | 15.8–20.7 | 19.6 | 16.7–22.5 | 20.9 | 18.1–23.7 |
| South America | 8 | 16141 | 15.9 | 13.5–18.2 | 15.7 | 13.0–18.5 | 18.4 | 15.5–21.3 |
| Central America | 2 | 3229 | 10.2 | 6.7–13.8 | 15.7 | 13.3–18.0 | 17.1 | 13.2–21.0 |
| Caribbean | 3 | 1137 | 33.5 | 28.8–38.3 | 33.5 | 28.8–38.3 | 34.7 | 29.5–39.9 |
| Asia | 20 | 98842 | 11 | 10.0–12.0 | 11.6 | 10.5–12.7 | 12.8 | 11.8–13.9 |
| Western Asia | 7 | 15124 | 14.3 | 11.-16.6 | 14.5 | 11.7–17.4 | 14.7 | 12.1–17.4 |
| Central Asia | 0 | … | … | … | … | … | … | … |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Southern Asia | 4 | 15838 | 10 | 8.3–11.6 | 10 | 7.5–12.6 | 12.5 | 10.2–14.8 |
| South-Eastern Asia | 6 | 4591 | 12 | 9.3–14.7 | 14.4 | 9.5–19.2 | 14.4 | 11.5–17.4 |
| Eastern Asia | 3 | 63289 | 9.2 | 7.6–10.8 | 9.1 | 8.2–10.0 | 11.1 | 9.9–12.4 |
| Africa | 19 | 36130 | 18.2 | 16.1–20.4 | 20.7 | 17.6–23.7 | 21.3 | 18.5–24.2 |
| Northern Africa | 3 | 1923 | 20 | 15.8–24.3 | 20.5 | 15.5–25.4 | 22.9 | 17.9–28.0 |
| Western Africa | 6 | 4860 | 13.6 | 9.0–18.3 | 17.2 | 6.2–28.3 | 17.5 | 10.8–24.1 |
| Middle Africa | 3 | 2058 | 18.6 | 16.9–20.3 | 19.3 | 15.9–22.7 | 23.9 | 14.7–33.1 |
| Eastern Africa | 6 | 14071 | 18.2 | 15.0–21.5 | 19.4 | 15.5–23.3 | 19.4 | 15.9–23.0 |
| Southern Africa | 1 | 13218 | 25.3 | 22.1–28.5 | 29.5 | 27.4–31.5 | 28.9 | 26.6–31.2 |
| Oceania | 1 | 440 | 19 | 6.8–31.3 | … | … | … | … |
| Melanesia | 1 | 440 | 2 | 0.6–3.5 | … | … | … | … |
| Overall | | 300176 | 15.2 | 14.3–16.0 | 17.4 | 16.3–18.5 | 18 | 16.9–19.1 |

[a]Recommended methods refers to studies including both rectal (or perianal) and vaginal swabs, and with selective enrichment or a selective agar proven to provide equivalent sensitivity.

[b]Adjusted prevalence for sample site and microbiological methods.

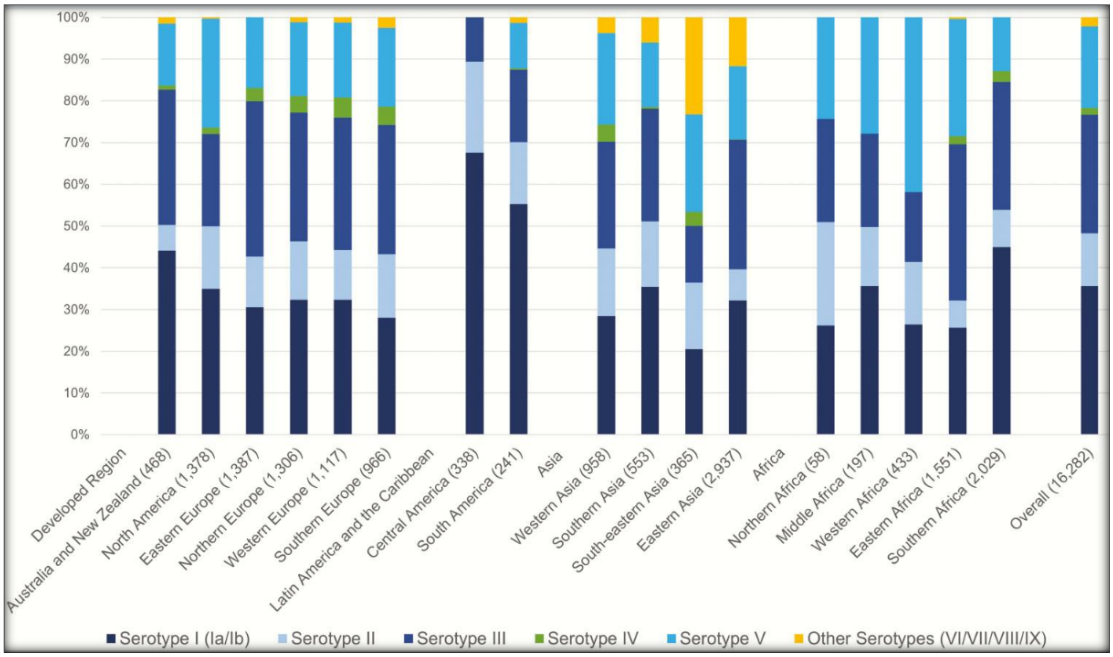



**Figure 1. 9. Maternal GBS colonizing serotype distribution by United Nations subregion.** Adapted with permission from [147].

### 1.11.2 Infant colonization rate of GBS

Understanding the global burden of GBS disease in young infants (0–89 days) is important to guide public health decision making on interventions [263]. According to the 2017 systematic review and meta-analysis on infant GBS incidence and serotypes worldwide [263], overall incidence risk for infant GBS disease has been recorded as 0.49 per 1000 live births, being highest (1.12) in Africa and the lowest (0.30) in Asia. Incidence risk of GBS-EOD per 1000 live births worldwide was 0.41, with the highest (1.47) in the Caribbean and lowest (0.20) in South Asia [263]. Incidence risk of GBS-LOD worldwide was 0.26, with the highest (0.93) in Southern Africa and lowest (0.0) in South America, Western Africa, and South-eastern Asia, based on single study captured from each of these regions (**Table 1.7**) [263].

**Table 1. 7. Incidence rate of iGBS disease in neonates per 1000 live births worldwide.** Reproduced from [263].

| Regions | overall | GBS-EOD | GBS-LOD |
|---|---|---|---|
| Asia | 0.3 | 0.32 | NA |
| Africa | 1.12 | 0.71 | NA |
| Latin America and Caribbean | 0.49 | 1.47 | NA |
| South Africa | NA | 1.07 | 0.93 |
| South Asia | NA | 0.2 | NA |
| Developed countries | 0.46 | NA | NA |
| South America | NA | NA | 0* |
| West Africa | NA | NA | 0* |
| South-eastern Asia | NA | NA | 0* |

*Rates are based on single study captured from each of these regions.

### 1.11.3 Serotype distribution of GBS in different populations

Whilst GBS colonization rates may look relatively analogous in various areas of the world, serotype distribution and prevalence are geologically distinctive [265, 339]. Serotypes Ia to V are frequent colonizers worldwide (**Figure 1.10**) [265, 339] in particular more prevalent in Australia and New Zealand, America, Europe, Africa, and Asia [147, 339]. Whereas, serotypes VI to VIII are found more frequently in south-eastern and eastern Asia based on serotype distribution as derived from one Bangladeshi and six Japanese studies included in the Madrid *et al* review (**Figure 1.9**) [263].

Another more recent review by Jassir *et al* of global distribution of GBS serotypes in young infants [265] described serotype III as the dominant invasive clone in developed countries, Central America, Eastern Africa, Eastern Asia and Southern Africa, but serotype III as less prevalent in South America where serotype Ia was the most common [263]. While serotype Ia and V are dominant invasive isolates in non-pregnant adults of North America [114, 152], the prevalence of invasive serotype IV strains in adults and neonates is increasing with time, despite being rarely isolated in North America prior to 2000 [68, 114, 344]. In the United States, longitudinal surveillance suggests that the proportion of serotype

IV GBS isolates from nonpregnant adults increased from 0.2% in 1998 to 1999 to 5.7% in 2005 to 2006 [132]. In 2010, serotype IV strains were shown to be responsible for 16% of early onset neonatal infections in the state of Minnesota [68, 344], and in 2014 were responsible for 6.2% of cases of iGBS disease in Canada [313].

An analogous situation could be estimated for serotype VI, which was rarely reported in Europe and America at 1% in infants younger than 90 days prior to 2012 [345] and 0.1% in nonpregnant adults prior to 2007 [132], yet serotype VI was the predominant serotype in Japanese pregnant women with GBS colonization (5.6–24.7%) over roughly the same time period [346-348] and also appeared as the most common serotype in Malaysian pregnant women [349-353]. In Taiwan, the past 20 years have seen serotype VI reported sporadically: 2.6% in patients with invasive disease during 1994–2004 [354], and 4% between 2001 and 2004 [355]. Still, prevalence rose to 10.5% in nonpregnant patients between 2006 and 2008 [356] and 12% in cases of invasive disease from 1998 to 2009 [357]. More recently, a single hospital study of Central Taiwan, [358] revealed a clonal dissemination of serotype VI among colonizing and invasive isolates. Similarly, several studies have also identified sporadic strains of serotype VIII both as colonizing strains and as causative agents for iGBS disease [114, 341, 346, 351, 359, 360].

### 1.11.4 CC and ST prevalence of GBS

Five main genetically unique lineages, clonal clades CC1, CC10, CC17, CC19 and CC23, are commonly reported in association with both invasive disease and asymptotic colonization in people of different age groups based on various GBS population epidemiology-based studies [1, 200, 361]. The predominance of these groups suggest they are well adapted to specific niches, regardless of population immune selective pressure or antibiotic resistance pressure [1, 362]. These five major lineages have been identified as colonisers of both humans and cattle [363, 364] [365, 366]. While two main GBS groups associated with fish-associated infections are CC7 and ST260/261 [367], CC7 has also been detected in humans and cattle [367], but ST260/261 is unique to fish as primary reservoir [368] (**Figure 1.10**).

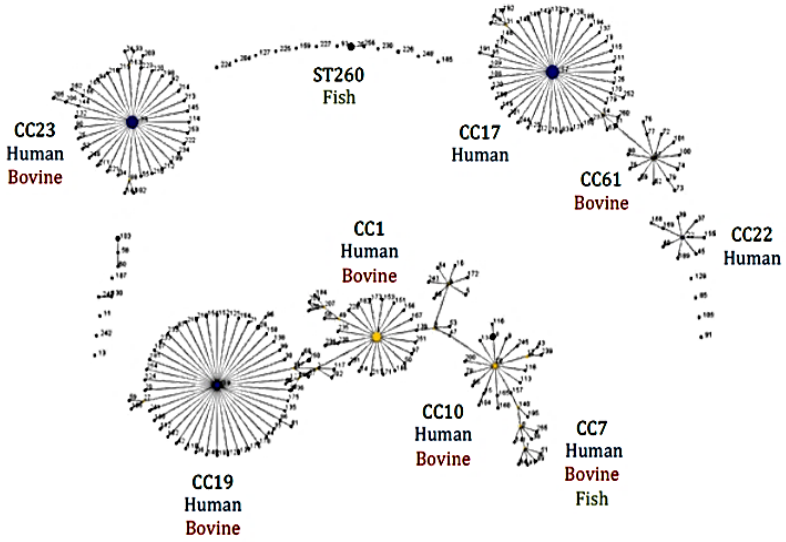



**Figure 1. 10. MLST defined GBS clonal complexes and their host origin.** Adapted with permission from [369].

In a most recent review, Jassir *et al* characterised global GBS maternal and infant colonization by MLST [265]. Seventy-eight studies were included with 45% (35/78) from Europe, 29% (23/78) from Eastern Asia, 9% (7/78) sub-Saharan Africa, 9% (7/78) Northern America and 3% (2/78) from Western Asia [265]. Additionally single studies from Latin America and the Caribbean, Northern Africa, Australia, and Southern Asia were included [265]. CC19 was the most common clonal complex reported for maternal colonisation at 22% overall prevalence, followed by 19% for CC23, 17% for CC1, and 15% for CC17 [265]. Within these clonal complexes the most common serotype was III representing 98% of CC17 strains and 70% of CC19 strains, while serotype Ia predominated in CC23 with 65% and serotype V in CC1 with 52% [265]. Similar to maternal colonisation, CC19 was the most frequent clonal complex identified for neonatal/infant colonisation with 39%, followed by CC23 (23%), but CC17 was much less common (5%) for maternal colonisation [265]. Sixty percent of CC19 strains expressed serotype III capsules and 52% of CC23 strains expressed serotype Ia [265].

For adult GBS invasive infections, CC17 strains are less frequently identified. In nonpregnant adults above 15 years of age, serotype V was the most prevalent serotype globally and in North America accounting for 43.48% (n=12926) and 46,72% (n=12184) of cases, respectively [370]. Serotype Ia was the second and serotype III was more prevalent in Europe (25.0%) and Asia (29.5%) [370]. Molecular epidemiology showed a notable correlation between ST1 and serotype V GBS causing invasive disease [40]. Serotype V GBS was first isolated from humans in 1975, and it was identified as a primary invasive serotype in the early 1990s [76]. Serotype V GBS strains (210/229; 92%) were almost exclusively isolated from the bloodstream of nonpregnant adults in the United States and Canada between 1992 and 2013 [76]. Analysis of the complete genome of a 1992 serotype V ST-1 strain revealed that this strain had the closest homology to a GBS strain causing cow mastitis [76]; however serotype V strains isolated in the late 1970s appear to have diversified significantly by acquisition of different cell surface proteins and antimicrobial resistance determinants that likely relate to serotype V ST1 strains adapting to become a more successful human pathogen [76].

## 1.12 Trends of GBS disease in neonates and adults in UK; 1999-2019.

Voluntary surveillance of GBS bacteraemia in England, Wales and Northern Ireland has been conducted by Public health England (PHE) for the last twenty-years (1999 to 2019) and is available online [8]. These reports consist of investigations on the trends, age and sex distribution, geographical distribution, and antimicrobial susceptibility of laboratory-reported cases of pyogenic and non-pyogenic streptococcal bacteraemia to study the regional and temporal incidence rates and antibiotic resistance profiles of these organisms circulating in England, Wales and Northern Ireland [8].

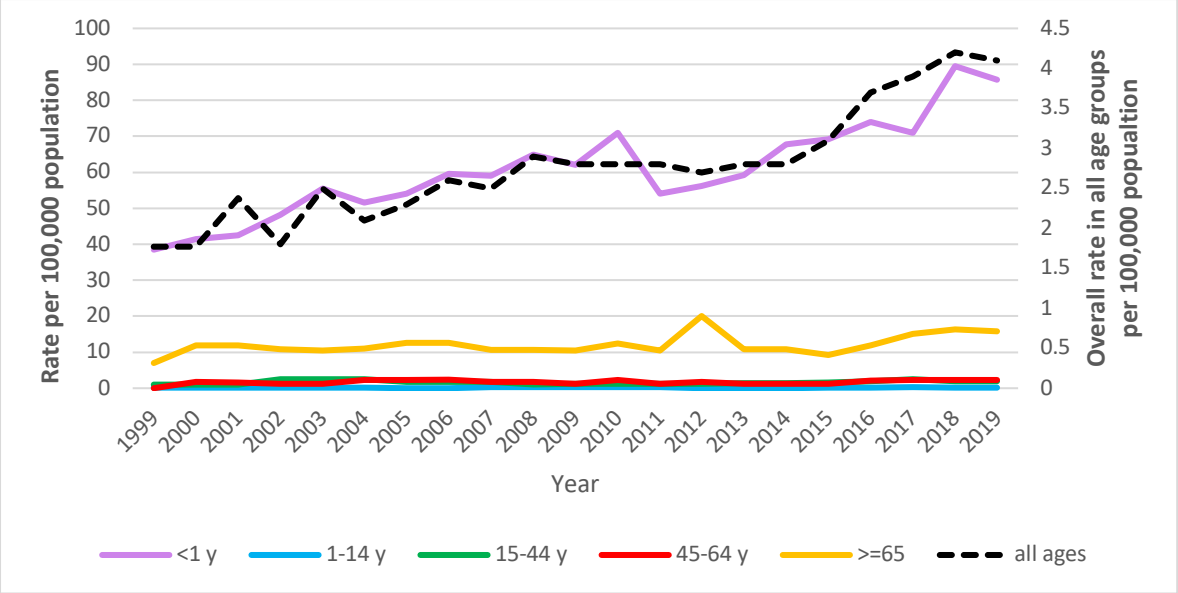



**Figure 1. 11. Age-specific rates of iGBS infection, England, Wales and Northern Ireland: 1999–2019.** Annual rates of GBS bacteraemia in different age groups per 100,000 population were extracted from PHE surveillance reports on GBS disease in England, Wales and Northern Ireland for 1999 to 2019 **[8].** The annual incidence rate of GBS in UK population of all age groups is plotted on secondary axis for better visualization.

The annual rate of GBS bacteraemia cases in all age groups in England, Wales and Northern Ireland doubled over the 20 years from 1.77 recorded in 1999 to 4.1 in 2019 per 100,000 population (**Figure 1.11**). The vast majority of GBS blood infections were concentrated in the under 1 year age group with infection rates increased enormously from 38.5 per 100,000 population in 1999 to 85.8 per 100,000 population in 2019 almost 20-fold compared to infection rates observed in all other age groups combined (**Figure 1.11**). The next highest reporting rates were in those aged 65 years and above, where infection rates recorded 1 per 100,000 population in 1991 and 9.25 per 100,000 population in 2019 on average (**Figure 1.11**). Overall, there was no significant change in GBS bacteraemia rates over time in people aged between 1 and 64 years (**Figure 1.11**). For the analysed period, rates were higher for males than females across most age groups, the exception being in the 15-44 age group where more females were found infected with GBS (**Figure 1.12**).

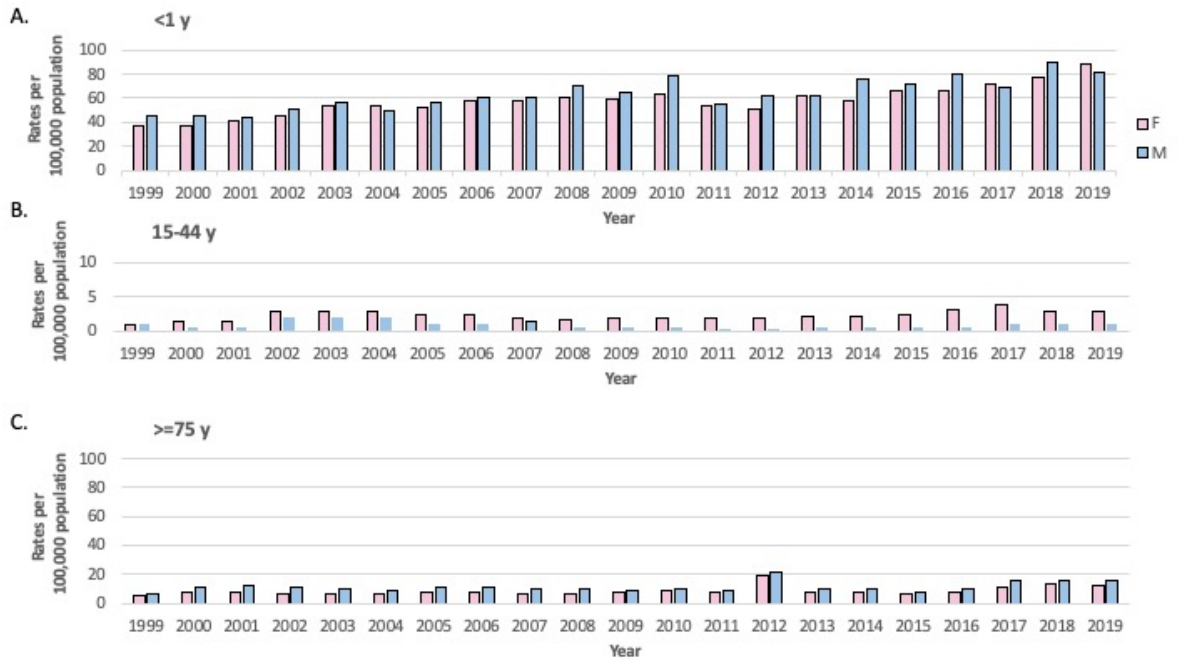



**Figure 1. 12. GBS bacteraemia age and sex rates per 100,000 population England, Wales and Northern Ireland; 1999-2019.** Reproduced using data extracted from PHE voluntary surveillance database for the given period **[8].** Rate of iGBS disease per 100,000 population is shown in three different study populations in A) infants aged < 1 year B) adults aged between 15 to 44 years and C) elderly above 75 years of age.

A general increase is seen both in GBS-EOD and GBS-LOD throughout the period between 1999 – 2019 (**Figure 1.13**). It even appears that there may have been a more rapid increase in GBS-EOD cases over the 2 decades with cases of 0.31 per 1000 live births recorded in 1999 to 0.55 per 1000 live births in 2019 (**Figure 1.13).** The rate of GBS-LOD as well, doubled over the 2 decades with 0.15 per 1000 live births cases reported in 1999 and 0.30 in 2019 (**Figure 1.13).** After 2012, cases of GBS-LOD slightly decreased but again showed a general increase in trend after 2017 (**Figure 1.13**).

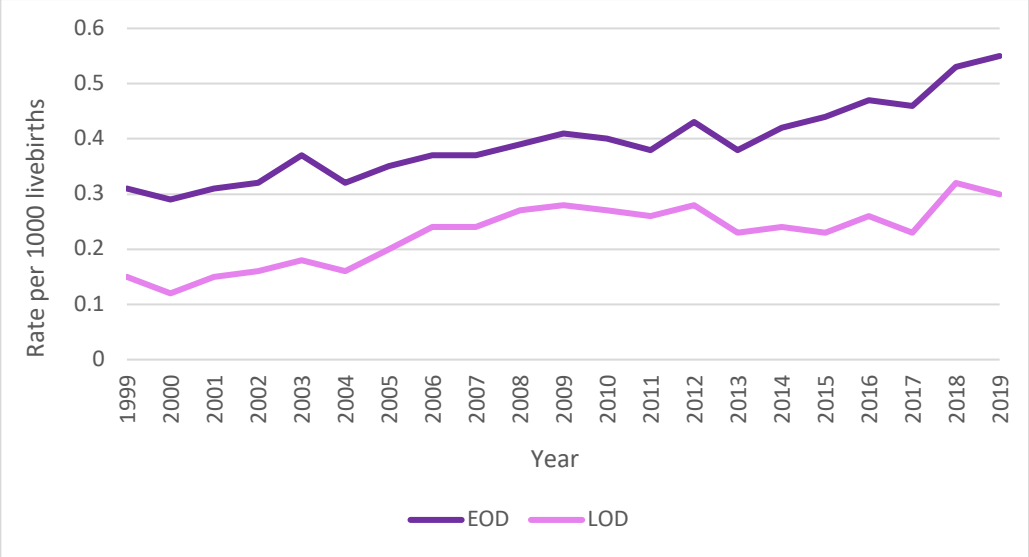



**Figure 1. 13. Rates of GBS-EOD and GBS-LOD infection in England, Wales and Northern Ireland; 1999–2019.** Rate of GBS infection (EOD and LOD) in neonates less than 3 months old were extracted from the PHE annual surveillance report of GBS bacteraemia **[8]** for the year 1999 – 2019 to plot the trends.

No information about GBS serotype distribution was provided in any of the published PHE annual surveillance reports for GBS causing bacteraemia in UK population between 1999-2019. In 2013, Lamagni *et al* [115] published the 20 year trends of GBS disease in England and Wales between 1991 and 2010. Overall, no significant change with age was observed in GBS serotype distribution for this report (**Figure 1.14A-C**). Among all age groups, serotype III and Ia were the foremost disease contributors, however, for infants (both LOD and EOD) serotype III was by far the most prevalent causing 70% of LOD and 37% of EOD (**Figure 1.14 A and B**). A slight change in serotype distribution is observed for each age group with stable differences between the groups through the 20-year epidemiological study (**Figure 1.14 A - C**).

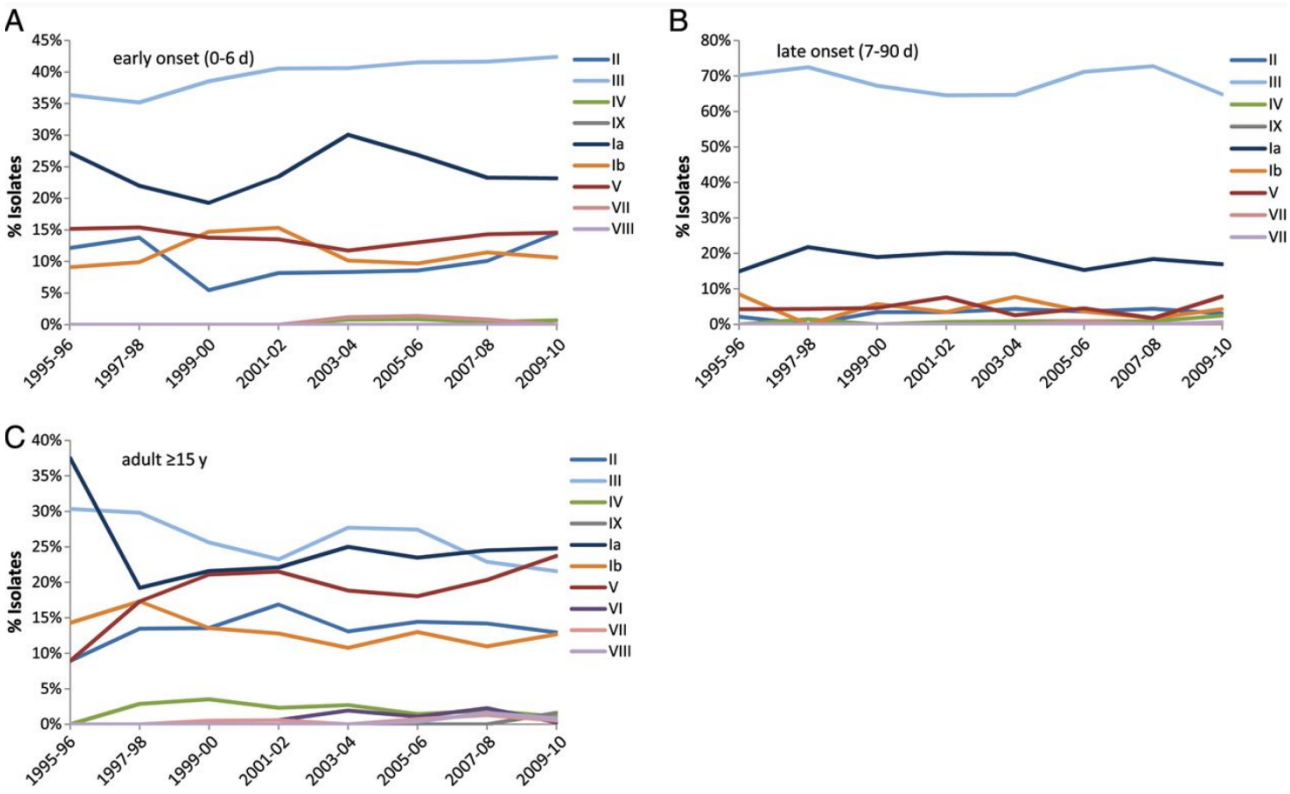



**Figure 1. 14. Distribution of GBS serotypes from sterile site isolates according to patient age, England, and Wales; 1995–2010.** A) representing serotypes distribution in EOD, B) representing serotypes distribution in LOD, and C) representing serotypes distribution in adult disease. Adapted with permission from [115].

PHE surveillance report for the years 2011 to 2018 were accompanied with the conventional AST results except for the year 2019, for which the AST results were not shared in public. Between 2011 and 2018, resistance to erythromycin and clindamycin were found steadily increasing with an average increase of 3% increase per annum (**Figure 1.15**). Analogous resistance rates were observed for erythromycin and clindamycin in 2011 (18%) which nearly doubled over the 8 years surveillance period for erythromycin 2018 (33%) and clindamycin (29%) in 2018, respectively (**Figure 1.15**). Increased tetracycline resistance rates were also observed for the period 2011 -2018, with an average of 83 to 86% GBS strains found resistant to tetracycline, per annum (**Figure 1.15**).

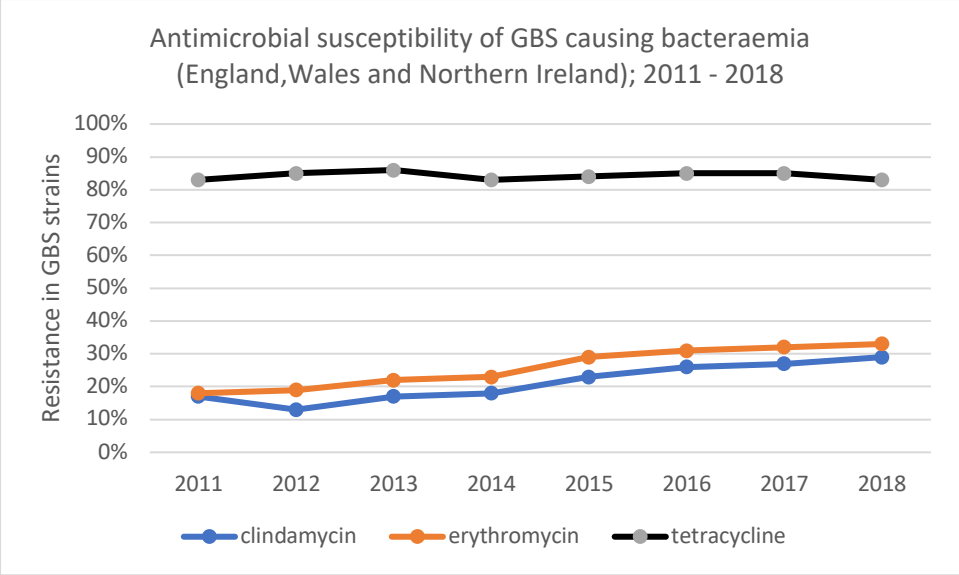



**Figure 1. 15. Resistance rate of clindamycin, erythromycin, and tetracycline against GBS causing bacteraemia, England, Wales and Northern Ireland; 2011 to 2018.** AST data for 2019 was not available. Reproduced using data extracted from PHE voluntary surveillance database for the given period **[8].**

## 1.13 Prevalence of GBS in Brazil

A systematic review by Madrid *et al* [263] included 15 studies from South America, of which 9 studies described incidence rate of GBS disease in infants in South America, 8 described case fatality rate while 4 studies described the serotype distribution in the same population. The incidence rate of GBS-LOD was lowest in South America, Western Africa, and South Eastern Asia (0.0, 0.0, and 0.03, respectively, based on the single study captured from each of these regions) compared to Southern Africa which had the highest incidence risk of GBS-LOD (0.93) [263]. Serotype III was the most prevalent serotype across the United Nations subregions, although it was lower in South America (34%) [263]. Unlike CDC and PHE, unfortunately there is no similar comprehensive and coordinated surveillance for GBS in Brazil; only sporadic and geographically limited papers are available on prevalence of GBS disease [371]. Most of the available data from GBS screening in Brazil in the last 10 years is from studies conducted on pregnant women and in the South East region, mainly in the states of São Paulo and Rio de Janeiro [48]. A lot of fluctuation is observed in the maternal GBS colonization rates from different geographic regions of Brazil, from 4.2% to 28.4%, based on studies published between 2008 and 2018

[372] compared to 23% based on studies published between 2007 to 2010 [373, 374]. This may be related to the absence of an official recommendation from the Federal Brazilian Health Authorities describing guidelines and protocols to perform GBS screening in pregnant women, in both public and private clinics. Divergence is also observed in the serotype distribution among pregnant women from different locations in Brazil [371, 374-376]. Few studies evaluated GBS serotypes, and from 2008 to 2017, most of them were performed in the South and South East region reporting highest prevalence of serotypes Ia, Ib, II, III, IV and V in pregnant women [374, 376-380]. In 2002, serotype III was the most abundant in pregnant women, followed by serotypes V and Ia [375]. In 2018, serotype Ia was reported as the most prevalent, followed by serotypes II and Ib [371], despite both of these studies being conducted in Rio de Janeiro. In São Paulo, serotype Ia was the most abundant followed by V and II in a study conducted in 2015–2016 [51]. In Paraná, serotype Ia was also the most prevalent; however, there was a surprisingly high number of serotype IV isolates [376]. For adults, including non-pregnant women, serotype Ia has been the most frequently reported in recent years [48, 379]. Examining global data for serotype prevalence, serotypes Ia, III, and V appear more common in pregnant women, serotype V and Ia in older adults, and most predominantly serotype III in infants with iGBS disease, combining data for all continents published between 1997 and 2019 [263, 265].

## 1.14 Prevalence of GBS in Australia

Despite widespread use of intrapartum antibiotic prophylaxis, GBS remains a leading cause of morbidity and mortality in infants in Europe, the Americas, and Australia [350]. There is currently no mandatory reporting for GBS infection or its complications in Australia, and recently published reports on incidences of neonatal and adult GBS disease are very limited or available from studies now over 10 years old and have relied on passively reported data from clinicians to the Australian Paediatric Surveillance Unit. As a result, the true incidence of GBS infection in Australia is largely unknown.

In 2012, a systematic review only identified three studies investigating the incidence of neonatal GBS infection in Australia [350], one of them was also included in a 2017 systematic review [263]. No

significant change was detected in the incidence rate of GBS-EOD disease in Australian infants over the last two decades. Between 2002 and 2012, the estimated incidence rates of GBS-EOD disease per 1000 live births in Australia was 0.39 and 0.43, respectively [381]. The most recent prospective surveillance study by the Australian Paediatric Surveillance Unit (APSU), published in 2015, identified GBS disease in infants aged 0–90 days between July 2005 and June 2008 at 12 major public hospitals across Australia [382]. Despite a reporting rate of one-third of that of known GBS cases to the APSU in this time, indicating the limitations of relying on passively reported data and not laboratory confirmed cases, the authors estimated rates of GBS-EOD to be 0.38 and GBS-LOD 0.19 per 1000 live births [382]. The maternal colonization rate was found to be 23.3 based on 6 Australian studies (tested 2639 pregnant women) included in a latest review which is comparable to the rate observed in Europe and North America [339].

GBS serotype distribution in Australian population belonging to different age groups seems stable in last three decades. Studies conducted after 1995 till 2020 described serotype III and Ia high association in causing neonatal invasive disease, and serotype V association in causing diseases in elderly [382, 383]. While serotype Ia, III V and Ib found prevalent in pregnant and non-pregnant women of childbearing age colonised with GBS [339, 383, 384]. These Australian studies were also included in a recent 2020 review to understand worldwide distribution of GBS serotypes by risk population group [265].

## 1.15 Trends of GBS disease in neonates and adults in US; 1997-2019.

Active Bacterial Core Surveillance (ABCs) was established in 1995 as part of the Centers for Disease Control (CDC) and Prevention Emerging Infections Program (EIP) [385] network that performs annual population and laboratory based iGBS disease surveillance for the last 22 years (1997 to 2019). These reports are published online [386] and contains investigations on the trends, age and geographical distributions of GBS disease (with no serotype information) in different age group people belonging to

10 EIP sites (Colorado, Connecticut, Georgia, Maryland, Minnesota, New Mexico, New York, Oregon, Tennessee and California,) across the United States [386].

The annual incidence rate of iGBS disease in US population of all age groups slightly increased over the last 22 years, reported as 6.2 in 1997 and 9.9 in 2019 per 100,000 population (**Figure 1.16**). Like UK elderly above 65 years, a continuous increase of iGBS disease is seen in US adults for this age group (20.7 in 1997 and 30.5 in 2019 per 100,000 population), but the annual rates of GBS disease were comparatively higher in US than UK seniors for the analysed period (**Figure. 1.16**). The rates of iGBS disease in US infants under 1 year age group has declined approximately (70%) since 1997, when implementation of intrapartum antibiotic prophylaxis to prevent GBS disease began in US and the 1996 GBS prevention guidelines were issued [387]. This is contradictory to the GBS infection rates seen in UK infants (<1 yr old) where a 20-fold increase was observed over the last two decades (**Figure 1.17**). During 1999 and 2001, the incidence of GBS disease in US infants <1 yr achieved a plateau of about 99.1 cases in 2000, from 81.2 cases noted per 100,000 population in 1999. After the issuance of 2002 guidelines, which changed previous rules of risk based GBS screening to universal GBS screening in pregnant women [388], incidence declined further and in 2019 rates were recorded as low as 57 cases per 100,000 population (**Figure 1.16**). Like the UK population, for individuals aged between 1 and 64 years, no significant change in the annual incidence rate of iGBS was noted over time (**Figure 1.11 and Figure 1.16**).

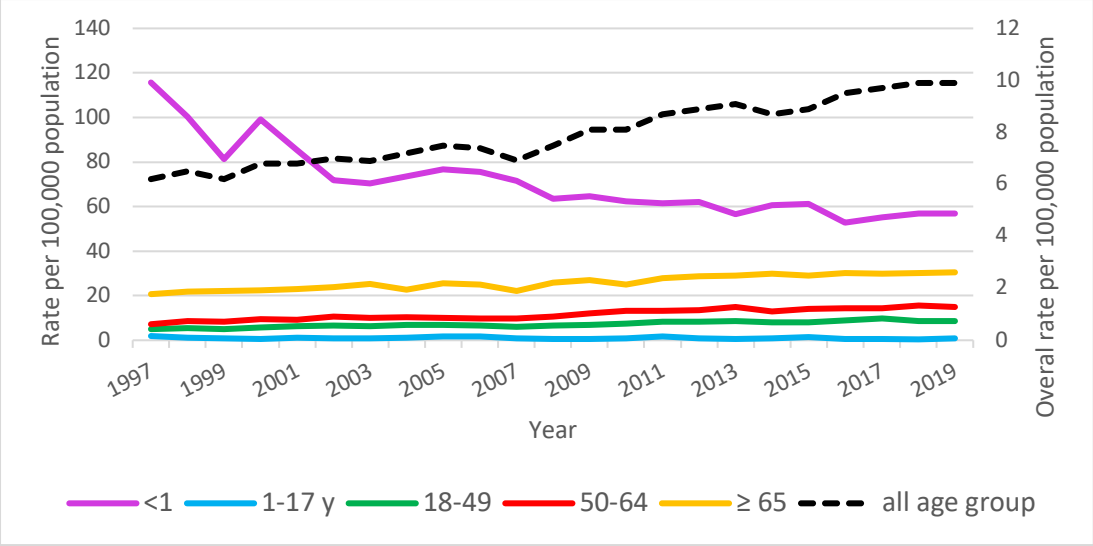



**Figure 1. 16. Age-specific rates of iGBS infection, in United States; 1997–2019.** Annual rates of GBS disease in different age groups per 100,000 population were extracted from ABC surveillance reports on GBS disease in United Sates for 1997 to 2019 [386]. The annual infection rate per 100,000 US population of all age groups is plotted on secondary axis for better visualization.

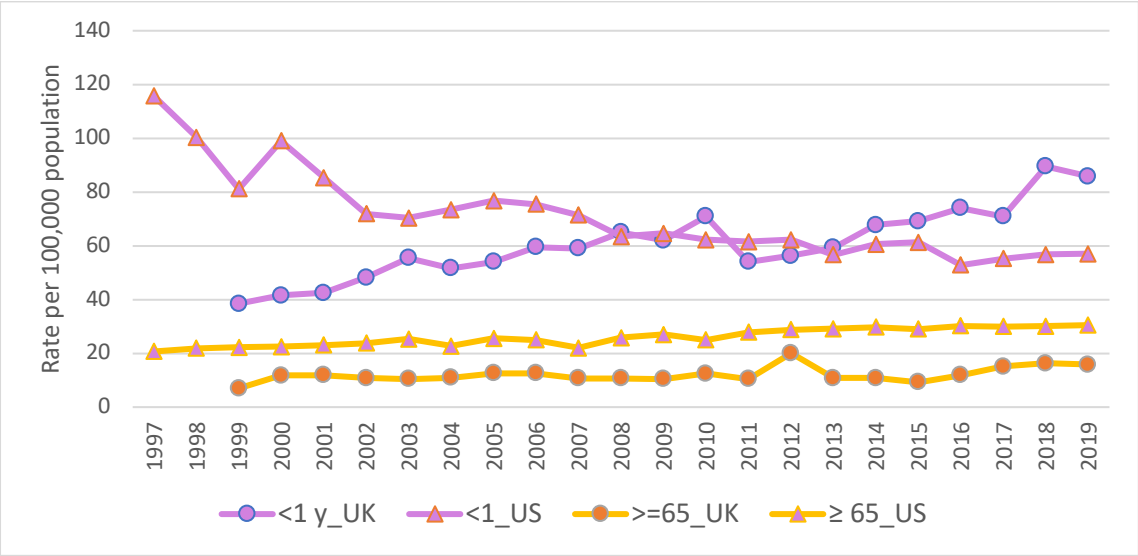



**Figure 1. 17. Comparison of rates of iGBS disease in infants (<1 yr) and elderly (>=65 yrs), in the United States and United Kingdom in late 1990s-2019.** Annual rates of GBS disease in two age groups per 100,000 population were extracted from ABCs reports [386] and PHE surveillance reports **[8].** ABC surveillance published their earliest report for 1997, while PHE published first surveillance report for the year 1999.

A more pronounced decrease is seen in US GBS-EOD throughout the period between 1997–2019 with a rate of 0.7 per 1000 live births recorded in 1997 to 0.19 in 2019 (**Figure 1.18**). For GBS-LOD, it may

appear as a more stable trend over the last 2 decades with 0.3 cases per 1000 live births recorded in 1997 and 0.33 in 2019 (**Figure 1.18**), implying that IAP has no effect in decreasing LOD cases.

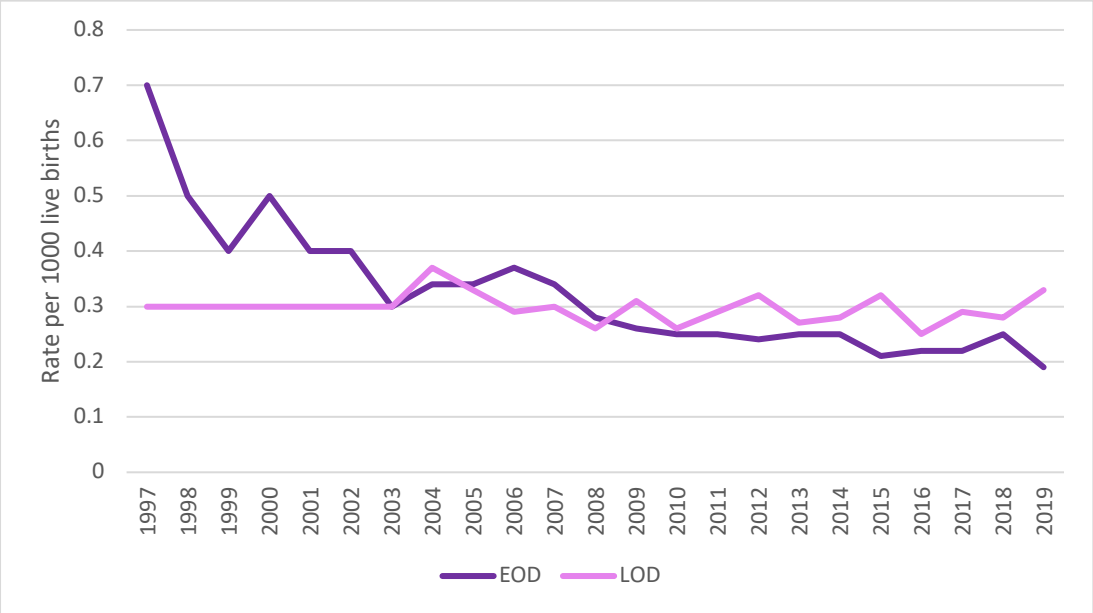



**Figure 1. 18. Rates of GBS-EOD and GBS-LOD infection per 1000 live births in United States 1997–2019.** Rate of GBS infection (EOD and LOD) in neonates <3 months old were extracted from the ABCs reports [386] for the year 1997–2019 to plot the trends.

No information about GBS serotype distribution and antibiotic susceptibility was provided in any of the published ABCs annual surveillance reports for GBS causing infections in US population between 1997-2019. In a recently published study, 6,340 US GBS isolates were characterized during 2015–2017 as a part of ABC surveillance program [389], including a majority of isolates from adults age more than 18 years (5778, 91.1%), from neonates age less than 3 month (506, 7.9%), from infants age between 90 days and 17 yrs (56, 0.8%) and from pregnant women (47, 0.8%). Of 6340 isolates, 98.4% were only serotyped among which the most frequent serotype was serotype Ia (21.8%), followed by V (17.66%), II (17.1%), III (15.6%), Ib (14.5%) and IV (11.8%) [389]. Serotype distribution was separately defined for GBS causing EOD (n=232), LOD (n=272) and disease in all other age groups (<18 yrs, n=5834) [389]. Five serotypes (Ia, Ib, II, III, and V) accounted for 203 of 232 (87.5%) EOD isolates, and 5028 of 5834 (86.2%) isolates of all age groups, while majority of GBS-LOD isolates were serotype III (69%) [389] (**Figure 1.19**). Fifty-four isolates (0.87%) had point mutations within pbp2x associated with

nonsusceptibility to 1 or more β-lactam antibiotics [389]. Overall, resistance to erythromycin and clindamycin was predicted to be 55.2% (n = 3497) and 43.9% (n =2783), respectively. Genes conferring resistance to tetracycline were found in 85.2% isolates [389]. Two isolates carrying *van*G were vancomycin non-susceptible (minimum inhibitory concentration = 2 µg/mL) [389].

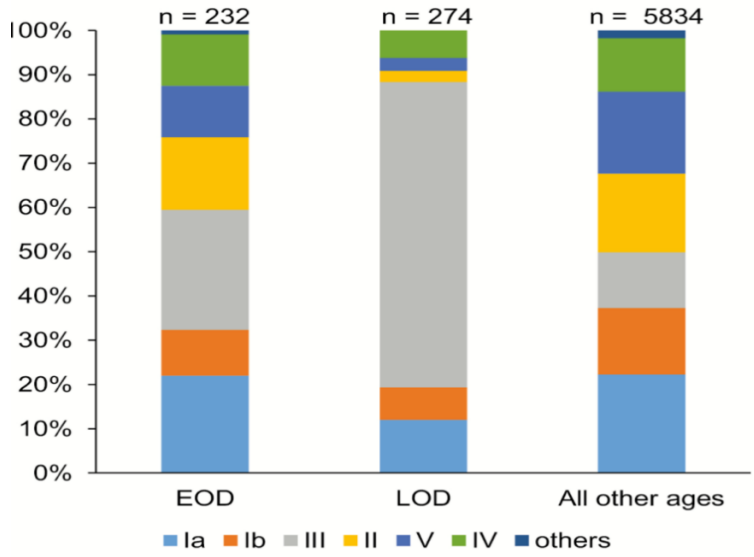



**Figure 1. 19. Distribution of iGBS serotypes among isolates recovered from neonates < 3 months old and from all other ages.** Adapted with permission from [389].

## 1.16 Bacterial evolution

Evolution of bacterial genomes occur by different means including point mutations, large genomic variations, including insertions, deletions, inversions, translocations, duplications, genome reduction (due to multiple genomic region loss), variation in numbers of short repetitive sequences and homologous recombination [390]. Advances in sequencing technologies have made it feasible to capture whole genome sequence variation for each sample under study, providing the potential to detect all changes at all positions in the genome from single nucleotide changes to large-scale insertions and deletions [391].

### 1.16.1 Point mutations

Point mutations are the small changes that include the substitution of one nucleotide with another (commonly known as single-nucleotide polymorphism [SNP]) or the insertion or deletion of a single nucleotide [390]. SNPs can either result in synonymous (silent) codon mutation, and thus no change in the encoded amino acid, or a nonsynonymous mutation, resulting in an amino acid change that may have an impact on the function of the encoded protein. [390]. Single nucleotide insertion or deletion in protein coding sequences will result in a frameshift such that the downstream codons, including stop codons, will be translated from a different reading frame, leading to significant alteration of the encoded protein [390]. Studying the function, rate, type and distribution of point mutations can give insights into the evolutionary pressures on bacterial pathogens [390].

### 1.16.2 Large scale insertion events

Large scale insertion events in bacterial genomes play a significant role in the emergence of epidemic clones or even the pathogen species itself [390]. Examples include insertion of prophages or ICE (integrative and conjugative elements). The cargo of these elements includes genes for virulence factors or drug-resistance determinants [390].

### 1.16.3 Gene Loss

Although gene loss can be mediated by insertion of transposable elements, gene loss via deletion, sometimes after gene disruption, is another major force that drives the evolution of pathogenic bacteria [390]. The sequences of bacterial genomes have revealed that many species have a mutational bias towards deletions events [392, 393], which results in pseudogenization (process by which a functional gene becomes non-functional) and gene loss [394].

### 1.16.3.1 Role of gene loss in GBS serotype diversity

GBS can be classified into 10 distinct serotypes varying in polysaccharide composition. Acquisition and deletion of genes encoding carbohydrate modifying enzymes relative to other serotypes is directly responsible for the variation in serotype reactions. Specific serotypes are defined by reactivity with characterised antisera by using Lancefield precipitation or latex agglutination tests [395]. Further, some isolates are considered as non-typeable (NT) because they do not react with any anti-capsular serum [395]. The frequencies of detected NTs in GBS epidemiological studies have decreased with the development of serotyping methods of improved sensitivity [232, 396]. However, they still account for 5–20% of the isolates colonizing or infecting human adults [341] and for 30–77% of those obtained from bovine mastitis [366]. The GBS NT phenotype could in principle be due to low or complete lack of capsule expression, or to the presence of yet unknown polysaccharide variants that do not react with the available typing antisera [396] or due to laboratory error in interpreting serotyping analysis. Molecular typing methods based on PCR amplification of the different GBS *cps* alleles have allowed assigning capsular genotypes to most NT strains [366, 396, 397]. Ramaswamy *et al* [398] identified an IS*1381* insertion in the *cps*E gene as possibly responsible for the NT phenotype of a capsular genotype V strain. Furthermore, a single strain lacking the entire capsule locus was described in 2012 [399]. Most recently in 2015 Rosini *et al* [400] undertook a comprehensive study to understand molecular basis for the loss of capsule expression in GBS, a large collection of GBS non-encapsulated colonizing and invasive, both human and cattle strains were investigated. A total of 125 different types of genetic events potentially responsible for the lack of capsule expression were detected in 89/96 NT isolates with single or multiple genetic alternations. Point mutations causing premature stop codons were the most prevalent (54/125, 463%), followed by transpositions (37/125, 29%), indels targeting a single gene (24/125, 20%) and deletions affecting more than one *cps* gene (10/125, 8%). Further, the main target was found to be *cps*E encoding the portal glycosyl transferase of capsule biosynthesis.

#### 1.16.3.2 Role of pilus genes in GBS diversity

In another study, Teatero *et al* [205] compared CC17 GBS strains causing neonatal and adult invasive disease in Toronto, Canada that were interspersed and cluster tightly in a phylogenetic tree, signifying that they are derived from the same genetic pool. Further CC17 strains showed low levels of intra-CC17 core genome recombination but all had a loss of Pilus Island 1 and acquisition of different mobile genetic elements carrying determinants of antimicrobial resistance occurred simultaneously that contribute to diversity among this group of organisms. Since PI-1 is a potential vaccine candidate [401], loss or replacement of this locus may have implications for the protection offered by pilus subunit-based vaccines. More recently in 2017, Almeida *et al* [402] conducted a thorough genomic analysis of the *in vivo* evolution of CC17 and identify genetic traits converging toward important phenotypes involved in the adaptation and pathogenicity of GBS in humans. The result shows that the most promising candidates for a GBS vaccine [265], such as the capsule, the pilus, and Rib, are frequent targets of evolution suggestive of that this high level of adaptability will need to be carefully considered in the development of preventive strategies against GBS colonization [403].

### 1.16.4 Short repetitive sequences

Repeats can be an important source of variation in bacterial genomes, however, studying them on a genome-wide scale can be challenging, due to the problems in differentiating one repeat copy from another either by assembly or read mapping [404, 405]. The ability to do this is also dependent on other sequencing parameters, including read length and fragment size [406]. Despite these difficulties, repeats should not be ignored, as they can be instrumental in the evolution and biology of the bacterium. With the recent development of software designed for detecting miss-assemblies [406] and the application of mate-pair sequencing, which provides information on the pair-wise constraints on the placement of reads [407], repeats can be detected with greater confidence. Moreover, the availability of an increasing number of whole-genome sequences has allowed new repeat elements to be discovered and illustrates the role of repeats in bacterial genome evolution [390].

### 1.16.5 Homologous recombination

Classical homologous recombination involves interaction between two sequences with a high nucleotide identity [408]. However, recombination can be responsible for horizontal gene transfer, where more distantly related sequences are exchanged or inserted into bacterial genomes [390]. Additionally, site-specific recombination can mediate the integration of phage genomes or conjugative elements, which involves short stretches of homology between specific sequences of foreign and bacterial DNA [409]. Identification of recombination events can provide insights into this important evolutionary force [390].

## 1.17 Mobile genetic elements (MGEs)

One of the major healthcare associated threats expanding worldwide is antibiotic resistance, particularly multidrug resistant organisms, which increase morbidity, mortality, and healthcare associated costs [410]. Numerous studies revealed that widespread multidrug resistance is frequently obtained through the acquisition of pre-existing gene cassettes followed by amplification in response to antibiotic pressure [411]. The accumulation, acquisition and dissemination of resistance genes are largely mediated through MGE, that promote and enable intracellular DNA mobility (between plasmids or from chromosome to a plasmid) [411].

### 1.17.1 Insertion sequence and Composite Transposons

Discrete DNA segments that can move themselves and carrying resistance genes to random new locations within a single cell's genome are referred as Insertion sequences (IS) and Transposons (Tn) [411]. Conventionally, IS were not considered to carry "passenger" genes but can reorganise resistance genes as part of a composite transposon, a region bounded by two copies of the same or related IS that can move as a single unit [412]. While other elements like integrons (In), move resistance genes between definite sites using site specific recombinases. Since these MGE types are frequently present in multiple copies at different locations within a genome, they facilitate homologous recombination (the exchange of DNA sequences between related or identical segments) [413]. Genetic exchange occurs

via intercellular mechanisms including i) conjugation or mobilization mediated by conjugative plasmids or Integrative conjugative element (ICE), ii) transduction mediated by bacteriophages and iii) transformation (uptake of extracellular DNA) [413].

The majority of IS are comprised of a single transposase (*tnp*) gene and can be divided into groups based on active site motifs found in the Tnp protein [414], including DDE (named for a conserved amino acid triad, Asp, Asp, Glu accumulated at the active site), DEDD (for Asp, Glu, Asp, Asp accumulated at the active site) and HUH (two His residues separated by a large hydrophobic amino acid) motifs [415] or whether the transposition is conservative or occurred through cut and paste mechanism where the IS excise itself from the donor and insert into the recipient [414, 416].

Typical ISs with DDE transposases are generally small (0.7 to 2.5 kb long) and are DNA segments comprised of one or two open reading frames (a transposase and possibly a regulatory protein) [415]. These ISs have inverted repeats (IRs) at their ends and generate direct repeats (DRs) on insertion [415]. The length of the DR is specific to each IS type. Several dominant families among the DDE IS group have been well documented for example IS*3*, IS*4* and IS*5* [412, 415].

### 1.17.2 Unit transposons

These transposons are larger elements than IS and are comprised of a transposase gene and internal "passenger" gene(s) which may encode antibiotic resistance, further they are bounded by IR rather than a pair of IS [417]. An extensive group of unit transposons that can encode large DDE transposases of more than 900 amino acids are Tn*3* family transposons [418, 419]. These transposons work via a replicative transposition mechanism and consist of a *tnp*R resolvase gene, and a resolution (*res*) site, made up of two or three subsites, and may include passenger genes [418]. Another transposon superfamily, known as Tn*7*-like transposons, includes members associated with antibiotic resistance, such as Tn*552* in *Staphylococcus* [420]. Members of this group share some features, such as multiple genes encoding products (including a transposase regulator) involved in transposition rather than the

single long *tnp*A gene found in the Tn*3* family but have different transposition mechanisms [412]. Unlike Tn*3* family transposons, members of this group may also target a particular site(s) [420].

### 1.17.3 Integron

Integrons are genetic elements that allow efficient capture and expression of exogenous genes [421]. They are widely known for their role in the dissemination of antibiotic resistance, particularly among Gram-negative bacterial pathogens [422]. However, since their initial discovery in clinical contexts, it has become apparent that integrons are a common component of bacterial genomes and that they have a long evolutionary history [421]. All integrons share three essential core features, whose combined activities capture and subsequently express exogenous genes as part of gene cassettes [423, 424] The first feature is a gene which encodes an integron integrase (*int*I), a member of the tyrosine recombinase family [425]. The integron integrase protein catalyzes recombination between incoming gene cassettes and the second core feature, an integron-associated recombination site, *attI* [426]. Once a gene cassette is recombined, it is expressed by the third core feature, an integron-associated promoter, (Pc) [427, 428]. Integrons acquire new genes as part of gene cassettes [429]. These are simple structures, usually consisting of a single open reading frame (orf) bounded by a cassette-associated recombination site, referred to as *attC* [430]. Circular gene cassettes are integrated by site-specific recombination between *attI* and *attC*, a process mediated by the integron integrase [431]. This process is reversible, and cassettes can be excised as free circular DNA elements [432]. Insertion at the *attI* site allows expression of an incoming cassette, driven by the adjacent Pc promoter [427].

### 1.17.4 Plasmids

Conjugation is responsible for the dissemination of virulence traits and antibiotic resistance among bacterial populations and is the major process involved in horizontal gene transfer between Gram-negative bacteria [433]. Hence considered as the primary driving force for bacterial evolution and plasticity and is mediated by several types of MGEs [434]. Among them, conjugative plasmids were extensively studied and discovered earliest due to the ease of their identification, wide distribution in

bacteria, massive impact on the spread of pathogenesis and commonly carried antibiotic resistance determinants [435]. Since plasmids are extra-chromosomal double stranded DNA (dsDNA) molecules and encode proteins for their own replication and their horizontal transfer, they are considered as conjugative elements [435]. Compared to Gram-negative and other gram-positive bacteria, plasmids are extremely rare in GBS, and this is probably the reason behind the fact that GBS has never become penicillin resistant and that mobile elements in GBS are particularly rare.

In 1976, for the very first time, a study confirmed the presence of two plasmids in a GBS strain isolated from clinical sputum specimen conferring resistance to tetracycline and to chloramphenicol, erythromycin, lincomycin, and pristinamycin, respectively [436]. Later after 2008, various reports confirmed the presence of macrolide resistance *erm*T on small, mobilizable, broad-host-range plasmids in GBS [204, 437, 438]. Most recently, in 2016, Sendi *et al* [212] published first report showing plasmid-mediated high level gentamicin resistance (HLGR) in GBS. Of 1,128 invasive and colonizing GBS isolates screened for the presence of HLGR, two strains displayed HLGR (BSU1203 and BSU452), both of which carried the *aac*A-*aph*D gene. The first isolate (BSU1203) carried the previously described chromosomal gentamicin resistance transposon, Tn*3706* and later isolate (BSU452) harboured plasmid (pIP501) that confer gentamicin resistance. However, plasmids do not represent a prominent mechanism of antimicrobial resistance gene movement in clinical or agricultural Gram-positive bacterial infections.

### 1.17.5 Genomic Islands

A genomic island (GI) is referred to as a distinct region of a bacterial chromosome capable of horizontal transfer as a unit [439]. They are usually flanked by DR, vary in size, and are classified according to the phenotypes they encode such as multiple resistance determinant carrying GI, referred to as resistance islands, and virulence factor carrying GI, known as pathogenicity islands [439, 440]. A comprehensive definition of GIs includes elements with mobility functions, such as ICEs, and integrative mobilizable elements (IME) [441]. GI can also be often observed as degenerative GI, in

which one or more gene(s) or sequence(s) necessary for intracellular and/or intercellular mobility are inactivated or deleted [442]. Sequence comparisons showed that various degenerative GIs derive from ICEs, and IMEs, by mutations or deletions of their mobility modules [442]. The size range of these elements is very large, going from 75 bp to 680 kb [442].

#### 1.17.5.1 Integrative Conjugative Elements

Besides conjugative plasmids, another family of conjugative elements known as ICE was discovered more recently in various bacteria [442]. Unlike plasmids, these ICEs can integrate into and transfer from the host chromosome to another bacterial chromosome and are replicated as a passive bystander during cell division [442]. ICE typically consist of a backbone (containing phage like integration/excision functions, plasmid-like conjugation/maintenance components, and a regulation module) into which accessory genes are inserted. Excision of the ICE as a circular form and integration of this circle (at low frequency), usually into a unique *attB* site in the host chromosome, are mediated by the ICE-encoded site-specific integrase (Int) such as serine recombinases, tyrosine recombinases or DDE transposases [433]. The 3′ ends of tRNA genes are often target sites, and integration creates DR at the ends of the ICE, known as *attL* and *attR* [439]. While less well-known than conjugative plasmids, the latest studies of bacterial genomes reveal a high frequency of ICEs suggesting their abundance surpasses that of conjugative plasmids [443, 444]. Further, the ICE transfer mechanism is analogous to that of conjugative plasmids although less well documented [409].

#### 1.17.5.2 Integrative and mobilizable elements

Another class of integrated MGE are the IMEs [442]. IMEs encode their own excision and integration proteins irrespective of their integration specificity and/or mechanism and are considered as non-self-transmissible due to the lack of conjugation apparatus [442]. Plasmids play an important in many bacteria in moving antibiotic resistance genes however, in the case of GBS specifically, plasmids are extremely rare and so ICE and IME are more prevalent.

## 1.18 Tn*916*/Tn*916* – like element

The Tn*916* element was formerly designated a conjugative transposon. More recently it has been realised that Tn*916* is an ICE, yet it retains its name as a conjugative transposon due to historical reasons. Tn*916* therefore illustrates a group of related ICE that is reported in a wide range of bacteria [445]. Originally discovered in the late 1970s in *E. faecalis*, Tn*916* was the first conjugative transposon encoding antibiotic resistance to be reported [446]. It is the smallest member of the Tn*916*/Tn*1545* family and contains 24 orfs (**Figure 1.20**), organized into functional modules involved in conjugal transfer, recombination (excision and insertion reactions), transcriptional regulation and accessory functions (antibiotic resistance) [447, 448].

In streptococci, resistances to macrolide and tetracycline are frequently due to the co-localization of *erm*B and *tet*M genes on the same Tn*916* derivatives [449]. An example of this is seen in the earlier GBS study of 2007, where *erm*B gene was found inserted adjacent to the *tet*M gene in a Tn*916*-like transposon and was subsequently designated as Tn*917*, analogous to one identified in *S. pneumoniae* [450]. Whereas, later in 2010 the majority of GBS isolates harbouring the *erm*B gene were found to be carried on another Tn*916* variant – a Tn*3872*-like transposon (**Figure 1.21**) [203]. Such Tn*916* variants have previously been reported in collection of macrolide-resistant *S. agalactiae* and *S. pyogenes* [451].

Other closely related Tn*916*-like elements include Tn*918* [452], Tn*925* [453], Tn*1545* [454, 455], Tn*1549* [456], Tn*2009* [457], Tn*2010* [457, 458], Tn*2017* [459], Tn*3702* [460], Tn*3872* [461], Tn*5031* [462], Tn*5251* [463], Tn*5381* [464], Tn*5383* [464], Tn*5397* [465], Tn*5801* [466], Tn*6000* [448, 467], Tn*6002* [468], Tn*6003* [455], Tn*6009* [469], Tn*6014* [470], Tn*6084* [471], Tn*6084* a and b [471]. These elements *s*hare the same modular structure as that of Tn*916* and confer resistance to tetracycline (via *tet*M or *tet*S) alone or in combination to other antibiotics due to the insertion of additional DNA elements into the basic backbone structure, this includes macrolide (*erm*B or *mef*E), kanamycin (*aph*A -3), mercury resistance (*mer*A or *mer*B) and/or vancomycin (*van*B*)* resistance genes [455]. The structure of various members of the Tn*916/*Tn*1545* family is shown in (**Figure 1.21**).


**Figure 1. 20. Schematic representation of Tn*916*.** The 24 orfs of Tn*916* element are represented using arrows where different colours highlight important coding sequence i.e., dark grey show accessory gene *tet*M, red – recombination genes including excisionase and integrase and light green showing remaining orfs. Adapted with permission form [470].

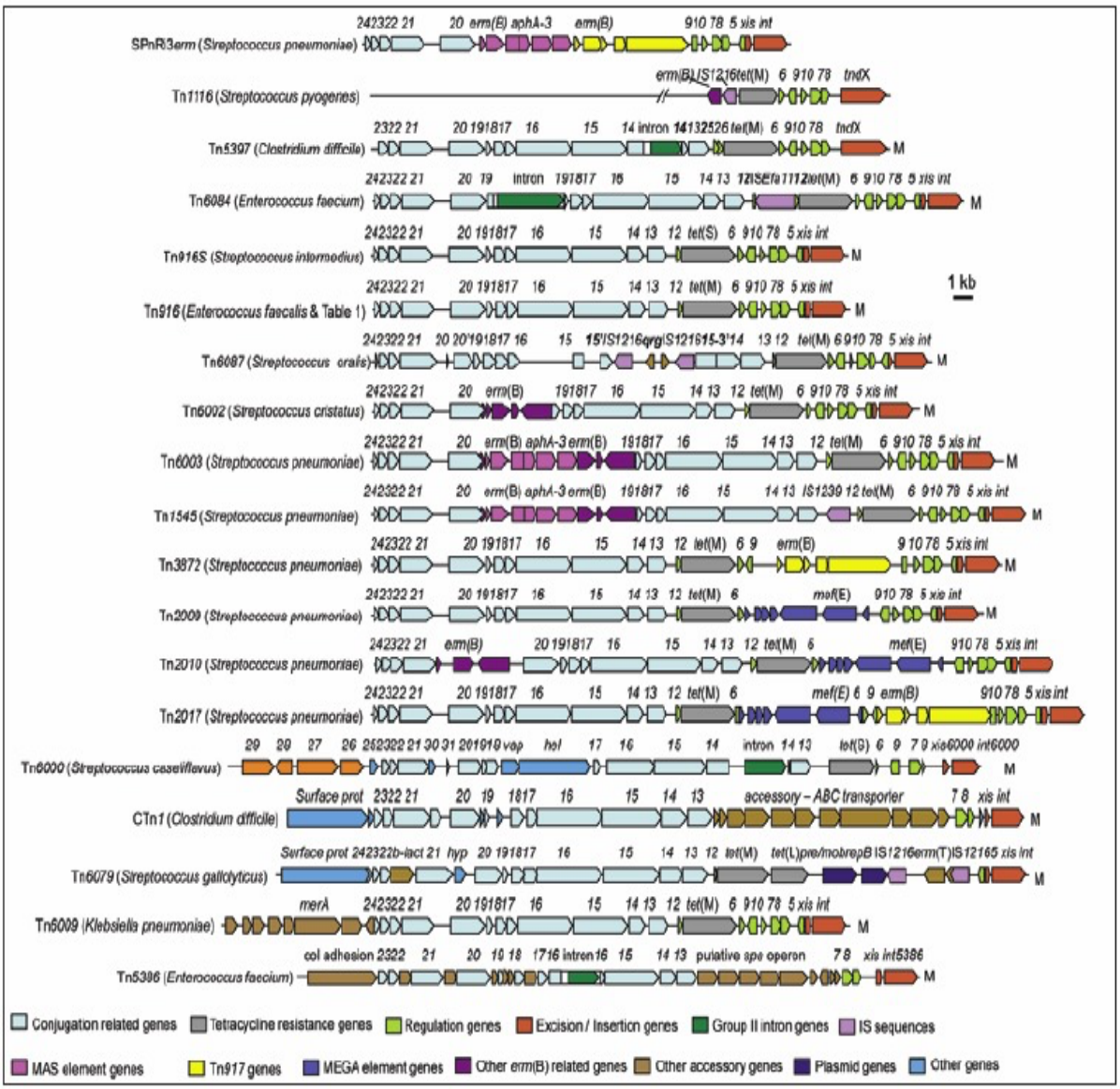



**Figure 1. 21. The structure of various members of the Tn*916*/Tn*1545* family.** Functional modules are represented as shown in the key. The organisms from which the elements were isolated are shown to the left in brackets. Mobility is denoted by a capital M on the right. Adapted with permission from [472].

Gram-positive pathogenic bacteria such as GBS represent an excellent example of the conjugation influence on the bacterial lifestyle, as mobilised DNA elements exchanged between strains are constantly driving selective advantage for resultant strains [345]. The GBS chromosome demonstrates a complex arrangement and organisation consisting of a conserved backbone and up to 69 variable regions [97, 473]. Comparative study of the first eight sequenced GBS genomes showed that 2/3rds of the variable regions corresponds to 12 putative ICEs and 6 IMEs, at 15 different locations in the GBS chromosome [473]. Of the various ICEs discovered in GBS, a novel family of ICE, named Tn*GBS*s, was identified which promotes mobilization of chromosomal DNA using high frequency (Hfr)-type mechanisms and whose movement is facilitated by a transposase with a DDE motif, as opposed to traditional serine/tryosine integrases [474]. In Hfr-type the conjugative transposon can transfer a portion of itself by conjugation without excising from the chromosome [475].

So far, many ICEs have been reported in GBS strains over the last few years including ICE*Sa*2603 and ICE*Sag*2603VR-1 in S. *agalactiae* 2603V/R [98] and Tn*GBS1* and Tn*GBS2* in *S. agalactiae* NEM316 [96]. ICE*Sa*2603 is 54kb in size and carry heavy metal resistance genes for mercury (Hg), cadmium (Cd) and copper (Cu) with no antibiotic resistance genes [98] (**Figure 1.22**), while ICE*Sag*2603VR-1 is a Tn*916* like element (~18kb) and carried *tet*M gene only [98]. The Tn*GBS1* and Tn*GBS2* elements did not carry any obvious cargo genes, such as antibiotic resistances or metabolic functions [476]. ICE*Sag*37 was found to carry virulence factors and multiple drug resistance genes for aminoglycosides (*aad*E, *aph*A, and *ant*(6)], erythromycin (*erm*B) and tetracycline (*tet*O) in a GBS strain causing neonatal bacteremia in Shanghai [477]. In 2016, a new mosaic ICE*Sag*236 carrying chloramphenicol (*cat*Q) and macrolide (*mef*I and *erm*TR) resistance genes was identified, suspected to result from the recombination of two parent ICEs; ICE*Sag*TR7 and ICE*Spn*529IQ [478]. Another study in 2016 reported the presence of two ICEs namely ICE*Sag*(RR1) and ICE*Sag*(RR2) harbouring multiple resistance to aminoglycoside (*ant*(6), *aph*A), macrolide (*erm*B), and tetracycline (*tet*O) antibiotic classes in GBS causing neonatal

invasive disease in China [479]. ICE*Sag*(RR1) carried two copies of the *ant*(6) gene while ICE*Sag*(RR2) carried three copies of *ant*(6) gene [479].


**Figure 1. 22. Schematic of ICE*Sa*2603.** EasyFig v 2.2.2 [336] was used to represent differet orfs of ICE*Sa*2603 (Accession no. AE009948.1). Annotation was performed using Artemis [337]. Different colours are indicating important features of ICE*Sa*2603 where red showing conjugal proteins, dark green – heavy metal resistance genes and blue – integrase gene. The genomic size of ICE*Sa2*603 is 54kb.

GBS is comprised of an evident surprisingly stable clonal structure, both geographically and temporally. However, bacterial diversification is an ongoing process and depends on local factors. Recombination leading to capsular switch is a key to diversification, however the frequency is not yet clearly defined and may be increased by selective pressure within a population receiving future vaccines. Such recombination should be considered in surveillance of the impact of and development of vaccines. Regardless of expanding reports on invasive disease and maternal colonization, an improved understanding of natural reservoirs of GBS and colonization in adults is necessary to manage GBS infections appropriately. Further continuous increase in antibiotic resistant GBS strains globally against normally prescribed drugs, the emergence of infrequent GBS serotypes causing invasive disease (and of conjugative element-positive strains acquiring antibiotic and virulence genes) highlight the importance of continuous surveillance for a better understanding of antibiotic resistance mechanisms, modes of transmission and the dynamic nature of GBS population.

# Chapter 2:

# Materials and Methods

## 2.1 Collection of clinical specimens

### 2.1.1 Ethics Considerations

Collection and analysis of bacterial isolates from the UK were classified as service evaluation and clinical investigation, and so was exempted from NHS Research Ethics Committee Review and Original ethical approval covering strains collected in Australia were approved by the Human Research Ethics Committee of Western Australian Department of Health, Women and Newborn Health Service (2056/EW) and for those collected in Brazil were approved by the Research Ethics Committee of the Faculty of Health Sciences of the Universidade de Brasília (UnB) CAAE (38856114.0.0000.0030).

## 2.2 UK isolates

### 2.2.1 Sample collection, DNA extraction and WGS in PHE laboratories

In 2015, a total of 1870 reports of GBS bacteraemia cases were reported by laboratories in England, Wales and Northern Ireland (EWNI) to PHE, a 7% increase compared to 2014 (n= 1676) and accounting for 36% and 32% of total pyogenic streptococci infections reported in 2014 (n=4,912) and 2015 (n= 6,061), respectively [480, 481]. Isolates received were from patients of all ages, with the majority from neonates (estimated >95%) and women of child-bearing age, with an estimated total adult isolate submission proportion of less than 11% of all invasive cases. Between 2014 and 2015, rates of GBS disease were higher in males than females in all age groups except the 15 to 44 years age group; also, the selection of isolates used in this thesis is heavily biased towards females representing >50% of isolates from females belonging to 18-44 years age group and do not represent entire UK GBS collection for these years.

The UK GBS isolates used in this thesis were 194 from invasive and non-invasive sites, isolated from adults aged >18 years. These GBS isolates were routinely submitted to the PHE reference laboratory as part of national surveillance between January 2014 and May 2015. The available 194 samples were received from PHE without prior knowledge of accompanying patient epidemiological or clinical data.

Later the WGS and the associated anonymised clinical data on age and sex of the patient, date of sample isolation, site of isolation and geographical region, was provided by PHE. Strains were assigned to three different groups regarding anatomical site of isolation: (i) invasive strains (n =179) for isolates recovered from blood, cerebrospinal fluid (ii) non-invasive strains (n =14) for isolates recovered from vagina, wound, pus and abscess specimens, tissue and placenta, and one isolate where no information about its source was provided. These 12 non-invasive isolates and one from unknown site were excluded from the analysis.

In the PHE laboratories, genomic DNA was extracted from 5-6 colonies per isolate after 24 h incubation at 37 °C on 5% horse blood agar (PHE media services) using the Qiasymphony (Qiasymphony DNA Mini kit, Qiagen) after pre-lysis at 37 °C for 30 min and 100 °C for 10 min with 30 units mutanolysin and 0.06 mg hyaluronisdae (Sigma-Aldridch). DNA was prepared using Wizard Genomic DNA Purification kit (Promega, USA) at PHE. Nextera XT DNA Library Preparation kit (Illumina, USA) was used to prepare multiplexed DNA sequencing libraries. Whole genome sequencing was undertaken using Illumina HiSeq 2500 System (Illumina, US) and 2 X 100bp paired-end mode.

Reads were submitted as PHE Pathogens BioProject PRJEB18093 at ENA - http://www.ebi.ac.uk/ena/data/view/PRJEB18093. Strep-B-Latex kit (Oxford, Biosystems) was used for the serotyping of 194 GBS strains based on ten known serotypes (Ia, Ib-IX). The (10 µl) of overnight bacterial cultures of GBS isolates (preferably in ToddHewitt broth) were mixed with the (10 µl) Strep-B-Latex suspension on a plastic slide. A positive reaction was indicated by an agglutination within 5 to 10 seconds, if the reaction time exceeds 30 seconds, false positive reactions may occur. Phenotypic identification methods such as catalase or gram staining were not performed.

### 2.2.2 Sample processing in Cardiff University

Isolated strains (n=194) archived at PHE were sent to Cardiff University on blood agar slants, where they were grown in Todd Hewitt broth and stored at -80ºC in glycerol for further investigation.

Antimicrobial susceptibility testing (AST) using Initial Kirby-Bauer disk diffusion method (described in detail in **section 2.5.2.1**), D-factor test (**section 2.5.2.2)**, and Minimum Inhibitory Concentration (MIC) test (**section 2.5.2.3)** were then performed. DNA extraction, and re-sequencing with combination of short and long read methods (**section 2.5.3)** was executed on few isolates (n=14) found to carry multiple antibiotic resistance genes on a potential mobile genetic element. Of 194 adult GBS samples, 1 sample did not pass the assembly quality check (**Table 2.4 and section 2.5.4**), therefore the final UK GBS data set consist of 193 GBS assembled genomes.

## 2.3 Brazil isolates

### 2.3.1 Sample collection at RJ Hospital, Brazil

Forty-six iGBS isolates were collected, when available, between June 2010 to May 2017 in a tertiary care hospital in the Southeast region of Brazil, Rio de Janeiro, Brazil, and does not represent the actual number of GBS isolates collected between this period (June 2010 – May 2017). Currently there is no routine GBS surveillance in place in Brazil, hence there is no information on the total number of GBS cases reported in different regions of Brazil per year. All isolates were isolated from blood culture. Clinical date and site of sample isolation and age/sex of the patient was provided.

### 2.3.2 Processing of Brazilian GBS isolates in Cardiff University

The collected 46 Brazilian iGBS isolates were sent to Cardiff University on charcoal swabs by Dr. Diego Andrey, (Infectious Diseases Specialist at University of Geneva, Switzerland) and Prof. Ana Gales, (Professor in the Department of Infectious Diseases, at the Federal University of São Paulo, Brazil) in UN3373 compliant containers following attainment of appropriate export/import paperwork. These isolates were cultured (**section 2.5.1**), processed for DNA extraction (**section 2.5.3**), and sequenced after transfer to Cardiff University. Out of 46, only 34 samples including 27 samples from adults, six from neonates and 1 with no information available on age were confirmed as GBS strains using WGS analysis. Six neonatal samples due to its odd representation in the total number of isolates

and one with no age information of patient were excluded from the analysis. Of 27 adult GBS samples, 1 sample did not pass the assembly quality check (**Table 2.4 and section 2.5.4**), therefore the final Brazil GBS data set consisted of 26 strains. AST and MIC tests (**section 2.5.2.1 and 2.5.2.3**) were performed to detect antimicrobial susceptibility profiles.

## 2.4 Australian Isolates

### 2.4.1 Sample collection at King Edward Memorial Hospital, Western Australia

Antenatal isolates were collected upon written consent as part of a larger cohort study, Predict1000, which was approved by the Women and Newborn Health Service Human Research Ethics Committee (201535EW). This study was conducted at King Edward Memorial Hospital, Perth, the largest Tertiary Obstetrics Hospital in Western Australia (WA), over the period 2015–2017 [384]. Specimen collection involved self vaginal and rectal e-swabs (Copan, Italy) collection from 815 pregnant women at ≤22 weeks (*n*=814) and ≥33 weeks' (*n*=567) gestation. Each self-collected swab per site (vaginal and rectal) were used for culture, which involved placement of the swab tip (no prior vortexing) into Strep B Carrot Broth (Hardy Diagnostics, USA) and overnight culture at 35 °C. The formation of orange pigment indicated pigmented/haemolytic GBS presence. All broths were sub-cultured onto StrepB CHROMagar (CHROMagar, France) and incubated at 35 °C for 24–48 h, regardless of result, to select for pure GBS colonies and to identify non-haemolytic isolates that would not be indicated by Carrot Broth. Mauve colonies (indicative of GBS) were sub-cultured onto CHROM and incubated as above to confirm purity, followed by tryptic soy broth (Becton Dickinson, Australia) sub-culture and storage in 15 % glycerol at −80 °C. For molecular serotyping, frozen samples were thawed (once only) and extracted using the InviMAG Universal kit (Stratec Molecular GMBH, Germany) as per the manufacturer's instructions on a KingFisher DUO (Thermo Fisher Scientific, USA) platform, and the resulting DNA was stored at 4 °C until analysis. The DNA was tested via multiplex PCR to detect GBS presence (ubiquitous *dltS* gene) and additional serotype Ia, Ib or III presence as previously described [384]. As this assay detects 3 of 10 serotypes, all *dltS*-positive isolates were further tested using the

Imperi standard PCR method [482] and amplicons were visualized using gel capillary electrophoresis on a QIAxcel Advanced platform (Qiagen, USA). Combined culture and PCR results were used to inform GBS colonization status (positive with one or both of these methods).

### 2.4.2 Sample processing in Cardiff University

A total of 171 GBS isolates were recovered from vaginal and/or rectal (often paired from same patient) at ≤22 weeks of gestation (*n*=567) from WA Australian pregnant women cohort [384] and were available for the analysis. These isolates were sent as e-swabs to Cardiff University by Dr Matthew S. Payne (research lead for this study) as purified isolates in glycerol stocks on dry ice in UN3373 compliant packaging following attainment of appropriate export/import paperwork. The isolates (n=171) were cultured (**section 2.5.1**) and tested for AST (**section 2.5.2.1**) but were not sequenced as per the initial plan due to Covid pandemic in UK that shut down Cardiff University lab for almost a year. Date of collection of specimens, age of pregnant women, weeks of gestation, and serotype information were provided in metadata. The serotype distribution and antibiotic resistance profiles of GBS population circulating in three geographic regions (Australia, Brazil and UK) were assessed to provides some insight into the serotype prevalence and antibiotic resistance rates of GBS in different study populations to inform vaccine development.

## 2.5 GBS isolates processed at Cardiff University

### 2.5.1 Bacterial culturing

The UK GBS isolates (n=194), Australian maternal vaginal and/or rectal e-swabs and Brazilian blood isolated GBS samples (171 and 46), respectively were inoculated onto on 5% horse blood agar (Oxoid, Basinstoke, UK) and placed in Todd-Hewitt broth with gentamicin, nalidixic acid, and yeast extract. The culture plates and the broth were incubated for 18 to 24 h at 36°C. Samples from the Todd-Hewitt broth were then sub-cultured onto blood agar plates. The medium was incubated anaerobically and examined at 18 to 24 h for orange-red colonies of GBS and re-examined at 48 h if results were negative

at 24 h. The blood agar was incubated in 5% $CO_2$, and colonies of GBS were identified based on colony morphology, Gram staining, and catalase reaction (Streptex B; Murex Biotech Ltd., Darford, England). All *S. agalactiae* positive cultures were confirmed by PCR and stored at -80°C as pure isolates for subsequent sequencing or AST analysis.

### 2.5.2 Antimicrobial screening

#### 2.5.2.1 Kirby-Bauer disk diffusion

##### 2.5.2.1.1 Australia, Brazil, and UK population isolated GBS

The Kirby-Bauer test for antibiotic susceptibility (also called the disc diffusion test) is a standard that has been used for years [483]. This test is used to determine the resistance or sensitivity of aerobes or facultative anaerobes to specific chemicals, which can then be used by the clinician for treatment of patients with bacterial infections [483]. The presence or absence of an inhibitory area around the disc identifies the bacterial sensitivity to the drug (**Figure 2.1**) [484].

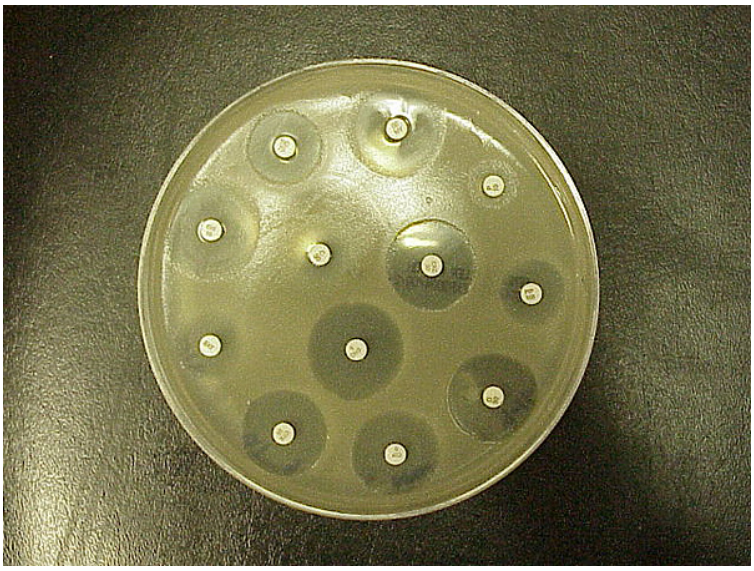


**Figure 2. 1. Kirby-Bauer test to measure antibiotic sensitivity.** Adapted with permission from [484].

Initial Kirby-Bauer disk diffusion methods were carried out for ampicillin (10µg), vancomycin (30 µg), tetracycline (30 µg), erythromycin (15 µg), clindamycin (2 µg), chloramphenicol (30 µg), gentamicin (120 µg), and levofloxacin (5 µg) (all purchased from Oxoid, Basinstoke, UK) on Columbia sheep blood agar plates (CA) (Oxoid, Basinstoke, UK), at 37ºC at 5% $CO_2$, under methods compliant with Clinical and Laboratory Standards Institute (CLSI) guidelines M100S [485] for definitions of inhibition zones consistent with sensitivity, intermediate resistance and full antibiotic resistance. A small number of colonies were transferred by sterile loop to 3ml of sterile saline to generate a 0.5 McFarland standard standardized inoculation. A uniform lawn of GBS for each isolate was then generated on the blood agar using a sterile cotton swab. The 8 different antibiotic disks were then "stamped" onto the surface of the inoculated plate and allowed to dry/set for 5 minutes prior to inversion (plate lid down) and transfer to the 5% CO2 incubator for incubation overnight. The antibiotic diffuses from the disc into the agar in decreasing amounts the further it is away from the disc. If the organism is killed or inhibited by the concentration of the antibiotic, there will be no growth in the immediate area around the disc: This is called the zone of inhibition.

These inhibition zones are measured the following morning and recorded. Zone sizes are compared to an internationally standardized chart (in this case interpreted according to the Clinical Laboratory Standards Institute M100-S27 Revised AST Breakpoints Document, Table 2H-1 for β-haemolytic Streptococci) to give a result of sensitive, resistant, or intermediate resistant (**Table 2.1**). This table also lists the accepted breakpoint thresholds for MIC (minimal inhibitory concentration) for that antibiotic. The MIC is currently the standard test run for antibiotic sensitivity testing because it produces more pertinent information on minimal dosages. The procedure defined below is adapted from CLSI guidelines [485].

**Table 2. 1. Zone Diameter and Minimal Inhibitory Concentration Interpretive Standards for *S. agalactiae* asper CLSI guidelines M100S** [485]**.**

| **Antimicrobial agent** | **Disc content (µg)** | **Zone diameter breakpoints (nearest mm)** | | | **MIC standard (µg/mL)** | | |
|---|---|---|---|---|---|---|---|
| | | **S** | **I** | **R** | **S** | **I** | **R** |
| Gentamicin | 120 | >=10 | 7-9 | < 6 | | | |
| Ampicillin | 10 | >=24 | - | - | <=0.25 | - | - |
| Chloramphenicol | 30 | >=21 | 18-20 | < 17 | <=4 | 8 | >=16 |
| Clindamycin | 2 | >=19 | 16-18 | < 15 | <=0.25 | 0.5 | >=1 |
| Vancomycin | 30 | >=17 | - | - | <=1 | - | - |
| Tetracycline | 30 | >=23 | 19-22 | < 18 | <=2 | 4 | >=8 |
| Levofloxacin | 5 | >=17 | 14-16 | < 13 | <=2 | 4 | >=8 |
| Erythromycin | 15 | >=21 | 16-20 | < 15 | <=0.25 | 0.5 | >=1 |

S – Susceptible, I – Intermediate and R – Resistant

### 2.5.2.2 D-factor test

#### 2.5.2.2.1 Brazil and UK GBS isolates

D-factor tests is a particular variant of the Kirby-Bauer method where the clindamycin and erythromycin disks are placed within 15 mm of each other [486]. This detects inducible clindamycin resistance in the presence of erythromycin and were performed for induction of clindamycin resistance by erythromycin according to CLSI (2008) instructions [486]. Interpretation of the D-factor effect is shown in **Figure 2.2**.

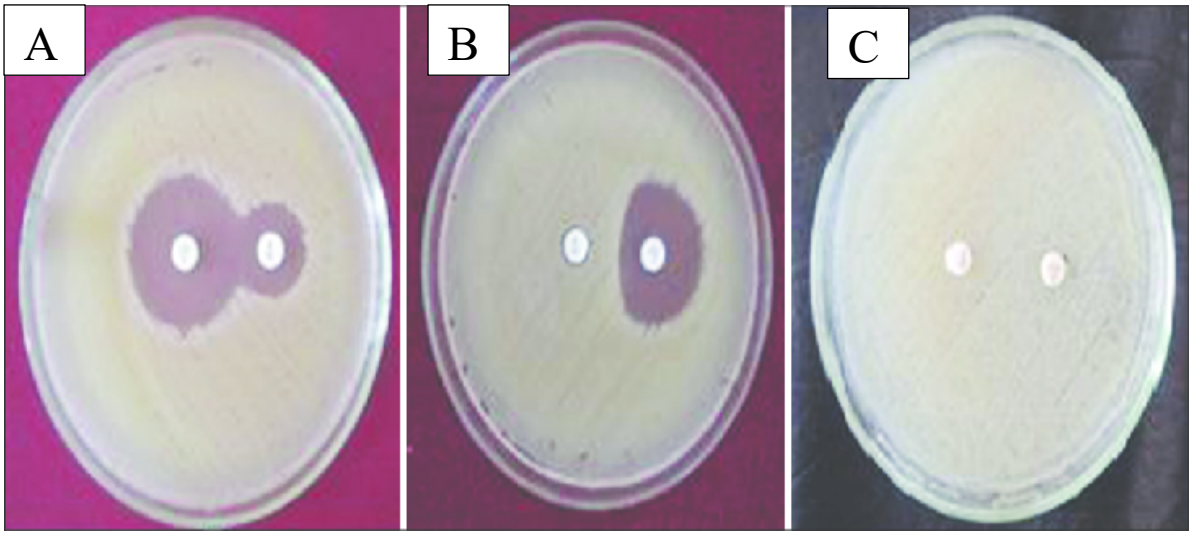



**Figure 2. 2. Disc diffusion test for inducible clindamycin resistance. (A)** Sensitive phenotype **(B)** inducible clindamycin resistance ($iMLS_B$) phenotype, clindamycin disk on the right and **(C)** constitutive macrolide and clindamycin resistance ($cMLS_B$) phenotype. Adapted with permission from [487]**.**

#### 2.5.2.3 Use of antimicrobial gradient strips to define accurate MICs

##### 2.5.2.3.1 Brazil and UK GBS isolates

The Oxoid M.I.C. Evaluator ™ Strip (Thermo Fisher Scientific, Basingstoke, UK) of antimicrobial gradient strip was used to determine MICs of strains showing resistance to gentamicin, vancomycin, chloramphenicol, levofloxacin, tetracycline, erythromycin, and clindamycin resistant isolates identified by Kirby-Bauer disk screening, by following the manufacturer's guidelines. A blood agar plate inoculated with a uniform GBS lawn for the strain of interest was prepared as above, before placing a strip impregnated with antimicrobials onto an agar plate (Figure 2.3) [488, 489]. A strain of bacterium will not grow near a concentration of antibiotic or antifungal if it is sensitive; the results can be used to determine a minimum inhibitory concentration (MIC) [489].

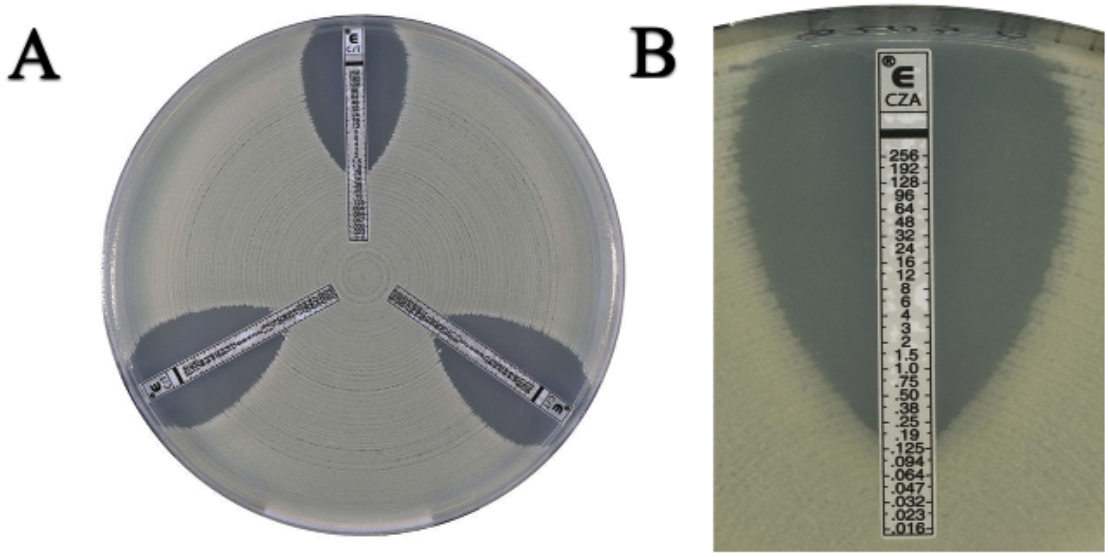



**Figure 2. 3. MIC test using strips. A)** Antimicrobial gradient strip showing MIC strip placed on CA plate inoculated with bacteria. **B)** showing zone of inhibition of a specific antibiotic strip used. Adapted with permission from [**https://microbenotes.com/epsilometer-test-e-test/**].

### 2.5.3 DNA extraction and genomic sequencing

Bacterial pellets were processed and sequenced using next generation sequencing (both Illumina and in some cases Nanopore) by the Bacterial WGS facility at Cardiff University School of Medicine. Physical preparations by Edward Portal and Ian Boostrom; MiSeq loading Edward Portal, Dr. Kirsty Sands and Jordan Matthias; fastq processing and analysis pipeline creation and maintenance by Dr. Kirsty Sands and Dr. Andrew Roberts.

#### 2.5.3.1 Brazil GBS isolates

All Brazilian GBS isolates (n=46) were extracted using the QIAGEN MagAttract Microbial DNA Extraction Kit (QIAGEN) as per manufacturer's instructions using the KingFisher Duo (ThermoFisher Scientific) automated platform. Extracted DNA was quantified on a QubitTM fluorometer (Life Technologies) with a dsDNA broad range kit (Life Technologies), and 1.0 ng of DNA was used in the Illumina Nextera XT library preparation protocol as per manufacturer's instructions (Illumina). The genomic DNA was fragmented, indexed by PCR and purified using AMPure XP beads. Library pooling

was performed after quantification of the library size using the LabChip® GXII. The pooled DNA library (12 pM) were sequenced on an Illumina Nextseq using a paired-end 300 bp read workflow.

#### 2.5.3.2 UK GBS Isolates

A total of 14 UK GBS isolates were re-sequenced with combination of short and long read methods. GBS isolates were regrown, and the same DNA extract was analysed by both Illumina MiSeq platform (Illumina, San Diego, California) and Oxford Nanopore MinION platform (**Table 2.2**). Colony purified GBS for each re-sequenced sample were grown on blood agar plates as a bacterial lawn overnight at 37 ºC at 5% $CO_2$ and the total growth scraped into 1 ml sterile PBS and pelleted for extraction. The genomic DNA (gDNA) was extracted using the QIAmp DNA mini kit (Qiagen, Germany), with an additional RNAse step, on the automated QIAcube platform (Qiagen, Germany) and was quantified using the Qubit fluorometer 3.0. Genomic libraries were prepared using Nextera XT V2 (Illumina, USA), with a bead-based normalisation, following the manufacturers' guidelines. A total of 40 isolates were multiplexed per sequencing run to provide a depth of coverage >30X. Paired end WGS, was performed on an Illumina MiSeq using the V3 chemistry to generate fragment lengths up to 300bp (600 cycles). The same gDNA extract was also subjected to additional long read sequencing using the Oxford Nanopore MinION device: The MinION library was prepared following the PCR-free rapid-barcoding kit (SQK-RBK004) and sequenced on a FLO-MIN106 (R9.4.1) flow cell, according to the manufacturer's instructions, for 48 hours. Base calling was completed using Guppy v2.1.3 (Oxford Nanopore), in MinKNOW v18.12, and the adapters were removed using Porechop v0.2.4. NanoPlot v1.19.0 was used to determine the number and length of reads generated on the MinION device. Long reads from MinION sequencing were assembled against the short reads from Illumina sequencing using Unicycler v0.4.7 [490] and associated dependencies.

**Table 2. 2. Sequence statistics generated for 14 UK GBS isolates sequenced by both Miseq and Nanopore technology.**

| Strains | Miseq | | | Nanopore MinION | | |
|---|---|---|---|---|---|---|
| | Genome length | Number of reads (mean) | Coverage (mean) | Length of genome | Number of reads (mean) | Coverage (mean) |
| PHEGBS0662 | 2179091 | 1089545.5 | 101.5 | 2185518 | 73907 | 157.9 |
| PHEGBS0206 | 2166025 | 1083012.5 | 164.8 | 2244608 | 12651 | 334.4 |
| 200682 | 2235218 | 1217950 | 127.5 | 2065099 | 36100 | 74.9 |
| PHEGBS0598 | 2191974 | 1095987 | 88.5 | 2173179 | 13063 | 0.1184 |
| PHEGBS0071 | 2115399 | 1057699.5 | 65.9 | 2209339 | 12892 | 0.0839 |
| PHEGBS0738 | 2126457 | 1063228.5 | 125.6 | 2128485 | 68794 | 174.1 |
| PHEGBS0608 | 2193800 | 1096900 | 188.3 | 2160381 | 104564 | 200.9 |
| PHEGBS0595 | 2109564 | 1054782 | 110.2 | 2180389 | 63868 | 138.6 |
| PHEGBS0139 | 2085983 | 1042991.5 | 45 | 2024578 | 31764 | 31.3227 |
| PHEGBS0207 | 2102437 | 1051218.5 | 258.3 | 2192944 | 148464 | 321.3 |
| PHEGBS0288 | 2159710 | 1079855 | 10 | 2223704 | 20875 | 18.9506 |
| PHEGBS0577 | 2156314 | 1078157 | 12.6 | 2227156 | 29,50 | 27.6303 |
| PHEGBS0559 | 2104667 | 1052333.5 | 113.2 | 2120175 | 59659 | 135 |
| 100414 | 2359530 | 1229972 | 79.2 | 2217174 | 17411 | 32.4855 |

### 2.5.4 Genomic Assembly and Annotation

#### 2.5.4.1 UK and Brazilian GBS isolates

Low quality reads and reads <100 bp were trimmed using Trimmomatic v0.35 [491]. Quality control was done on raw sequencing data using FastQC v0.11.2 [492]. Low quality reads with per base sequence quality <28 was discarded from the analysis. Good quality reads were selected based on per sequence base quality (33-40) and per sequence quality Phred scores (36-39), duplication level (<10%), per base N content (NIL) and per base GC content (~37%). High-quality reads were *de novo* assembled using Spades v3.9.0 [493] using k-mer sizes ranging from 43 to 127 base. Pilon v1.22 [494] and flash v1.2.11 [495] were used to improve draft genome assemblies by correcting bases, fixing miss-assemblies and filling gaps. Quast v2.1 [496] was used to summarize statistics for genome assemblies and results were used to discard assemblies with poor quality, based on three metrics including total genome length, total number of contigs and N50 generated. Overall, the 194 UK and 46 Brazilian GBS assemblies had among (13 to 1.14 Kb) contigs, (1.73 Mbp and 7.48 Mbp) total genome length and (8.0 Kbp and 1.17

Mbp) N50. Mean and range values of essential parameters observed for 194 UK and 46 Brazil GBS draft genome assemblies are listed in **Table 2.3**. Based on these parameters, one UK and thirteen Brazilian GBS assemblies could not pass the assembly quality check (**Table 2.4**). Genome annotation was performed using Prokka v1.12 [497] and Geneious R11 (Biomatters ltd. New Zealand) [498].

**Table 2. 3. Quast v2.1 [496] essential assembly metrics observed for 193 UK and 27 Brazilian GBS genomes that passed quality control test.**

| Metrics | Results |
|---|---|
| Closest Reference Sequence from NCBI | *S. agalactiae* 2603V/R (Accession No. NC_004116.1) |
| Mean coverage per genome | ~ 86 (min=22, max=201) |
| Mean genome length | 2.11 Mbp (min=1.73 Mbp, max=2.8 Mbp) |
| Mean no. of contigs | 34 (min=13, max=55) |
| Mean contig length | 2.11 Mbp (min=1.95 Mbp, max =2.7 Mbp) |
| N50 | 1.35 Mbp (min=8.0 Kbp, max= 1.17 Mbp) |
| GC content | 33.4% (min=33%, max= 37%) |

**Table 2. 4. Poor assembled genomes of UK and Brazil GBS isolates that were excluded from the analysis.**

| Assemblies | Country | No. of contigs | Total length | GC (%) | N50 |
|---|---|---|---|---|---|
| BS-S3341 | Brazil | 1283 | 4440487 | 33.96 | 8083 |
| BS-S1975 | Brazil | 967 | 2758505 | 37.22 | 9066 |
| BS-S2506 | Brazil | 340 | 4449014 | 33.71 | 48643 |
| BS-S3132 | Brazil | 247 | 2252086 | 35.12 | 234420 |
| BS-S3007 | Brazil | 149 | 4515245 | 33.54 | 67959 |
| BS-S3008 | Brazil | 37 | 3055286 | 37.24 | 219673 |
| BS-S3315 | Brazil | 44 | 3070116 | 37.18 | 114009 |
| BS-S2895 | Brazil | 28 | 1763390 | 38.41 | 161528 |
| BS-S3133 | Brazil | 27 | 1763919 | 38.42 | 14500 |
| BS-S2361 | Brazil | 25 | 2082441 | 35.38 | 119458 |
| BS-S2041 | Brazil | 18 | 2109430 | 35.46 | 181247 |
| BS-S3267 | Brazil | 17 | 1732733 | 38.38 | 163049 |
| BS-S2293 | Brazil | 15 | 2097195 | 35.38 | 382955 |
| PHEGBS0511 | UK | 1146 | 7480374 | 56.64 | 144101 |

### 2.5.5 Capsular serotyping and Multi-locus Sequence Typing

Capsular serotype was assigned *in silico* using oligonucleotide primers published previously (**Table 2.5**) [50] in Geneious R11 (Biomatters ltd. New Zealand) [498]. Further validation of isolate serotype was performed using BLAST tool in Geneious R11 (Biomatters ltd. New Zealand) [498] and sequence-based serotype allocation using serotype references as follows: serotype Ia (AB028896.2) [499], serotype Ib (AB050723.1) [500], serotype II (LT671985.1) [45], serotype III (AF163833.1) [401], serotype IV (AF355776.1) [500], serotype V (AF349539.1) [500], serotype VI (AF337958.1) [500], serotype IX (LT671992.1) [45]. MLST was assigned using SRST2 v0.2 [501] and double checked using online software MLST v2.0 available at Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/MLST/) [502]. Alleles and STs not previously described were deposited in the *S. agalactiae* MLST database (https://pubmlst.org/sagalactiae/). Pyloviz v2 [503] and its goeBURST algorithm was used to establish relationships between STs using "3" as minimum SLV count for subgroup definition and "6" as minimum number of identical loci for group definition.

**Table 2. 5. Oligonucleotide primers used for GBS serotyping using Geneious R11 (Biomatters ltd. New Zealand) [498].** Reproduced from [50].

| Primer name | Sequence (5′ to 3′) |
|---|---|
| Ia-F | GGTCAGACTGGATTAATGGTATGC |
| Ia-R | GTAGAAATAGCCTATATACGTTGAATGC |
| Ib-F | TAAACGAGAATGGAATATCACAAACC |
| Ib-R | GAATTAACTTCAATCCCTAAACAATATCG |
| II-F | GCTTCAGTAAGTATTGTAAGACGATAG |
| II-R | TTCTCTAGGAAATCAAATAATTCTATAGGG |
| III-F | TCCGTACTACAACAGACTCATCC |
| III-R | AGTAACCGTCCATACATTCTATAAGC |
| IV-F | GGTGGTAATCCTAAGAGTGAACTGT |
| IV-R | CCTCCCCAATTTCGTCCATAATGGT |
| V-F | GAGGCCAATCAGTTGCACGTAA |
| V-R | AACCTTCTCCTTCACACTAATCCT |
| VI-F | GGACTTGAGATGGCAGAAGGTGAA |
| VI-R | CTGTCGGACTATCCTGATGAATCTC |
| VII-F | CCTGGAGAGAACAATGTCCAGAT |
| VII-R | GCTGGTCGTGATTTCTACACA |
| VIII-F | AGGTCAACCACTATATAGCGA |
| VIII-R | TCTTCAAATTCCGCTGACTT |

| | |
|---|---|
| IX-F | CTGTAATTGGAGGAATGTGGATCG |
| IX-R | AATCATCTTCATAATTTATCTCCCATT |

F- forward and R – Reverse

### 2.5.6 Resistome

The presence of Antimicrobial Resistance Genes (ARGs) were determined using ResFinder v2.1 [504] and penicillin-binding protein transpeptidase amino acid sequence types (PBPs 1a, 2b and 2x ) as previously described by Metcalf *et al* 2017 [505]. Geneious R11 (Biomatters ltd. New Zealand) [498] and SRST2 v0.2 [501] were further used to validate presence of resistance genes in each isolate with minimum gene coverage of 90%. A manual search for known mutations in the quinolone resistance determining regions (QRDRs) of gyrase and topoisomerase IV (*gyr*A*, gyr*B*, par*C*,* and *par*E) was performed using ClustalW v1.2.4 [506] by comparing these genes to the equivalent genes present in a reference *S. agalactiae* NEM316 strain (Accession No. AL732656.1) where the mutations and resistance in QRDR regions were defined earlier [237].

### 2.5.7 Pilus Island, *hvg*A, and Surface protein gene profiling

Detection of potential markers for protein vaccine development including alpha and alpha-like surface protein genes (*alp*1, *alp*2, *alp*3, *bca*, and *rib*) [78, 313], and pilus island genes (PI-1, PI-2a and PI-2b) along with other virulence genes i.e. Cβ (*bca*), bacterial adhesins (*bib*A and *hvg*A) and serine rich repeats (*srr*-1 and *srr*-2) was performed using SRST2 v0.2 [501]. Some of the alpha like family protein members i.e. AlphaC (*bca* and *rib*) are attractive candidates for the development of protein based GBS vaccine [507]. Currently, MinervaX is developing a novel innovative, adjuvanted protein-only vaccine based on bacterial antigens derived from the family of alpha-like surface proteins of Group B Streptococcus [507]. The antigens are based on the N-terminal domains of the most prevalent alpha-like protein serotypes (AlpCN, RibN, Alp1N and Alp2/3N), covering close to 100% of clinical GBS isolates [507], therefore in this study appealing surface proteins for vaccine development were screened in all GBS datasets. The results of SRST2 were validated through Geneious R11 (Biomatters ltd. New

Zealand) [498]using previously published oligonucleotide primers for all target genes except *cba* for which no primer was found in the literature (**Table 2.6**).

**Table 2. 6. Oligonucleotide primers used for the detection of virulence genes in UK and Brazil GBS genomes.**

| Primers | Sequence | References |
|---|---|---|
| *alp*1-R | CCAGATACATTTTTTACTAAAGCGG | [21] |
| *alp*2-R | CTTCATCTGTTGACTTATCTGGATAG | [22] |
| *alp*3-R | CTTTTGGTTCGTTGCTATCCTTAAG | |
| *bca*-F | TAACAGTTATGATACTTCACAGAC | [23] |
| *bca*-R | ACGACTTTCTTCCGTCCACTTAGG | |
| *rib*-F | CAGATGCCGATAAGA | |
| *rib*-R | TACGCGGATCGACAA | |
| PI-1-UP | GGTCGTCGATGCTCTGGATTC | [22] |
| PI-1-DN | GTTGCCCAGTAACAGCTTCTCC | |
| PI-2a-UP | CTATGACACTAATGGTAGAAC | |
| PI-2a-DN | CACCTGCAATAGACATCATAG | |
| PI-2b-UP | ACACGACTATGCCTCCTCATG | |
| PI-2b-DN | TCTCCTACTGGAATAATGACAG | |
| *bibA*-F | AATCGAAAACAACGTTGGAAAG | [23] |
| *bibA*-R | AAACCAGGCTTCATCAGTCATT | |
| *hvg*A-F | ATACAAATTCTGCTGACTACCG | |
| *hvg*A-R | TTAAATCCTTCCTGACCATTCC | |
| *srr-1*-F | AGTGTCTGATACTGAAATGTTAGGTA | [24] |
| *srr-1*-R | TCGCATCAGGACTTGGGAATCTA | |
| *srr-2*-F | TCACGCAAAGTTCGAGTTAAAA | |
| *srr-2*-R | AGATTTAGTAGCTCCTAA | |

### 2.5.8 Mobile genetic element (MGEs) analysis

The first step was to search some well-known integration sites, like well-conserved housekeeping genes such as *rpl*L - L7/L12 ribosomal protein (protein_id: AWZ30204.1), *rps*I - S9 ribosomal protein (protein_id:.AKI94631.1), *rip*L - S9 ribosomal protein (protein_id:.AQY24521.1), *rpl*M - L13 ribosomal protein (protein_id:.AKI94630.1) and *rum*A - 23S rRNA (uracil-5-)-methyltransferase (protein_id: AQY23883.1) in respective genomes using Geneious 11 [498]. The second step was to search DRs on both sides of the putative ICEs using find repeats tool in Geneious 11 [498] with a

minimum repeat length of 15 bps. Blastn (https://blast.ncbi.nlm.nih.gov/) was run as the final step on the query nucleotide sequences present between DRs of the putative ICEs in each genome. If the DRs were not found, then the whole stretch of DNA present downstream the integration sites was used as a query. Insertion sites of Tn*916*/Tn*916*-like elements were identified relative to the completely sequenced SS1 (NZ_ CP010867.1) strain carrying Tn*916*, CZ-NI-013 strain carrying Tn*3872* [24] and DK-NI-005 strain carrying Tn*5801*, respectively [1]. Artemis [337] was used to colour the detected signature proteins of an ICE such as tyrosine/serine integrases/recombinases, relaxases, plasmid mobilization protein genes (*mob*A/*mob*L), conjugal transfer protein genes (*vir*B4, *vir*D4, *tra*L and *trb*L) and insertion sequences. Pairwise BLASTn alignment was performed and visualized using Easyfig v2.2.2 [336].

### 2.5.9 Phylogenetic Analysis

Raw reads were mapped to the reference genome and Single Nucleotide Polymorphisms (SNPs) were called using snippy v3.2 (https://github.com/tseemann/snippy). Reference sequences were selected based on the most common ST and serotype observed within the CCs and were as follows: for the analysis of CC1 isolates, reference sequence SS1 (NZ_CP010867.1) was used; for the CC17 – reference sequence COH1 (NZ_HG939456.1), for the CC19 – reference sequence SG-M25 (CP021867), for the analysis of CC8/CC10 – reference sequence Sag37 (NZ_CP019978.1) and for CC23 – reference sequence FDAARGO_512 (NZ_CP033822.1). Phylogenetic trees were build based on core SNPs alignments using RAxML v8.2.8 [508] and phylogenetic trees were visualized using the Interactive tree of life (iTOL) (https://itol.embl.de/) [509].

### 2.5.10 Recombination analysis to detect potential capsular switching events

Gubbins [510] was used to highlight recombination hotspots using the core SNP alignment file generated by snippy v3.2 (https://github.com/tseemann/snippy) and the BLAST ring image generator (BRIG) [511] was used to visualize these recombination hotspots.

### 2.5.11 Pan-genome construction and genome wide association analysis

Roary v3.13 [512] was used to generate pan-genome for GBS genomes clustered in main CCs (CC1, CC17, CC8/CC10, CC19 and CC23). Parameters for each run were: 95% of minimum Blastp identity; MLC inflation value 1.5; with 99% as the percentage cut-off in which a gene must be present to be considered as core. Scoary v1.6.16 [513] was used to establish which genes were typical of each CC via a Pan-genome Wide Association Study (pan-GWAS). The CC of each isolate was depicted as a discrete phenotype, e.g., belonging to CC17 or not, and defined as "positive" or "negative" respectively with the Scoary algorithm evaluating which gene feature is statistically associated with a particular CC [513]. The cut-off for a significant association was a p-value lower than 1e-10 and a sensitivity and specificity greater than 90 percent. The gene is CC specific if present in all or at least 90% of isolates belonging to that CC but completely absent in all other CC isolates.

# Chapter 3:

## Identifying large scale recombination and capsular shifting events in *Streptococcus agalactiae* causing disease in adults in the United Kingdom from January 2014 - May 2015

## 3.1 Summary

Recombination is a key contributor to bacterial clonal diversity and can lead to capsule switching in GBS, potentially combining virulence factors with altered capsular loci outside vaccine coverage. Whole genome analysis of 179 invasive clinical GBS strains (iGBS) isolated from adults submitted to the UK national reference laboratory was used to determine capsule type, MLST, presence of virulence factors, antimicrobial resistance genes, phylogeny, and genetic recombination. Capsular serotypes III (26.8%), Ia (26.2%) and V (15%) were most common in adults, with different gender/age distribution. Most isolates (n=161) grouped to 5 clonal complexes: CC1, CC8/CC10, CC17, CC19 and CC23 with common associations between specific serotypes and virulence genes. Serotype V/*alp3*, serotype II and III/*bca*+*cba* and serotype Ia/*bib*A predominantly clustered in CC1, CC8/CC10 and CC23, respectively, whereas serotype III/*rib* clustered in CC17 and CC19. All isolates carried at least one pilus island; mostly PI-1+PI-2a, but PI-1+PI-2b exclusively co-located to CC17 isolates. Large recombination events mediating capsular switching from CC1-serotype V to serotype Ib (n=2), II (n=2) and VI (n=2) were found; with CC19-serotype III to serotype V (n=5); with CC23 serotype Ia to serotype V (n=2) and a hypervirulent CC17 (*hvg*A positive) serotype III to serotype IV (n=1). Extensive genomic investigation of GBS isolates from adults in the UK found a high number of capsular serotype-CC mismatches (14/179 iGBS) with a concerning recombination of hypervirulent *hvg*A core genome expressing a non-vaccine covered serotype IV capsule. This highlights the need for improved surveillance to inform vaccine development. The high genetic diversity of disease-causing isolates and multiple recombination events reported in this study, highlight the need for routine surveillance of the circulating disease-causing GBS strains. This information is crucial to better understand global spread of GBS serotypes and genotypes and will form the baseline information for any future GBS vaccine research in the UK and worldwide.

## 3.2 Introduction

GBS is a Gram-positive encapsulated pathobiont bacterium of the human genitourinary and gastrointestinal tract. GBS is well established as a leading cause of bacterial sepsis and meningitis in neonates [514] and is increasingly associated with invasive infections in adults [115, 515, 516]. In the UK, cases of iGBS in the adult population, particularly the elderly and women of childbearing age, have increased over the last fifteen years; with previous studies reporting annual incidences from 0.92 to 2.39/100,000 population during 1991-2010 [115], increasing to 3.48/100,000 population during 2015–2016 [517].

Bacterial recombination involves the exchange of genetic material either between multiple chromosomes or between different regions of the same chromosome [518] and is believed to be a major factor explaining the bacterial evolution [519]. This process is generally mediated by three means i) transformation - uptake of exogenous DNA from the surrounding environment, ii) transduction - virus-mediated transfer of DNA between bacteria, and iii) conjugation - the transfer of DNA from one bacterium to another via cell-to-cell contact [520]. Two major types of recombination are i) homologous recombination in which transfer is mediated if some degree of homology (sequence identity) is present between recipient and donor genetic material [521]. Various cases of nonhomologous (homeologous) recombination (other type of recombination) do exist, in which exchange between slightly diverged DNA sequence takes place [522]. Homologous recombination is an evolutionarily fundamental process that is essential for genome plasticity and plays a pivotal role in maintaining genomic stability by repairing complex DNA damage such as DNA double-stranded breaks and interstrand cross-links while facilitating genetic diversity [521, 523]. However, if the homologous recombination was prevalent, it would expect that recombination between these different repetitive elements would produce adverse gross rearrangements, loss of genetic information (through gene conversion and deletion) and general genetic instability [524-526]. The rate of homologous recombination varies greatly: in some species it appears to be rare and leads to the evolution of distinct clonal lineages, whereas in others these localised recombinational imports arise much more frequently than mutations [527]. In GBS, a very low rate of

recombination was observed for CC17 lineage involving 3% of the genome compared to a high recombination rate observed for CC1 comprising 47% of the genome [1].

GBS is sub-divided into ten serotypes (Ia, Ib, II-IX) based on capsular polysaccharides. However, only five serotypes, Ia, Ib, II, III, and V, account for 93-99% of cases of neonatal and adult infections globally [265]; with serotypes V and Ia accounting for 58% of elderly iGBS and serotype III dominating in neonatal iGBS cases [265]. Currently, there are three main GBS vaccines in development: multi-capsular polysaccharide vaccines: trivalent, covering serotypes Ia, Ib and III [528]; hexavalent, covering serotypes Ia, Ib, II, III, IV and V [264, 529]; and a multivalent adjuvanted protein vaccines (NCT03807245) based on the N-terminal domains of the most prevalent alpha-like protein serotypes (AlpCN, RibN, Alp1N and Alp2/3N), covering close to 100% of clinical GBS isolates [46]. With these GBS vaccines on the horizon, questions arise to predict if implementation of GBS polysaccharide-based vaccines will drive serotype replacement, or capsular switching, which has been reported for other vaccine preventable bacterial pathogens such as *S. pneumoniae* [530, 531] and leads to the emergence of non-vaccine serotypes. Increasing numbers of GBS genomic epidemiology studies has suggested that GBS could undergo capsular switching through horizontal transfer of the capsular locus [77, 338]. To date, higher frequency of capsular switching events have been recorded in ST1 GBS lineage [77, 338], which is dominated by serotype V and commonly found in non-pregnant adult iGBS cases, and recombination of capsular sites for serotypes Ib, II and VI have been reported for this lineage [76, 389]. Later, study from Toronto, Canada described multiple episodes of capsular switching events reporting conversion of serotype IV to II ST196, serotype III to V ST19, serotype Ia to V ST7 colonizing GBS strain [338]. Capsular switching within members of CC17 rather than ST17 lineage where serotype III strains switched to express serotype IV capsule have also been documented in France [269] and Ireland [532] CC17 iGBS strains, respectively. Although most of the reported capsular switching events described is the change of serotype to a vaccine covered serotype, there is potential for the less common serotypes to become established in iGBS. Therefore, structured longitudinal surveillance of carriage and clinical GBS strains is important to help underpin the prevalence and conditions required for

capsular switching events in the GBS population, and implications of such events on clinical disease incidence and the GBS vaccine implementation policy. In the UK, a detailed genomic surveillance of GBS isolates causing early and late-onset GBS disease in infants has been recently carried out which uncovered clusters of infection [533], however, detailed information on genomic features of adult GBS isolates in the UK is lacking. Here, we used whole genome sequencing to characterise disease-causing GBS strains from the adult population in the UK and report putative capsular switching events that occurred in this GBS population.

## 3.3 Results

### 3.3.1 Population characteristics and serotype distribution

Sterile site isolates from adults are, where possible, routinely submitted to the Public Health England (PHE) reference laboratory as part of national surveillance. Between January 2014 and May 2015, isolates were mainly submitted from laboratories in England with a small proportion of submissions from elsewhere in the UK. Isolates received are from patients of all ages, with the majority from neonates (estimated >95%) and women of child-bearing age, with an estimated total adult isolate submission proportion of less than 11% of all invasive cases. [480, 481]. Information about pregnant/non-pregnant and postpartum status was not available for the women GBS was collected from.

A total 179 iGBS isolates (178 from blood, and 1 from CSF) were available for analysis (**Table 3.1**). In this dataset, age of the patients ranged from 20 to 100 years (median age 59 years) with 64.2% of GBS isolates from female adult patients (n=115/179). As expected, due to the potential for post-partum infection in females a lower median age-group was observed for females (median age 39 years old, IQR: 40; 32.5-72.5) (**Figure 3.1a**), compared to males (median age 73.5 years old, IQR: 19; 62-81). More GBS were isolated from younger women, 18 to 44 years old, (n=65/115) relative to women over 44 years of age (n=50/115) (**Figure 3.1b**), whereas only 7/64 were isolated from male patients 18-44 years old (**Figure 3.1c**). The bias towards isolates from females (56.5%, 65/115) of childbearing age

(18-44 years) in this dataset was not deliberate and the proportions are not representative of the gender or age prevalence of invasive GBS documented between 2014 and 2015 for the UK. The samples in this thesis simply reflect those that were viable and available at PHE for whole genome sequencing, no selection of isolates was performed: analysis was performed on ALL viable isolates available from adults for sequencing during the collection period.

In total, eight serotypes were identified: III (26.8%, 48/179), Ia (26.2%, 47/179), V (15%, 27/179), II (13.9%, 25/179), Ib (11.7%, 21/179), IV (3.3%, 6/179), VI (1.1%, 2/179) and IX (1.6%, 3/179). Serotype III was associated with female patients ($X^2(1, N=179) =11.07$, $p=0.0008$) whilst serotype V was more commonly identified in male patients ($X^2(1, N=179) =5.3$, $p=0.02$). Serotypes III and Ia were more common among female patients, especially in the younger age group (18-44 years old), with the exception for the 45-64 years age group, where serotypes II and Ib pre-dominated, though this group contains a small number of isolates for comparison to other groups (**Figure 3.1b**). In males, serotypes Ia and V were more prevalent compared to other serotypes (**Figure 3.1c**).

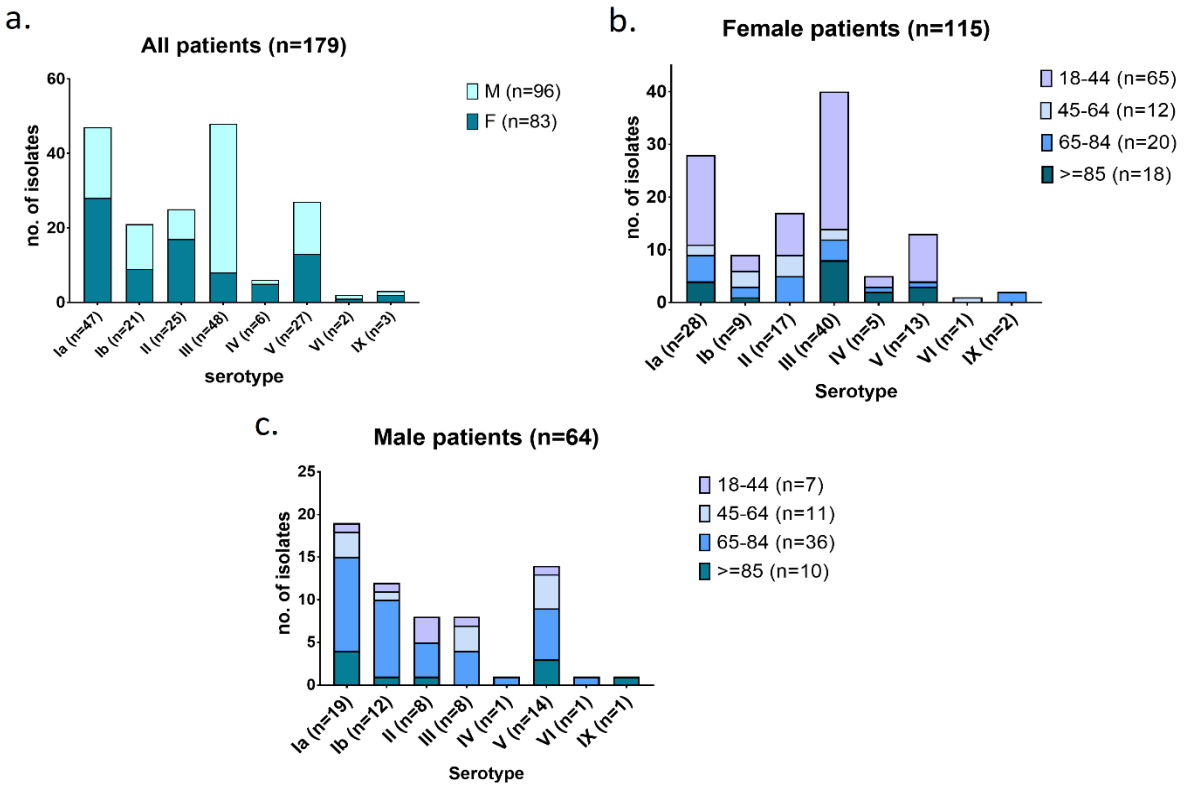



**Figure 3. 1. Summary of demographic information available for 179/193 (only the invasive) GBS isolates submitted to Public Health England between January 2014 and May 2015. (a)** Overall number of GBS serotypes observed in the dataset segregated by sex; **(b)** graphical distribution of GBS serotypes identified in female patients, segregated by age group; **(c)** graphical distribution of GBS serotypes identified in male patients, segregated by age group.

### 3.3.2 Genotypes and virulence genes prevalent in disease-causing GBS in the UK

Of the 42 sequence types (STs) found, 5 predominated including ST23 (n=37), ST17 (n=25), ST1 (n=23) and ST19 (n=15) and ST12 (n=13). Additionally, 13 novel STs were identified (**Table 3.1**). The STs based on single (SLV) were grouped into five main clonal complexes: CC23 (23.4%, 42/179), CC1 (17.8%, 32/179), CC8/CC10 (16.2%, 29/179), CC17 (16.7%, 30/179) and CC19 (15.6%, 28/179) (**Table 3.2**).

**Table 3. 1. Summary of characteristics of 179 iGBS whole genome sequences used in this study.** During 17 months of surveillance by PHE between January 2014 and May 2015, a total of 179 iGBS isolates causing disease in adults were collected and whole genome sequenced. Isolates were collected from blood (n=178) and CSF (n=1). Novel STs identified in this study are marked in bold. * Uncertainty detected for reporting ST for this isolate. F- female, M – male, n/a – information not available.

| Sample ID | ERA number | Sex | Age (yrs) | Serotype | ST | Specimen |
|---|---|---|---|---|---|---|
| PHEGBS0041 | ERR1742042 | M | 74 | V | ST1 | Blood |
| PHEGBS0042 | ERR1742087 | F | 69 | IX | ST130 | Blood |
| PHEGBS0044 | ERR1742051 | F | 34 | Ia | ST23 | Blood |
| PHEGBS0047 | ERR1741500 | M | 55 | Ia | ST24 | Blood |
| PHEGBS0048 | ERR1741806 | M | 84 | Ia | ST23 | Blood |
| PHEGBS0049 | ERR1741385 | M | 85 | Ia | ST144 | Blood |
| PHEGBS0052 | ERR1741861 | F | 99 | III | ST17 | Blood |
| PHEGBS0054 | ERR1742070 | M | 65 | III | ST283 | Blood |
| PHEGBS0060 | ERR1741459 | F | 88 | Ia | ST23 | Blood |
| PHEGBS0061 | ERR1741474 | F | 47 | Ib | ST8 | Blood |
| PHEGBS0066 | ERR1741366 | F | 76 | III | ST17 | Blood |
| PHEGBS0067 | ERR1741677 | M | 44 | Ia | ST23 | Blood |
| PHEGBS0068 | ERR1741595 | F | 36 | Ib | ST12 | Blood |
| PHEGBS0070 | ERR1742116 | F | 80 | Ia | ST23 | Blood |
| PHEGBS0071 | ERR1742139 | M | 63 | III | ST19 | Blood |
| PHEGBS0072 | ERR1741514 | M | 43 | Ib | ST15 | Blood |
| PHEGBS0080 | ERR1741859 | M | 73 | II | ST22 | Blood |
| PHEGBS0081 | ERR1741497 | F | 86 | V | ST1 | Blood |
| PHEGBS0082 | ERR1741501 | F | 92 | V | ST1 | Blood |
| PHEGBS0084 | ERR1741695 | F | 93 | IV | ST459 | Blood |
| PHEGBS0090 | ERR1742102 | M | 48 | V | **ST1314** | CSF |
| PHEGBS0091 | ERR1741407 | F | 60 | II | ST12 | Blood |
| PHEGBS0092 | ERR1741722 | M | 73 | II | ST12 | Blood |
| PHEGBS0095 | ERR1741918 | F | 58 | Ia | ST23 | Blood |
| PHEGBS0097 | ERR1741888 | F | 92 | Ib | ST104 | Blood |
| PHEGBS0098 | ERR1741573 | M | 23 | V | ST1 | Blood |
| PHEGBS0100 | ERR1741744 | F | 49 | II | ST3 | Blood |
| PHEGBS0106 | ERR1742039 | F | 34 | Ia | ST23 | Blood |
| PHEGBS0117 | ERR1741635 | F | 23 | III | ST17 | Blood |
| PHEGBS0121 | ERR1741801 | F | 31 | Ia | ST23 | Blood |
| PHEGBS0122 | ERR1741539 | M | 86 | Ia | ST23 | Blood |

| | | | | | | |
|---|---|---|---|---|---|---|
| PHEGBS0123 | ERR1742074 | F | 36 | Ia | ST23 | Blood |
| PHEGBS0127 | ERR1741367 | M | 86 | V | ST1 | Blood |
| PHEGBS0128 | ERR1742140 | M | 80 | V | ST1 | Blood |
| PHEGBS0132 | ERR1741466 | F | 51 | III | ST529 | Blood |
| PHEGBS0134 | ERR1741692 | M | 54 | Ia | **ST1218** | Blood |
| PHEGBS0135 | ERR1741673 | M | 77 | Ib | ST8 | Blood |
| PHEGBS0139 | ERR1741954 | M | 63 | V | ST19 | Blood |
| PHEGBS0144 | ERR1741588 | F | 28 | III | ST17 | Blood |
| PHEGBS0145 | ERR1741818 | M | 72 | Ib | ST12 | Blood |
| PHEGBS0151 | ERR1741944 | M | 28 | II | ST10 | Blood |
| PHEGBS0152 | ERR1742041 | M | 69 | Ia | ST24 | Blood |
| PHEGBS0153 | ERR1741822 | F | 70 | II | ST104 | Blood |
| PHEGBS0154 | ERR1741575 | M | 77 | Ib | ST9 | Blood |
| PHEGBS0156 | ERR1741857 | F | 63 | III | ST19 | Blood |
| PHEGBS0164 | ERR1741749 | M | 59 | V | ST1 | Blood |
| PHEGBS0170 | ERR1742029 | F | 88 | Ia | ST23 | Blood |
| PHEGBS0171 | ERR1741963 | F | 28 | II | ST12 | Blood |
| PHEGBS0176 | ERR1742030 | F | 32 | III | ST23 | Blood |
| PHEGBS0188 | ERR1741906 | F | 34 | Ia | ST24 | Blood |
| PHEGBS0193 | ERR1742119 | F | 35 | Ia | ST23 | Blood |
| PHEGBS0194 | ERR1741932 | M | 87 | Ia | ST7 | Blood |
| 200682 | ERR3531626 | F | 35 | III | ST27 | Blood |
| 200683 | ERR3531633 | M | 76 | Ib | ST8 | Blood |
| 200684 | ERR3531637 | M | 71 | Ib | **ST1215** | Blood |
| PHEGBS0206 | ERR1741687 | M | 76 | IV | ST196 | Blood |
| PHEGBS0207 | ERR1741997 | F | 85 | III | ST17 | Blood |
| 200690 | ERR3531636 | F | 91 | V | ST1 | Blood |
| 200693 | ERR3531627 | M | 76 | III | **ST1219** | Blood |
| 200694 | ERR3531635 | M | 77 | III | ST17 | Blood |
| 141439 | ERR3531632 | M | 81 | Ib | ST10 | Blood |
| 200696 | ERR3531628 | F | 23 | III | ST550 | Blood |
| 200697 | ERR3531629 | F | 42 | III | ST17 | Blood |
| PHEGBS0219 | ERR1741590 | F | 100 | III | ST550 | Blood |
| 200698 | ERR3531634 | M | 49 | Ib | ST8 | Blood |
| 200702 | ERR3531630 | M | 59 | Ia | ST23 | Blood |
| PHEGBS0230 | ERR1741957 | F | 86 | Ia | ST23 | Blood |
| 200706 | ERR3531625 | F | 48 | Ib | ST12 | Blood |

| | | | | | | |
|---|---|---|---|---|---|---|
| 200710 | ERR3589625 | F | 32 | III | ST17 | Blood |
| PHEGBS0243 | ERR1741633 | F | 33 | III | ST17 | Blood |
| PHEGBS0246 | ERR1741679 | M | 81 | Ia | ST23 | Blood |
| PHEGBS0248 | ERR1742038 | F | 34 | Ia | ST23 | Blood |
| PHEGBS0252 | ERR1741658 | F | 32 | III | ST19 | Blood |
| PHEGBS0265 | ERR1741375 | F | 34 | Ia | ST23 | Blood |
| PHEGBS0266 | ERR1741584 | F | 32 | IV | ST136 | Blood |
| PHEGBS0270 | ERR1742082 | M | 81 | Ia | ST23 | Blood |
| PHEGBS0275 | ERR1742034 | F | 31 | III | ST17 | Blood |
| PHEGBS0283 | ERR1741523 | F | 33 | Ia | ST23 | Blood |
| PHEGBS0286 | ERR1741880 | F | 39 | III | ST17 | Blood |
| PHEGBS0287 | ERR1742120 | M | 88 | Ia | ST23 | Blood |
| PHEGBS0288 | ERR1741473 | F | 88 | III | ST17 | Blood |
| PHEGBS0295 | ERR1741889 | F | 38 | Ia | ST23 | Blood |
| PHEGBS0296 | ERR1741462 | M | 49 | III | ST19 | Blood |
| PHEGBS0300 | ERR1741553 | F | 27 | Ia | ST23 | Blood |
| PHEGBS0306 | ERR1741485 | F | 27 | V | ST1 | Blood |
| PHEGBS0308 | ERR1741488 | F | 80 | II | ST12 | Blood |
| PHEGBS0318 | ERR1741792 | F | 89 | III | ST17 | Blood |
| PHEGBS0320 | ERR1741605 | F | 72 | Ia | ST7 | Blood |
| PHEGBS0336 | ERR1741747 | M | 65 | Ia | ST23 | Blood |
| PHEGBS0337 | ERR1741620 | M | 91 | V | ST498 | Blood |
| PHEGBS0355 | ERR1741387 | F | 24 | III | ST17 | Blood |
| PHEGBS0359 | ERR1741657 | F | 91 | III | ST19 | Blood |
| PHEGBS0360 | ERR1742032 | M | 78 | V | ST1 | Blood |
| PHEGBS0367 | ERR1742111 | M | 71 | Ia | ST23 | Blood |
| PHEGBS0368 | ERR1742016 | F | 43 | V | ST1 | Blood |
| PHEGBS0372 | ERR1741524 | F | 30 | III | ST19 | Blood |
| PHEGBS0373 | ERR1741478 | M | 67 | V | ST1 | Blood |
| PHEGBS0377 | ERR1741496 | F | 31 | III | ST19 | Blood |
| PHEGBS0378 | ERR1741669 | F | 30 | V | ST1 | Blood |
| PHEGBS0380 | ERR1741867 | F | 30 | III | ST17 | Blood |
| PHEGBS0382 | ERR1741847 | M | 67 | Ib | ST8 | Blood |
| PHEGBS0383 | ERR1741449 | F | 27 | III | ST17 | Blood |
| PHEGBS0389 | ERR1741939 | F | 20 | III | ST17 | Blood |
| PHEGBS0390 | ERR1741684 | M | 40 | II | ST12 | Blood |

| | | | | | | |
|---|---|---|---|---|---|---|
| PHEGBS0393 | ERR1741475 | F | 70 | Ib | ST12 | Blood |
| PHEGBS0394 | ERR1741769 | F | 48 | II | ST9 | Blood |
| PHEGBS0398 | ERR1741711 | M | 85 | II | **ST1213** | Blood |
| PHEGBS0399 | ERR1741537 | F | 74 | II | ST28 | Blood |
| PHEGBS0401 | ERR1741671 | M | 66 | Ia | ST23 | Blood |
| PHEGBS0407 | ERR1741602 | F | 63 | Ia | ST23 | Blood |
| PHEGBS0408 | ERR1741526 | F | 29 | II | ST28 | Blood |
| PHEGBS0411 | ERR1742019 | F | 40 | Ib | ST8 | Blood |
| PHEGBS0416 | ERR1741402 | F | 35 | III | ST17 | Blood |
| PHEGBS0428 | ERR1741389 | M | 87 | Ib | ST8 | Blood |
| PHEGBS0429 | ERR1741515 | F | 92 | III | ST17 | Blood |
| PHEGBS0446 | ERR1741823 | M | 77 | VI | ST1 | Blood |
| PHEGBS0447 | ERR1741785 | F | 35 | Ia | ST1065 | Blood |
| PHEGBS0448 | ERR1741564 | F | 83 | IV | ST196 | Blood |
| PHEGBS0450 | ERR1741493 | F | 74 | Ib | ST1 | Blood |
| PHEGBS0463 | ERR1742031 | F | 21 | IV | ST196 | Blood |
| PHEGBS0464 | ERR1741702 | F | 38 | II | ST28 | Blood |
| PHEGBS0467 | ERR1741981 | F | 32 | V | ST1 | Blood |
| PHEGBS0476 | ERR1741458 | F | 85 | IV | **ST1351** | Blood |
| PHEGBS0480 | ERR1741862 | F | 25 | III | ST17 | Blood |
| PHEGBS0483 | ERR1741662 | F | 34 | III | **ST1212** | Blood |
| PHEGBS0491 | ERR1741535 | F | 82 | II | ST12 | Blood |
| PHEGBS0492 | ERR1741922 | F | 89 | III | ST19 | Blood |
| PHEGBS0493 | ERR1741913 | F | 22 | V | **ST1217** | Blood |
| PHEGBS0501 | ERR1741911 | F | 36 | III | **ST1221** | Blood |
| PHEGBS0503 | ERR1741579 | M | 75 | Ia | ST23 | Blood |
| PHEGBS0509 | ERR1741644 | F | 22 | II | ST12 | Blood |
| PHEGBS0512 | ERR1741958 | F | 32 | Ia | ST23 | Blood |
| PHEGBS0513 | ERR1741701 | F | 44 | III | ST17 | Blood |
| PHEGBS0518 | ERR1741606 | F | 32 | Ia | ST498 | Blood |
| PHEGBS0520 | ERR1741361 | F | 70 | V | ST1 | Blood |
| PHEGBS0524 | ERR1741696 | M | 83 | III | ST23 | Blood |
| PHEGBS0527 | ERR1742059 | F | 26 | Ia | ST23 | Blood |
| PHEGBS0532 | ERR1741494 | F | 29 | II | ST28 | Blood |
| PHEGBS0547 | ERR1741621 | F | 88 | Ia | ST23 | Blood |
| PHEGBS0549 | ERR1741850 | F | 32 | II | ST1 | Blood |

| | | | | | | |
|---|---|---|---|---|---|---|
| PHEGBS0551 | ERR1741634 | M | 65 | II | ST28 | Blood |
| PHEGBS0552 | ERR1741992 | M | 92 | IX | **ST1216** | Blood |
| PHEGBS0554 | ERR1741388 | F | 36 | II | ST28 | Blood |
| PHEGBS0555 | ERR1741852 | F | 20 | Ib | ST8 | Blood |
| PHEGBS0556 | ERR1741948 | M | 65 | II | ST28 | Blood |
| PHEGBS0559 | ERR1741492 | F | 34 | III | ST17 | Blood |
| PHEGBS0561 | ERR1741542 | F | 31 | III | ST17 | Blood |
| PHEGBS0566 | ERR1741860 | F | 73 | III | ST19 | Blood |
| PHEGBS0568 | ERR1741647 | F | 49 | II | ST1 | Blood |
| PHEGBS0577 | ERR1741835 | F | 88 | III | ST17 | Blood |
| PHEGBS0581 | ERR1741660 | M | 41 | II | ST12 | Blood |
| PHEGBS0589 | ERR1741794 | M | 73 | Ia | ST23 | Blood |
| PHEGBS0592 | ERR1741448 | F | 73 | IX | ST130 | Blood |
| PHEGBS0593 | ERR1741754 | M | 42 | III | ST19* | Blood |
| PHEGBS0595 | ERR1741483 | M | 74 | V | ST19 | Blood |
| PHEGBS0598 | ERR1741680 | F | 66 | III | **ST1316** | Blood |
| PHEGBS0599 | ERR1741659 | M | 88 | V | ST1 | Blood |
| PHEGBS0608 | ERR1741534 | F | 23 | V | ST19 | Blood |
| PHEGBS0610 | ERR1742130 | F | 29 | III | ST17 | Blood |
| PHEGBS0616 | ERR1742012 | F | 42 | Ia | ST24 | Blood |
| PHEGBS0618 | ERR1741842 | F | 84 | Ia | **ST1214** | Blood |
| PHEGBS0622 | ERR1741630 | M | 78 | Ib | ST8 | Blood |
| PHEGBS0623 | ERR1741456 | F | 67 | Ia | ST23 | Blood |
| PHEGBS0624 | ERR1741580 | F | 38 | II | ST28 | Blood |
| PHEGBS0625 | ERR1742108 | M | 68 | Ia | ST23 | Blood |
| PHEGBS0626 | ERR1741728 | F | 29 | III | ST19 | Blood |
| PHEGBS0627 | ERR1741512 | F | 75 | II | ST2 | Blood |
| PHEGBS0630 | ERR1741525 | M | 69 | V | ST498 | Blood |
| PHEGBS0639 | ERR1741887 | M | 53 | V | ST110 | Blood |
| PHEGBS0643 | ERR1741746 | M | 62 | III | ST19 | Blood |
| PHEGBS0648 | ERR1741752 | F | 38 | V | ST1 | Blood |
| PHEGBS0654 | ERR1741549 | F | 67 | Ia | ST23 | Blood |
| PHEGBS0657 | ERR1741517 | F | 41 | V | ST19 | Blood |
| PHEGBS0658 | ERR1741901 | M | 82 | Ia | ST23 | Blood |
| PHEGBS0662 | ERR1741427 | F | 51 | VI | ST1 | Blood |
| PHEGBS0663 | ERR1741436 | F | 21 | Ia | ST144 | Blood |

| | | | | | | |
|---|---|---|---|---|---|---|
| PHEGBS0667 | ERR1741656 | M | 81 | Ib | ST12 | Blood |
| PHEGBS0670 | ERR1742049 | F | 48 | Ib | ST1 | Blood |
| 100414 | ERR2560245 | F | 40 | V | **ST1350** | Blood |

**Table 3. 2. Sequence types and serotypes of UK adult iGBS isolates belonging to five main CC.**

| CC | ST (n) | Serotype (n) |
|---|---|---|
| **CC23 (42)** | ST23(n=38), ST144(n=2), ST1065(n=1), ST1218(n=1) | Ia(n=40), III(n=2) |
| **CC1 (n=32)** | ST1(n=23), ST2(n=1), ST196(n=3), ST1217(n=1), ST459(n=1), ST136(n=1), ST1350(n=1), ST1314(n=1) | V(n=20), IV(n=5), II(n=3), Ib(n=2), VI(n=2) |
| **CC8/CC10 (n=29)** | ST12(n=13), ST8(n=9), ST15(n=1), ST104(n=2), ST10(n=2), ST9(n=2) | Ib(n=18), II(n=11) |
| **CC17 (n=30)** | ST17(n=25), ST550(n=2), ST1219(n=1), ST1351(n=1), ST1221(n=1) | III(n=29), IV(n=1) |
| **CC19 (n=28)** | ST19(n=17), ST28(n=8), ST27(n=1), ST1316(n=1), ST110(n=1) | II(n=8), III(n=15), V(n=5) |
| **Others (n=18)** | | |
| **CC130/CC1216 (n=3)** | ST1216(n=1), ST130 (n=2) | IX (n=3) |
| **CC22/CC1213 (n=2)** | ST1213 (n=1), ST22 (n=1) | II (n=2) |
| **CC24/CC498 (n=7)** | ST24(n=4), ST498(n=3) | Ia (n=5), V (n=2) |
| **Singletons (n=6)** | ST104(n=2), ST15(n=1), ST1212 (n=1), ST26 (n=1), ST283 (n=1) | II (n=1), Ib (n=2), III (n=2), V (n=1) |

The majority of CC23 strains were serotype Ia (95.2%, 40/42), CC1 isolates serotype V (84.3%, 27/32) and predominantly ST1/*alp*3-positive (n=17). Serotype Ib and serotype II isolates were mainly located within CC8/CC10 with majority of the isolates being ST8 (n=9) and ST12 (n=13) (**Table 3.2**). Most serotype III isolates aligned either with CC19 (15/28 CC19 isolates) or hypervirulent clone CC17 (29/30 CC17 isolates). All CC17 serotype III and one serotype IV isolates contained the *srr*-2, *rib* and *hvg*A genes (**Figure 3.2**).

All 179 iGBS isolates harboured at least one pilus islet (PI), the PI-1+PI-2a was the most frequent combination observed (44.6%, 80/179), with only CC17 being dominated by the presence of PI-1+PI-2b gene combination (**Figure 3.2**). The (93.3%, 28/30) of *bca* genes found in this GBS collection were

found in isolates belonging to CC8/CC10. The *rib* gene was exclusively expressed in only CC17 (n=30/30) and CC19 (n=23/29) isolates (**Figure 3.2**). The *srr*-1 genes were disseminated among all CC isolates except CC17 isolates that exclusively carried *srr*-2 gene (**Figure 3.2**). Singletons that could not be assigned to a CC (**Table 3.2**) made up 3.3% (6/179) of the total iGBS isolates.

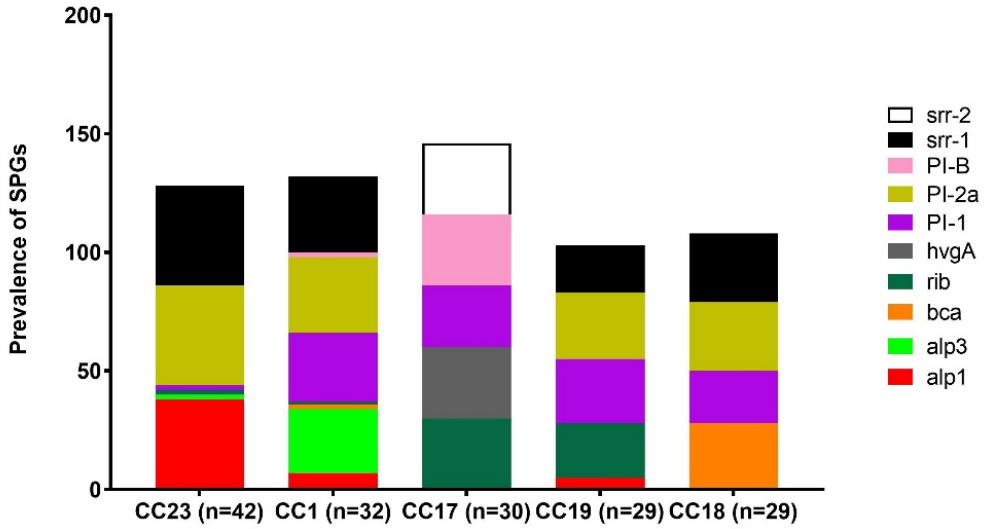



**Figure 3. 2. Distribution of surface proteins genes (SPGs) in iGBS isolates belong to five main CCs.** The prevalence of surface protein genes identified in 179 GBS isolates belonging to five main CCs are represented using stacked chart using different colours where red shows *alp*1, light green – *alp*3, orange – *bca*, dark green – *rib*, grey – *hvg*A, purple, olive and pink – Pilus Island genes (PI-1, PI-2a and PI-b) respectively, black – *srr*-1 and white – *srr*-2 gene.

### 3.3.3 Phylogenetic relationship between GBS strains causing disease in adults

#### 3.3.3.1 Clonal complex 23 (CC23)

The largest clonal complex CC23 included isolates of ST23 (serotype Ia, n=38; and serotype III, n=2) and additional four STs of serotypes Ia (**Figure 3.3**). Phylogenetic analysis of CC23 isolates showed two separate clusters based on serotype and SPGs disparity; cluster 1 (salmon) solely made up of serotype Ia (n=40) ST23 or SLV of ST23 isolates that lacks *alp*2 and cluster 2 (yellow) comprised serotype III (n=2) ST23 isolates that contains *alp*2 gene (**Figure 3.3**). A high level of homogeneity was observed between all serotype Ia ST23 isolates which differed on average by 294 SNPs per isolate

(range: 77-1009 SNPs) (**Figure 3.3**). No specific clustering of isolates based on age group was identified (**Figure 3.3**)

### 3.3.3.2 Clonal complex 1 (CC1)

Of the 32 CC1 isolates in the study the majority were ST1 (23/32; 71.8%) serotype V (17/23,73.9%), while the remaining were serotype Ib (n=2), II (n=2) and VI (n=2). Phylogenetic analysis indicated three clusters based on serotypes and SPGs: cluster 1 (serotypes VI - purple) only cluster positive for *bca* gene, cluster 2 (serotype IV, ST196, ST136 and ST459 - orange) – exclusively contains *bib*A ; and cluster 3 (serotype V, serotype Ib and II strains - green) – contains *alp*3 and *bib*A (**Figure 3.4**). No clustering based on age group was noted.

### 3.3.3.3 Clonal complex 8/10 (CC8/CC10)

Within CC8/CC10 there were 18 serotype Ib isolates of which 9 ST8 co-located phylogenetically showing less divergence between them (**Figure 3.5**). No clustering based on serotype, patient's age group or SPGs was detected however, based on STs, three distinct clusters were noticed consisting of serotype Ib and II isolates in combination: cluster 1 (plum) contains ST12 isolates only, cluster 2 (green) comprised ST (8, 9 and 10) isolates whereas cluster 3 (yellow) includes ST (15 and 104) isolates (**Figure 3.5**).

### 3.3.3.4 Clonal complex 19 (CC19)

Three distinct clusters based on serotypes were found in CC19, cluster 1 (blue) comprised of serotype III and V ST19 or SLV of ST19 isolates in combination with an average 472 SNP difference; (range :64-1059) observed between five serotype V isolates (**Figure 3.6**). The cluster 2 (bisque) consist of serotype II ST28 isolates and cluster 3 (green) made up of serotype III ST19 or SLV of ST19 isolates (**Figure 3.6**). Cluster 2 largely consist of GBS (5/8, 62.5%) collected from younger patients belonging

to 14-64 years age group. No dominant age group was observed in remaining cluster 1 and 3. All CC19 isolates carried *rib*, *bib*A and PI (1a and 2b) in combination except four ST19 serotype V isolates that lack *rib* gene (**Figure 3.6**).

#### 3.3.3.5 Clonal complex 17 (CC17)

Phylogenetic analysis indicated three main clusters based on serotypes expressed by CC17 isolates, the cluster 1 and 2 (salmon and blue) consist of serotype III isolates (n=30) and (n=7), respectively being ST17 or SLV of ST17, whereas cluster 3 (brisque) comprised serotype III (n=2) and serotype IV (n=1) in combination (**Figure 3.7**). Overall, serotype III isolates of all three clusters (average 717 SNP difference; range: 136 - 1075 SNPs) were more commonly identified in the 18 – 44 years age group (19/29, 65.5%) (**Figure 3.7**). Whereas serotype IV ST1351 (a novel ST assigned in this study) isolate - PHEGBS0467 clustered with two serotype III isolates in cluster 3, does not express the expected serotype III and additionally also carried the *hvg*A gene exclusively detected in all serotype III CC17 isolates (**Figure 3.7**).

Overall, none of the other CCs had specific clustering observed in isolates based on geographical origin, patient age group or antimicrobial profile. No clusters of isolates were noted indicating potential for infection in cases with the same strain.

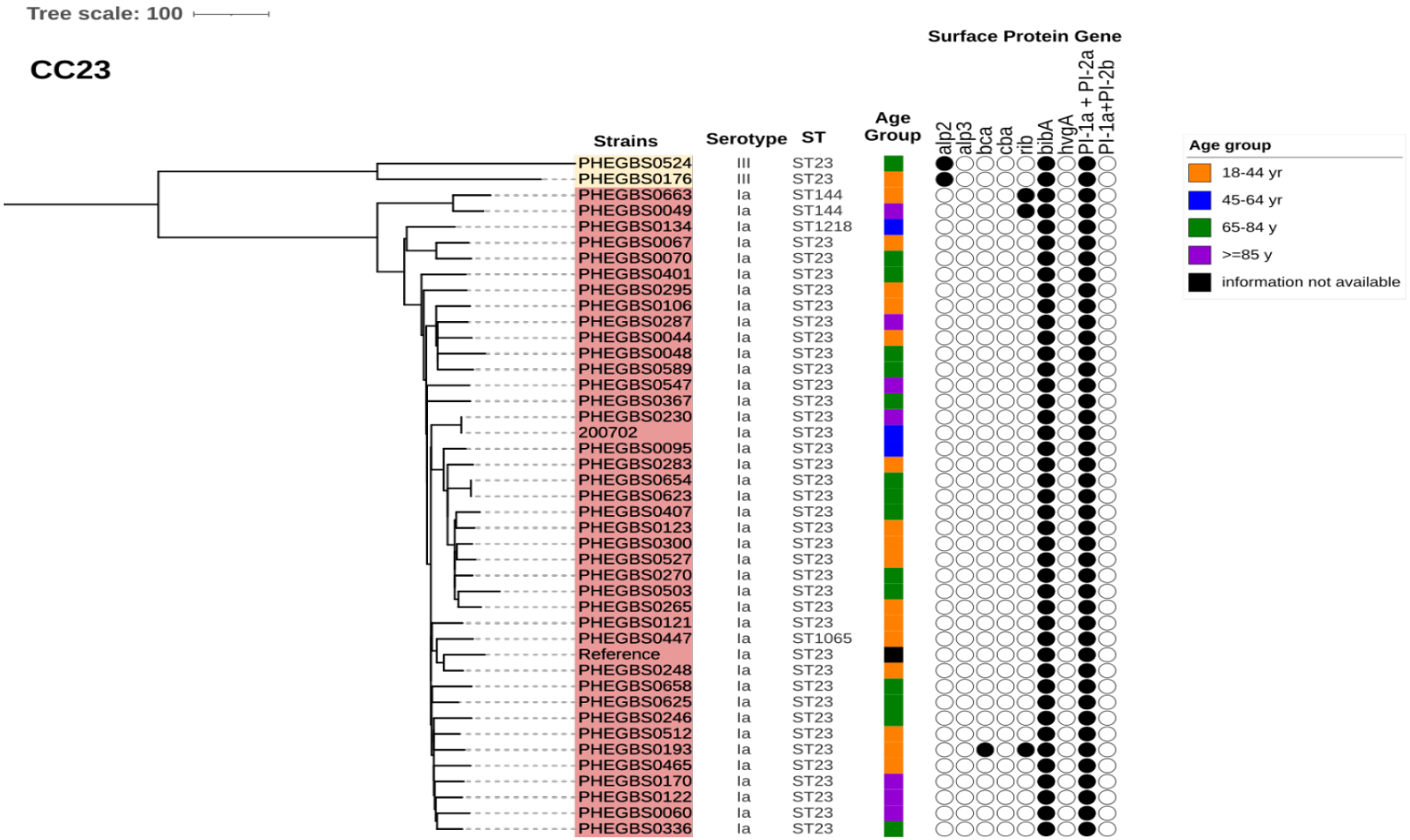



**Figure 3. 3. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 23 isolated from January 2014 and May 2015.** Forty-two isolates belonged to CC23 and are represented by a mid-point rooted maximum likelihood phylogeny tree using SNPs based approach (6,826 core SNPs for this analysis) relative to reference sequence FDAARGOS 512 (NZ_CP033822.1). Each isolate has a unique identifier, listed next to the capsular serotyping result and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Age group (18-44: orange; 45-64: blue; 65-84: green; ≥85 years of age: purple; or not available: black) and surface protein genes (SPGs – are indicated in black circle if present in an isolate otherwise blank shows absence) are indicated on the right of the figure. Distinct clusters seen in phylogeny are coloured where cluster 1 is shown in salmon and comprised of serotype Ia isolates only; cluster 2 – lime and consist of serotype III isolates only. Scale bar, represents a distance of 100 point mutations.

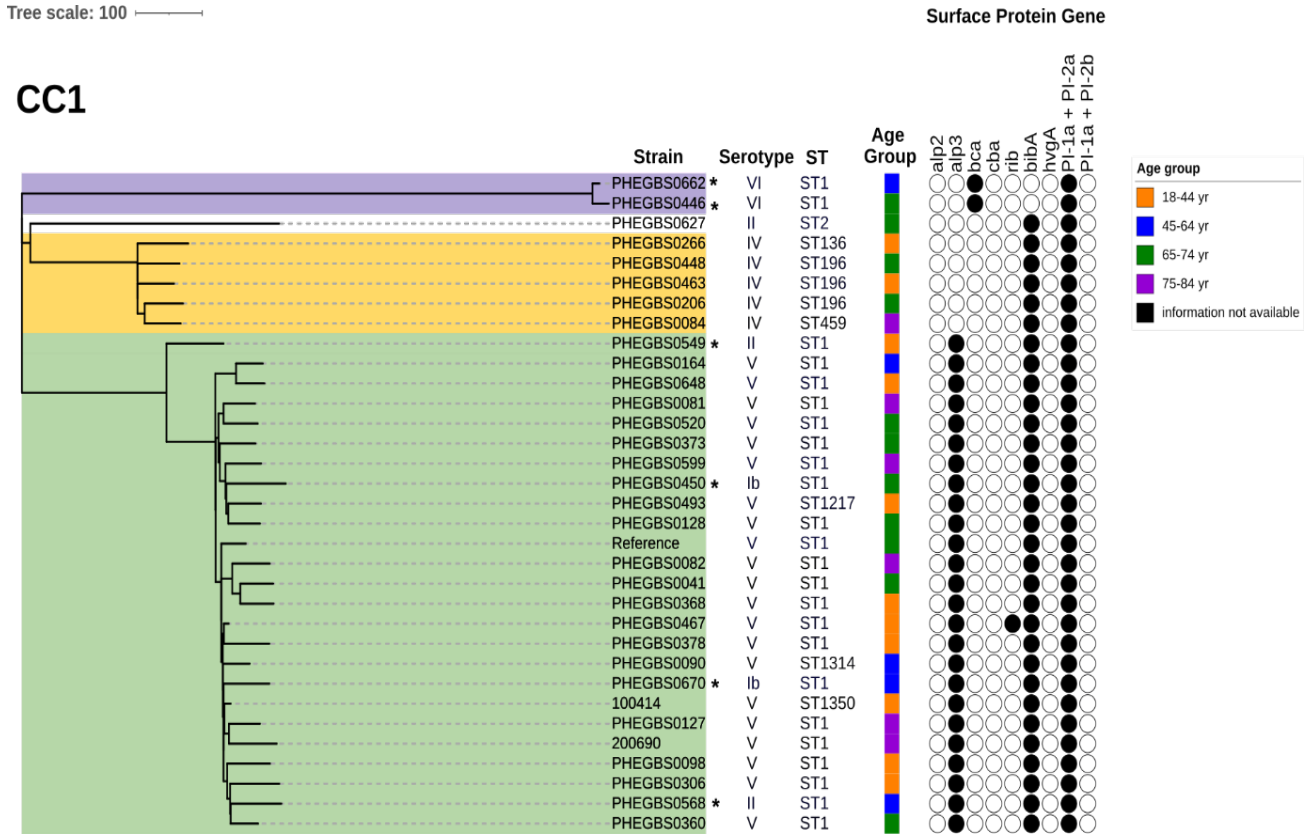



**Figure 3. 4. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 1 isolated from January 2014 and May 2015.** Thirty-two isolates belonged to CC1 and are represented by a mid-point rooted maximum likelihood phylogeny tree using SNPs based approach (5,524 core SNPs for this analysis) relative to reference sequence SS1 (NZ_CP010867). Distinct clusters seen in phylogeny are coloured where cluster 1 is shown in purple and comprised of serotype VI strains only; cluster 2 – orange and consist of serotype IV strains only and cluster 3 – green consist of serotype V, Ib and II strains. Each isolate has a unique identifier, listed next to the capsular serotyping result and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Age group (18- 44: orange; 45-64: blue; 65-84: green; ≥85 years of age: purple; or not available: black) and surface protein genes (SPGs – are indicated in black circle if present in an isolate otherwise blank shows absence) are indicated on the right side of the figure. Strain identifiers followed by an asterisk (*) indicate an unexpected combination of capsular serotype and sequence type. Scale bar represents a distance of 100-point mutations.

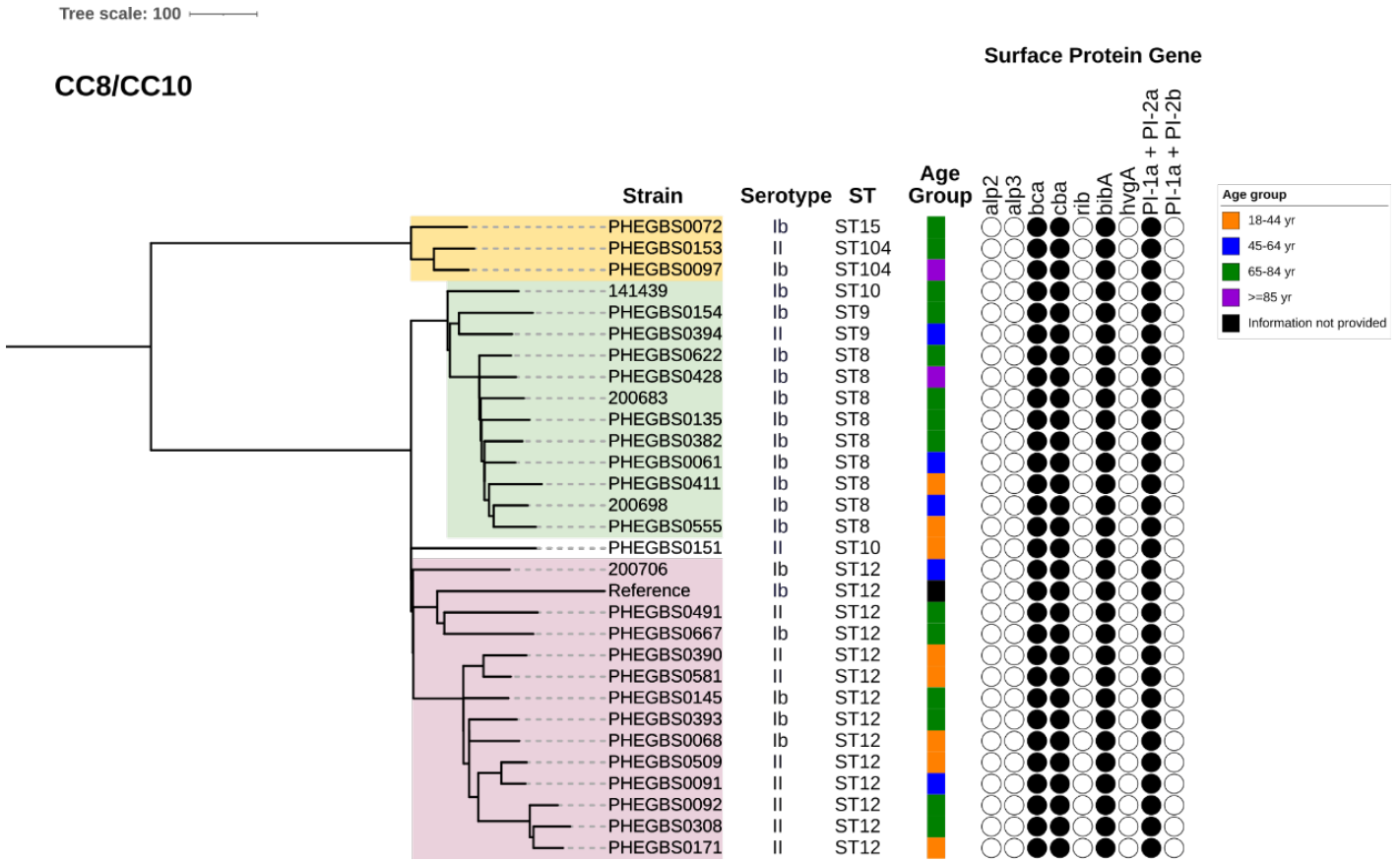



**Figure 3. 5. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 8 isolated from January 2014 and May 2015.** Twenty-nine isolates belonged to CC8/CC10 and are represented by a mid-point rooted maximum likelihood phylogeny tree using SNPs based approach (9,153 core SNPs for this analysis) relative to reference sequence Sag37 (NZ CP019978.1). Each isolate has a unique identifier, listed next to the capsular serotyping result and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Age group (18-44: orange; 45-64: blue; 65-84: green; ≥85 years of age; purple; or not available: black) and surface protein genes (SPGs – are indicated in black circle if present in an isolate otherwise blank shows absence) are indicated on the right side of the figure. Distinct clusters seen in phylogeny based on STs are coloured where cluster 1 is shown in plum and comprised of ST12 isolates only; cluster 2 – green and consist of ST (8, 9 and 10) isolates and cluster 3 – yellow consist of ST (15 and 104) isolates. Scale bar, represents a distance of 100 point mutations.

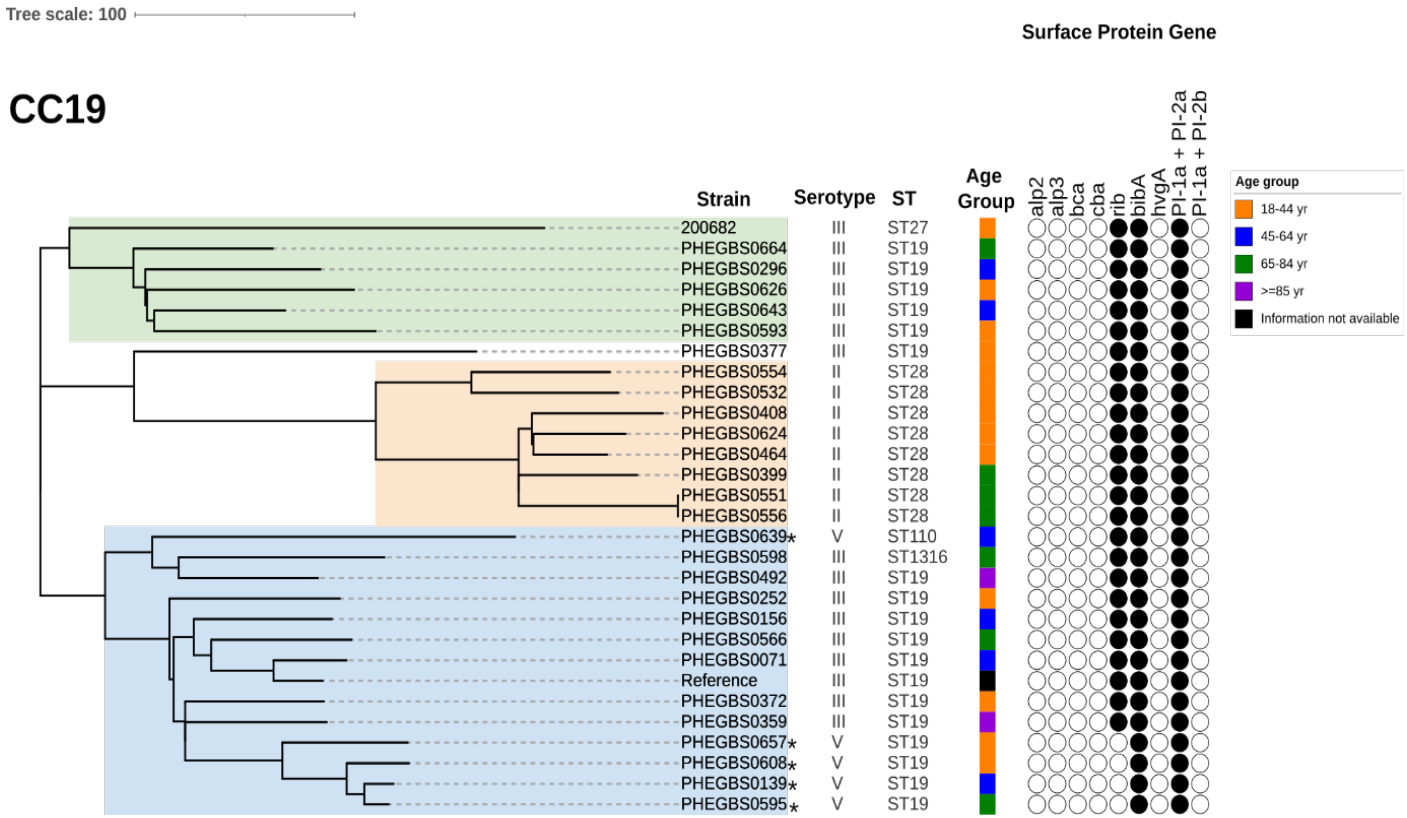



**Figure 3. 6. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 19 isolated from January 2014 and May 2015**. Twenty-eight isolates belonged to CC19 and are represented by a mid-point rooted maximum likelihood phylogeny tree using SNPs based approach (5,260 core SNPs for this analysis) relative to reference sequence SG-M25 (NZ_CP021867.1). Each isolate has a unique identifier, listed next to the capsular serotyping result and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Age group (18-44: orange; 45-64: blue; 65-84: green; ≥85 years of age; purple; or not available: black) and surface protein genes (SPGs – are indicated in black circle if present in an isolate otherwise blank shows absence) are indicated on the right side of the figure. Three distinct clusters based on serotype are indicated in different colours where blue shows cluster 1, bisque – cluster 2, and green – cluster 3. Strain identifiers followed by an asterisk (*) indicate an unexpected combination of capsular serotype and sequence type. Scale bar, represents a distance of 100 point mutations.

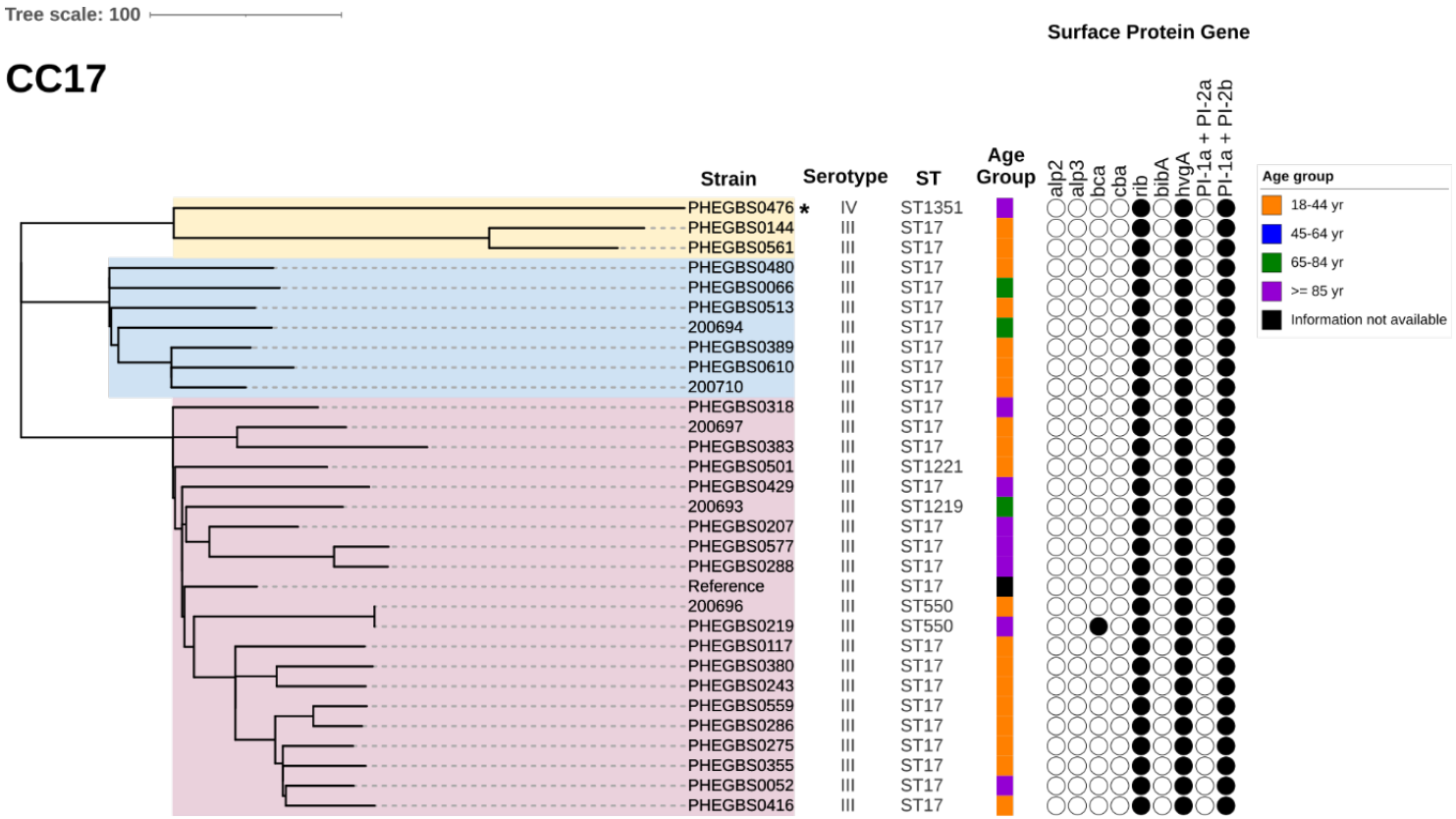



**Figure 3. 7. Phylogenetic relationship between iGBS from adult patients belonging to clonal complex 17 isolated from January 2014 and May 2015.** Thirty isolates belonged to CC17 and are represented by a mid-point rooted maximum likelihood phylogeny tree using SNPs based approach (2,699 core SNPs for this analysis) relative to reference sequence COH1 (NZ_HG9394456.1). Each isolate has a unique identifier, listed next to the capsular serotyping result and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Age group (18-44: orange; 45-64: blue; 65-84: green; ≥85 years of age; purple; or not available: black) and surface protein genes (SPGs – are indicated in black circle if present in an isolate otherwise blank shows absence) are indicated on the right side of the figure. Three separate clusters based on serotype found in CC17 phylogeny are cluster 1 – salmon, cluster 2 blue and cluster 3 – brisque. Strain identifiers followed by an asterisk (*) indicate an unexpected combination of capsular serotype and sequence type. Scale bar, represents a distance of 100 point mutations.

### 3.3.4 Recombination events detected uncommon serotype within specific genotype

Among this collection serotypes were associated with specific CCs, as expected (**Figure 3.8**). However, several exceptions were notable, where closely related strains sharing the same genotype were found to have a different capsular type (asterisks in **Figure 3.4**), these combinations have been documented [12,14,16,38], but not always investigated in details.

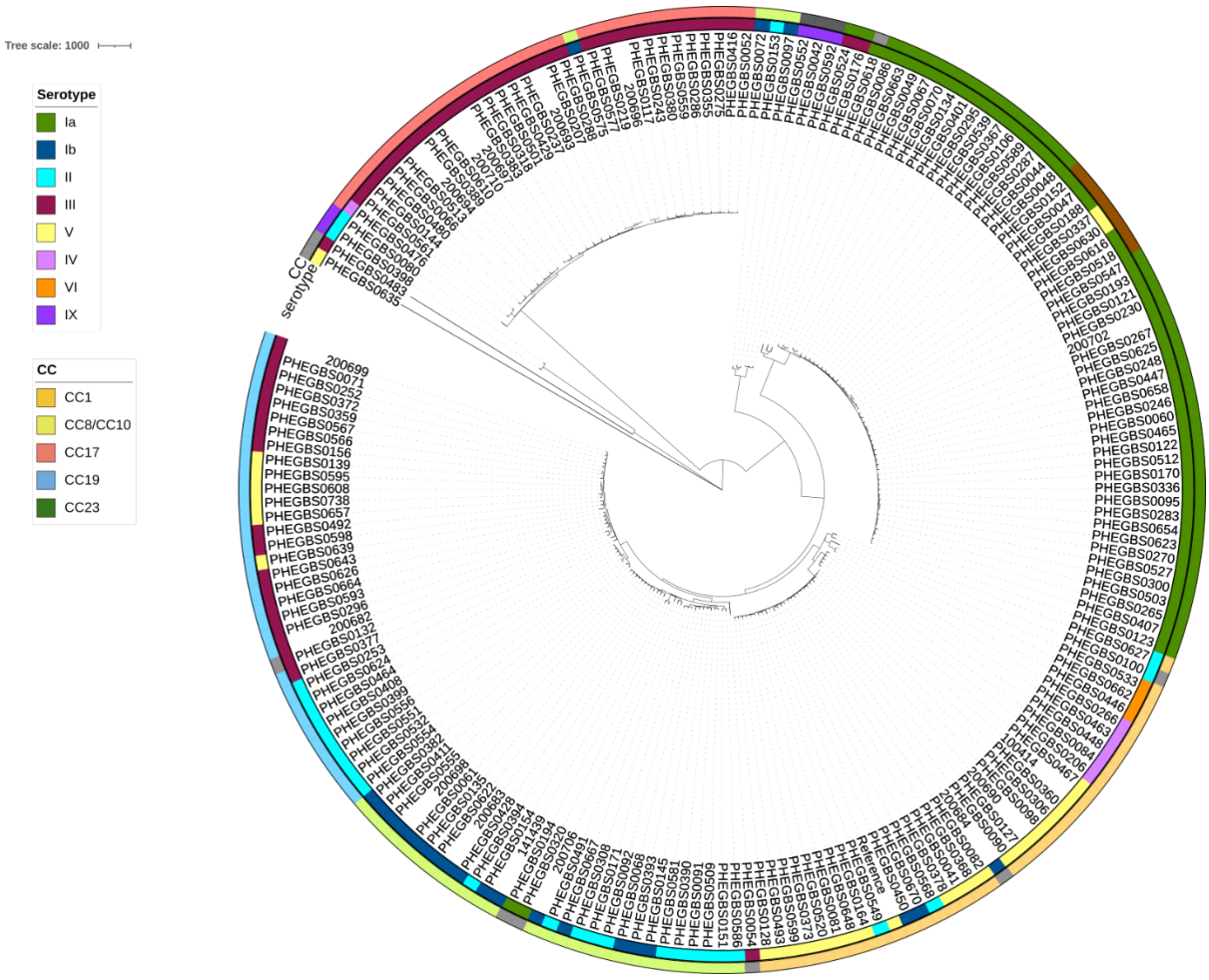



**Figure 3. 8. Phylogenetic relationship between 179 iGBS isolates causing disease in adult population in the UK.** A total of 179 genomes were analysed for the study. Mid-point rooted maximum likelihood phylogeny tree built using SNPs after correcting for recombination. The inner ring indicates one of the eight serotypes identified amongst 179 iGBS isolates in this study. The outer ring indicates five major clonal complexes CC23, CC1, CC19, CC17 and CC8/CC10 identified in the study, purple indicates CC22/CC1213, and brown indicates CC24/CC498, dark grey indicates CC130/CC1216, and light grey colour indicates isolates identified as singleton based on goeBURST SLV analysis. Scale bar, represents a distance of 1000 point mutations.

Six isolates of ST1 lineage were identified in this study that carry ST1 core genome that is more consistent with serotype V but show a potential capsular switching event with serotypes Ib, II and VI (**Figure 3.4**). PHEGBS0670 and PHEGBS0450 (both serotype Ib, ST1) acquired DNA regions of 217kb and 209kb, respectively, including the *cps*Ib locus (**Figure 3.9a**), PHEGBS0549 and PHEGBS0568 (both serotype II) acquired DNA regions of 232 kb and 144 kb, respectively, including the *cps*II locus (**Figure 3.9Ib**). In contrast, serotype VI ST1 isolates PHEGBS0662 and PHEGBS0446 had larger scale recombination events (including the *cps*IV locus) of 883 kb and 884 kb, respectively, involving approximately 45% of the genomes, with the resultant genomes containing all of the ST1 MLST determinants (**Figure 3.9c**). All six isolates retained high similarity to reference strain SS1 and all contained the *alp*3 gene and PI-1 and PI-2a pilus combination common to serotype V strains [39,40]. To confirm the identified capsular switching events, ST1 isolates in this study were compared to previously published ST1 isolates with recombination events; ST1-serotype Ib (SH5446) from Portugal [78], and ST1-serotype II (NGBS748) and ST1-serotype VI (NGBS209) from Canada [77]. Phylogenetic analysis of these three GBS isolates with ST1 isolates in this study showed dispersal throughout the UK ST1 phylogenetic cluster with an average of 207 and 583 SNPs per serotype Ib and VI ST1 isolates, respectively (**Figure 3.10**). The recombination borders were nearly identical for all isolates, confirming the potential capsular switching events have taken place between serotypes V and Ib and serotype V and II (**Figure 3.11a**).

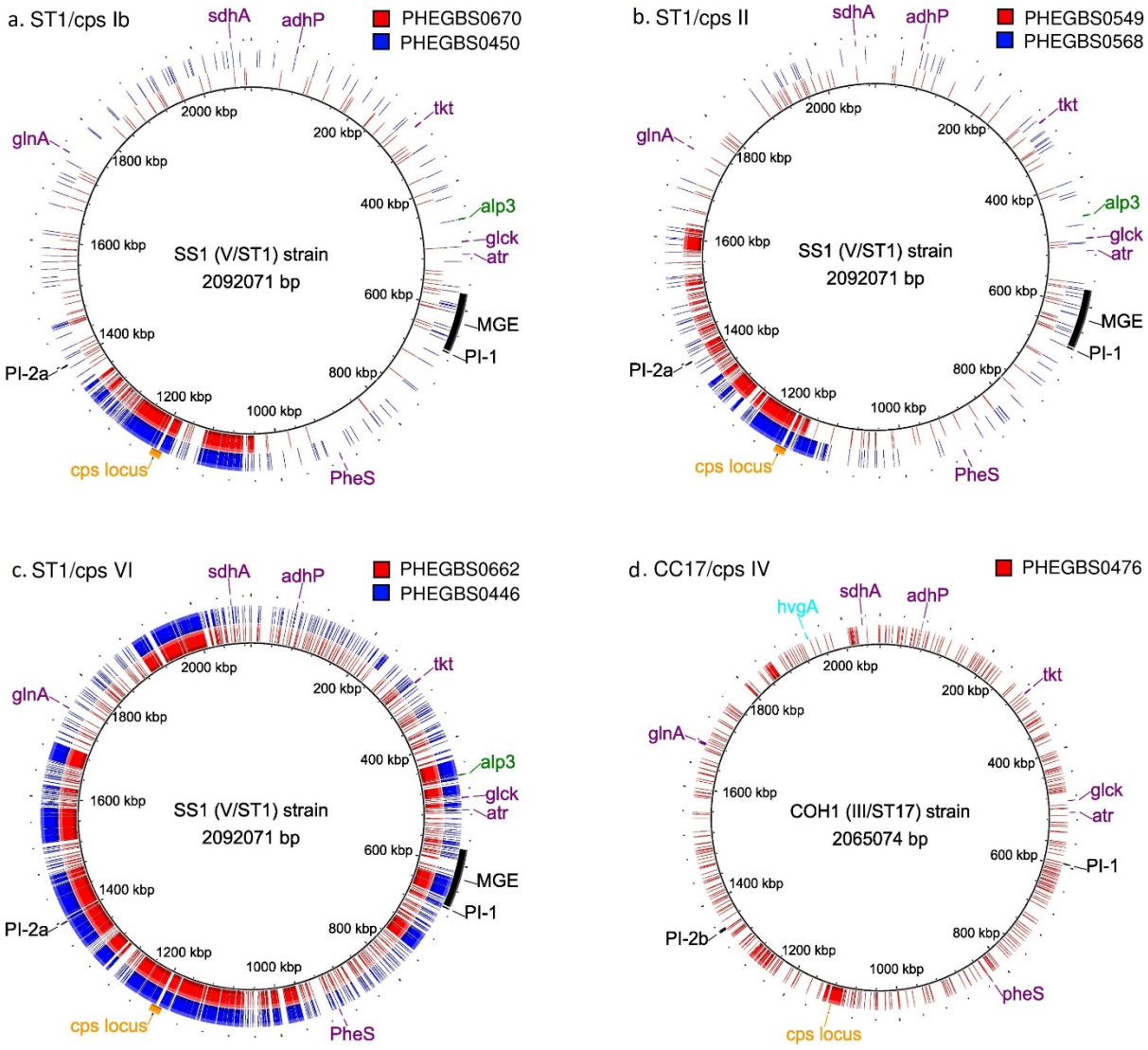



**Figure 3. 9. Recombination leading to serotype switching for ST1 and CC17 GBS strains from adult patients isolated from January 2014 and May 2015.** Polymorphisms determined in unexpected serotypes within ST1: serotype Ib ST1 strains PHEGBS0670 and PHEGBS0450 (a); serotype II ST1 strains PHEGBS0549 and PHEGBS0568 (b); and serotype VI ST1 strains PHEGBS0662 and PHEGBS0446 (c) are plotted in BRIG [511] against the reference serotype V ST1 strain SS1 (NZ_CP010867). Serotype IV ST1351 (CC17) PHEGBS0476 strain (d) is plotted against serotype III ST17 reference strain COH1 (NZ_HG9394456.1). The innermost (black circle) represents the respective reference SS1 or COH1. Following genome landmarks are shown in the outermost circle: *cps* locus: capsular polysaccharide locus in orange, *alp*3: alpha like surface protein encoding gene in green, *hvg*A: hypervirulent surface anchored adhesin gene in aqua, seven multilocus sequence typing (*adhP, atr, tkt, glcK, sdhA, glnA,* and *pheS*) genes in purple, MGE: Mobile genetic element in black, *PI-1* and *PI-2a*: Pilus Island genes also in black. Polymorphisms mapping to MGE identified in the various non–serotype V ST1 strains were not included in the analysis.

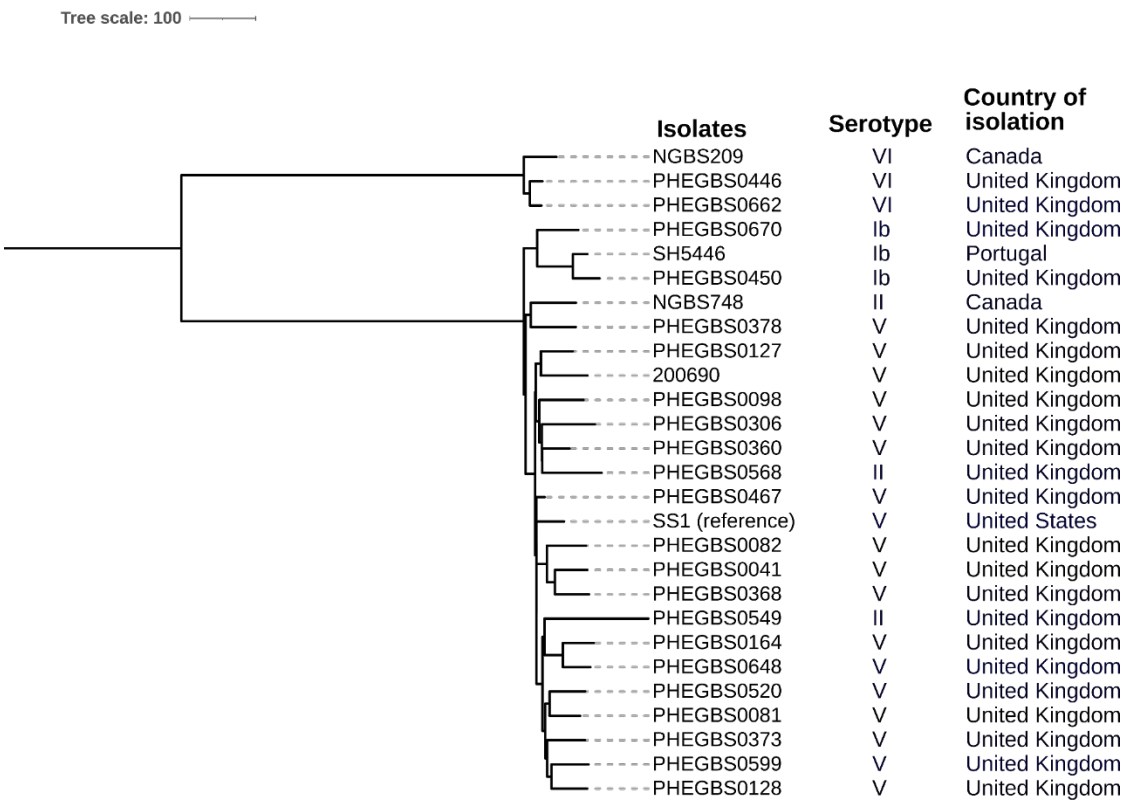



**Figure 3. 10. Phylogenetic relationship between UK ST1 isolates and previously reported recombined Portugal and Canadian ST1 non-serotype V isolates.** Maximum likelihood phylogenetic tree of UK (n=23), Portugal (n=1, [78]) and Canada (n=2, [77]) GBS isolates based on 1,964 core SNPs with recombination removed was generated using a reference sequence SS1 (NZ_CP010867.1). Majority of UK ST1 isolates were serotype V (n=17), followed by serotype Ib, II and VI (n=2 per serotype), Portugal ST1 isolate (SH5446) was reported as serotype Ib, Canadian isolates (NGBS748) and (NGBS209) were serotypes II and VI, respectively. From left to right: mid-point rooted maximum likelihood phylogeny tree, followed by information on serotype, and country of isolate origin. Scale bar, represents a distance of 100 point mutations.

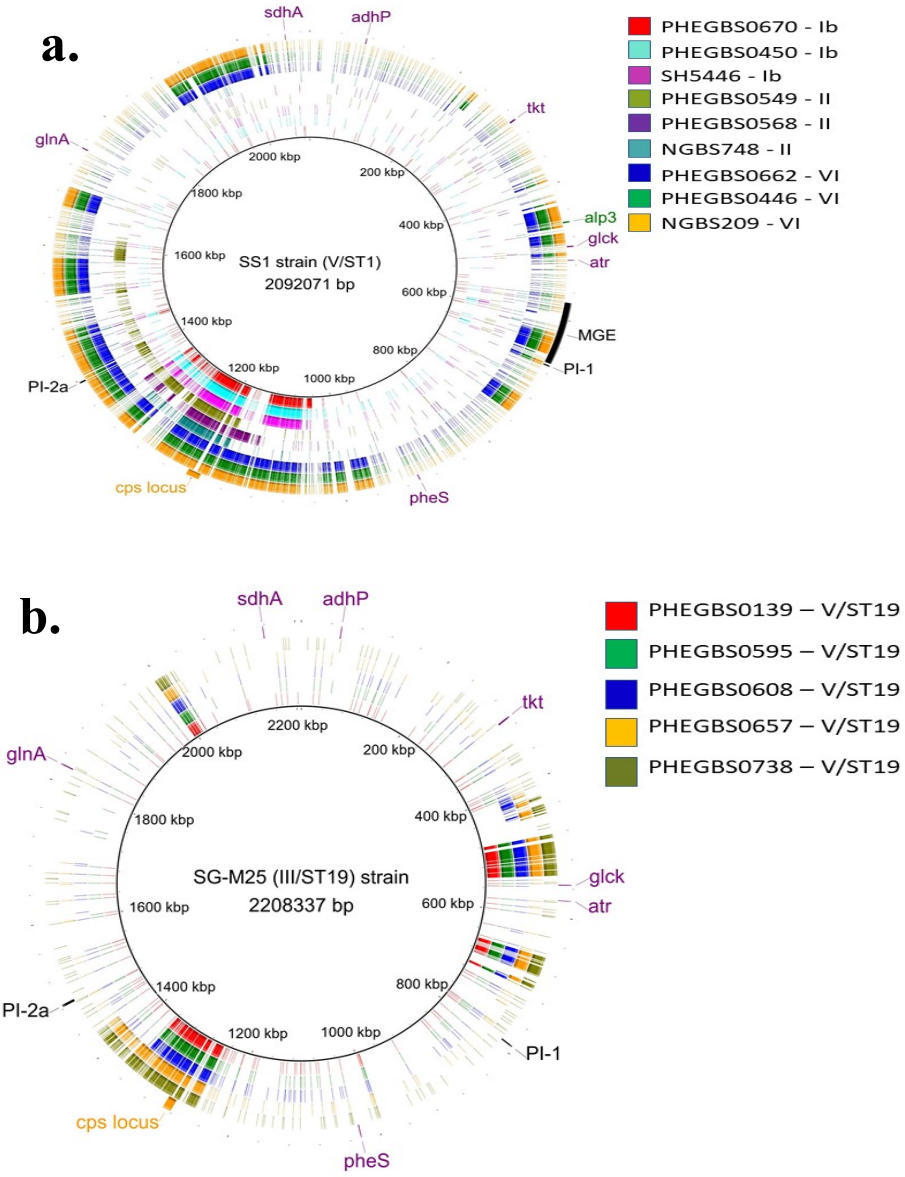



**Figure 3. 11. Comparison of recombination events among selected GBS groups. BRIG [511] was used to show Single Nucleotide Polymorphisms (SNPs) distribution. (a)** Among ST1 serotypes Ib, II and VI. ST1 isolates of serotype Ib (n=3), II (n=3) and VI (n=3) from UK, Portugal and Canada were compared against reference strain SS1 (NZ_CP010867.1). From inner to outermost circle: inner circle (black line) is reference SS1, then polymorphisms identified in each of the GBS isolates: serotype Ib (red – PHEGBS0670, UK; aqua - PHEGBS0450, UK; and magenta – SH5446, Portugal [78]; serotype II (olive - PHEGBS0549, UK; purple - PHEGBS0568, UK; and teal – NGBS748, Canada [77]); serotype VI (blue – PHEGBS662, UK; green – PHEGBS446, UK; and orange – NGBS209, Canada. **(b)** Among ST19 serotype V isolates. ST19 serotype V isolates (n=5) identified in the study and compare them to the ST19 serotype III reference genome SG-M25 (NZ_CP021867.1). From inner to outermost circle: inner circle (black line) is reference SG-M25, then polymorphisms identified in each of the GBS isolates. Outermost circle indicates genome landmarks as follows: *cps* locus - capsular polysaccharide locus (in orange), *alp*3 - alpha like surface protein encoding gene (in green), seven multi-locus housekeeping genes (*adhP*, *atr*, *tkt*, *glcK*, *sdhA*, *glnA*, and *pheS*, in purple), MGE - Mobile genetic elements and pilus island genes PI-1 and PI-2a (in black).

On examination of the potential capsular switch event in serotype IV isolate ST1351 within CC17 (ST1351 is a SLV of ST17), this isolate carried *hvg*A, which is commonly associated with hypervirulent T17 strains. Recombination analysis indicated PHEGBS0476 potential for capsular switching could be attributed to event with a 31.3 kb DNA fragment exchange that contained the entire *cps*IV operon, comparable to the DNA transfer of 35.5 kb reported in serotype IV/ST17 isolate from France (CC209361, [16]) (**Figures 3.9d** and **3.12**). Comparison of this isolate to other IV isolates from CC1 revealed the potential transfer of *cps*E-*cps*G and *neu*B-*neu*A genes was highly conserved, irrespective of whether they were present in isolates with determinants designating them CC1 or CC17, not only in isolates from this study, but also in isolates previously reported to undergo the capsular switching in CC1 and CC17 serotype IV isolates from France (CC209361, [16]) and Ireland (GBS148, GBS150 and GBS175, [41]), where two SNPs, one in each of the *cpsE* and *cpsF* genes (PHEGBS0476) were observed (**Figure 3.12**).

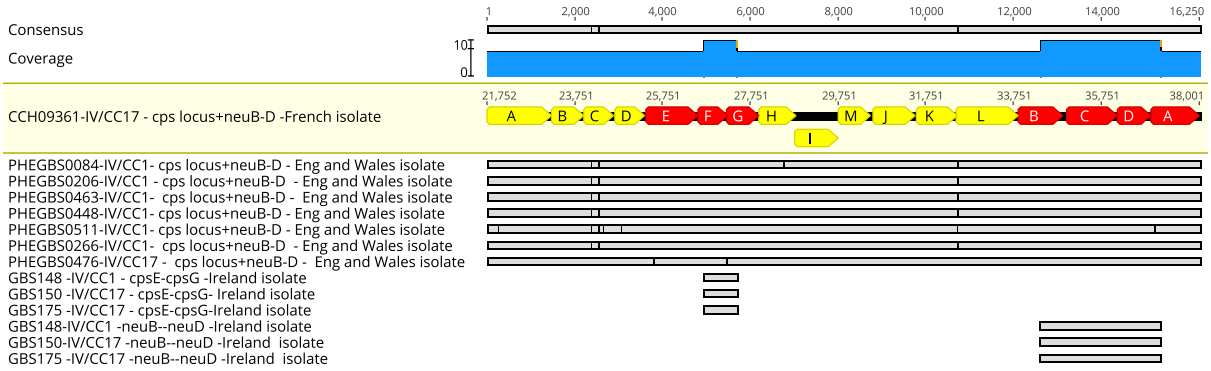



**Figure 3. 12. SNPs distribution in conserved *cps*E-*cps*G and *neu*B-*neu*A regions of *cps*IV locus of CC1 and CC17 GBS isolates from Ireland, France and England and Wales.** Mapping, SNP calling, and visualization was performed by Geneious 11 [498] using entire *cps*IV sequences of current study CC1 and CC17 serotype IV isolates (PHEGBS0084, PHEGBS0206, PHEGBS0463, PHEGBS0448, PHEGBS0266 and PHEGBS0476) and available partially (*cps*E, *cps*G, *neu*B and *neu*A) and fully sequenced (*cps*F, *cps*G, *neu*C and *neu*D) sequences of Irish CC1 and CC17 serotype IV isolates (GBS148, GBS150 and GBS175) [4] using a reference full *cps*IV locus of French CC17 serotype IV isolate (CC209361) [5]. Different colours are used to highlight important landmarks in *cps*IV locus of CC209361 with yellow representing *cps*(A-D, H-M) genes, red showing conserved *cps*(E,F,G) and *neu*(B, C, D and A) genes respectively. Grey region showing similarity whereas each vertical black line representing a single-nucleotide polymorphism (SNP) in each isolate against a reference *cps*IV locus.

Investigation of potential recombination events between five serotype V ST19 isolates (PHEGBS139, PHEGBS595, PHEGBS608, PHEGBS657 and PHEGBS738), which formed a cluster in the CC19 clade (**Figure 3.6**). These isolates were compared to a serotype III ST19 reference. Recombinant DNA regions identified included *cps*V locus in PHEGBS139, PHEGBS595, PHEGBS608 of 301 kb in size and in PHEGBS657 and PHEGBS738 acquisition of 334 kb and 344 kb regions respectively (**Figure 3.11b**).

Potential recombination and capsular switching events were also observed in serotype V ST498 (CC23; **Figure 3.13**) isolates (PHEGBS337 and PHEGBS630) with 353 and 352kb DNA recombination boundary segments (**Figure. 3.13**). No transposons were detected in close vicinity of recombination events identified in any of serotype-genotype mismatches.

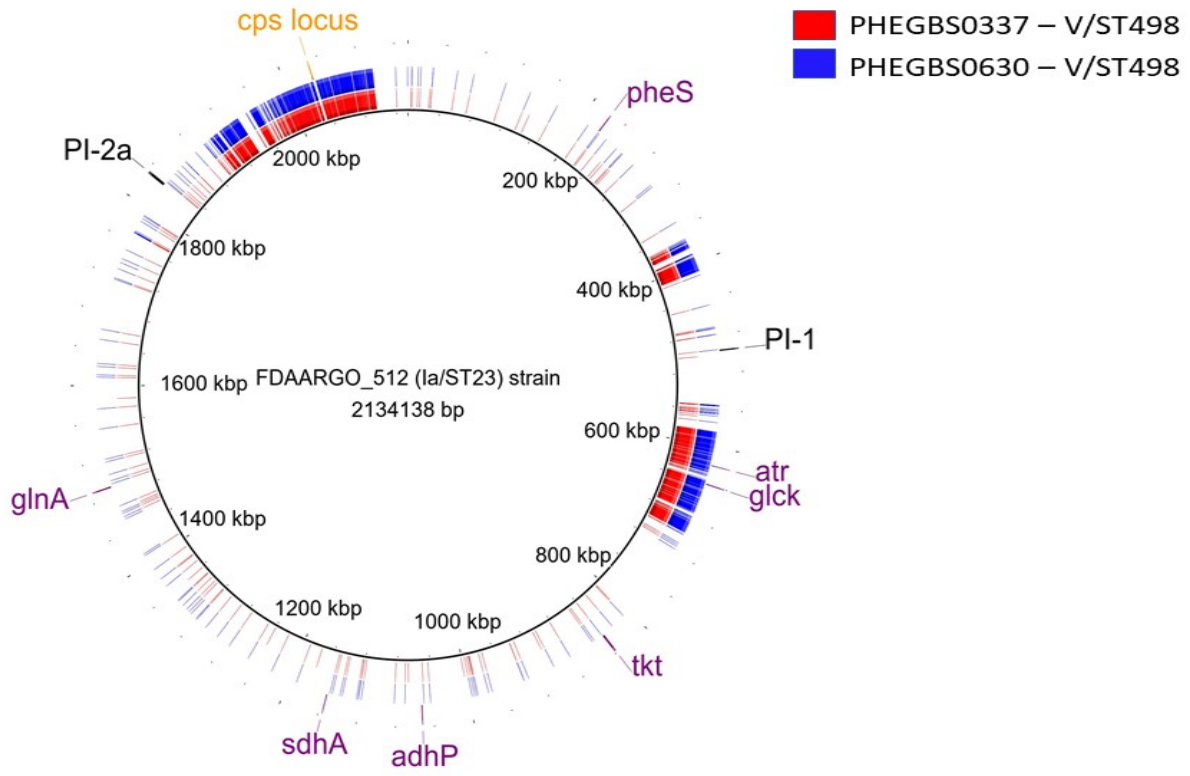



**Figure 3. 13. Comparison of recombination among serotype V ST498 (CC23) isolates.** BRIG [511] was used to show Single Nucleotide Polymorphisms (SNPs) distribution in CC23 isolates ST498 serotype V (n=2) against ST23 serotype Ia reference genome FDAARGO_512 (NZ_CP033822.1). From inner to outermost circle: inner circle (black line) is reference FDAARGO_512, then polymorphisms identified in each of the two GBS isolates: PHEGBS0337 and PHEGBS0630;

outermost circle indicates genome landmarks as follows: cps locus - capsular polysaccharide locus (in orange), seven multilocus housekeeping genes (*adhP, atr, tkt, glcK, sdhA, glnA,* and *pheS, in purple*), and pilus island genes - PI-1 and PI-2a (in black).

## 3.4 Discussion

Although bacteria reproduce clonally, their genomes evolve not only by point mutations but also by DNA exchange through homologous recombination which follows unidirectional horizontal DNA transfer mediated by transformation, transduction, or conjugation [520, 534]. GBS genomes are genetically very isolated, and rate of recombination in GBS are rare compared to other bacterial species [473, 535, 536]. There is an interesting dichotomy between very clear evidence of large-scale recombination between different lineages of GBS [296, 537, 538] with very little recombination within expanding clones, which instead evolve largely by accumulation of mutations [76, 335, 402, 539]. Interestingly, there have been several noted instances of serotype switching reported in countries including USA [76], Canada [338], France [269], Ireland [532] and Portugal [78], likely through large-scale recombination (most clearly seen in originally serotype V ST1 strains that have converted to serotype Ib, II, and VI through apparently single recombination, ranging from 79 to 200Kbp, encompassing the capsule determining *cps* locus) and also seen in the UK for the very first time through this study. Capsular switching in GBS may still be a rare occurrence and have mostly seen from a serotype to hexavalent vaccine covered serotype that could provide up to 98% coverage [265], yet an introduction of capsular serotype-based vaccines could create a greater selection pressure leading to serotype replacement in disease as seen with childhood pneumococcal conjugate vaccine (PCV) immunisation [540] specifically an increase in non-vaccine serotypes.

Previous reports note an increase of iGBS infections in adults between 1991-2010 and 2015-2016 in England and Wales [115, 517] which is consistent with the similar reports of rising GBS invasive cases in adults in other geographical regions [516, 541]. This increase of identified iGBS infections in adults is not fully understood. However, increasing longevity of populations, increasing emergence of contributing diseases (e.g., diabetes) and immunosuppression have been postulated as contributors [516,

542]. In 2014, 1676 and in 2015, 1870 cases of GBS bacteraemia were reported by laboratories in England, Wales and Northern Ireland to PHE. Isolates received are from patients of all ages, with the majority from neonates (estimated >95%) and women of child-bearing age, with an estimated total adult isolate submission proportion of less than 11% (n=179) of all invasive cases. In both years, GBS rates were higher in males than females in all age groups except the 15 to 44 years age group (females 2.2 and males 0.5/100,000) [480, 481] which agrees to most of our findings where more GBS samples collected in 2014-2015 were from male patients belonging to age groups (> 65 yrs) respectively, and for the 18-44 years age group more GBS were collected from female patients. A discrepancy observed for the number of iGBS isolates observed for the 45-64 years age group in this study where more isolates were recovered from female (n=12) than male (n=11), contrary to the PHE GBS surveillance report of 2014 and 2015, this difference could possible arise due to the total number of GBS bacteraemia cases reported in PHE reports but not the actual GBS samples recovered in the laboratories.

This chapter further demonstrates serotype and genotype distribution of GBS infections reported in adults between 2014 and 2015 in the UK. Serotypes III, Ia and V predominate in this collection, indicating no change in the UK population since the 2010 [115]. Forty-seven different MLST genotypes, including 15 new STs, were identified; these isolates were largely clustered into five major clonal clusters (CC1, CC8/CC10, CC17, CC19 and CC23) as reported previously [1, 515, 532]. This suggests an overall limited genomic diversity within UK GBS population, as also observed in other studies [1, 665]. Of importance, no geographic or phylogenetic clusters of isolates were noted indicating no evidence of clonal transmission.

For the most part, expected correlates between virulence gene and pilus islands segregation to particular serotypes or CCs reviewed previously [78, 353, 543]. For example, serotype V/*alp*3, serotype Ib and II/*bca*+*bac*, and serotype Ia/*bib*A that predominated in CC1, CC8/CC10 and CC23, respectively. While serotype III/*hvg*A exclusive to CC17 and serotype III/*rib* divided between CC17 and CC19. All isolates harboured at least one of PI genes, of which the combination PI-1+PI-2a was most prevalent, whereas PI-1+PI-2b was found exclusively in serotype III CC17 isolates.

Genes encoding the most abundant GBS surface Alp family proteins such as Alp2, Alp3, AlphaC (*bca*) and Rib, were readily identified in this GBS dataset. AlphaC protein in GBS is important for interactions with cervical and epithelial cells to promote intracellular invasion and dissemination [44,45]. The other Alp family surface proteins are also of great interest as they provide immunological reaction and potentially protection against GBS disease [46]. Due to this, studies report associations observed between GBS serotypes and genotypes. A recent meta-analysis study on global GBS isolates have reported that 79% of invasive adult GBS had at least one of the Alp2, Alp3, AlphaC (*bca*) or Rib protein genes [6], similar distribution has been observed in our study as well with 72.5% (n=140/193) of the isolates carrying one of the mentioned alp family proteins. We have found serotype III isolates were reported to more commonly be associated with Rib protein and serotype V isolates to be associated with Alp3, which concurs with previous reports [6]; and AlphaC protein (*bca* gene) was found to be associated with serotype Ib in our study, though it has been reported to be associated with serotypes Ib and Ia [6]. Due to their antigenic properties, Alp family proteins have been proposed and potential targets for protein-based GBS vaccines, one of such using the N-terminal domain of AlphaC and Rib has been already studied in Clinical Trials Phase I (NCT02459262). Vaccines targeting GBS pilus proteins were also proposed, considering that these proteins are present universally in GBS strains, with PI-1+PI-2a being the most common [6] and were identified in this study also. However, there is a lack of data supporting its potential as vaccine candidate at this time and further analysis is needed [47,48].

This study has several limitations; the sample set is a cross sectional observational study limited to only referred isolates within adults over the time-period in question. It does not represent an accurate number of isolates from cases with disease burden or include isolates from screening adults. However, though this dataset is not representative of all GBS infection in adults during the study period, isolates analysed here are likely to be representative of infection of greatest severity, which are more likely to be sent to the reference laboratory. Irrespective, this study highlights the need for continuous genomic surveillance to monitor genotype, genotypic capsular serotype, genes associated with hypervirulence and antibiotic resistance to help inform future treatment of cases and potential vaccination programmes,

including polysaccharide- or protein-based GBS vaccines.

# Chapter 4:

**Description of antimicrobial resistance (AMR) mechanisms and identification of mobile genetic elements carrying AMR genes in *Streptococcus agalactiae* causing bloodstream infections in adults in UK between January 2014 and May 2015**

## 4.1 Summary

Antimicrobial resistance in bacteria is increasing globally and the extensive use of antimicrobials is inevitably driving expanding prevalence and evolution of diverse multidrug resistant (MDR) strains. The pool of mobile genetic elements (MGE) in microbial communities consists of viruses, plasmids, and associated elements (insertion sequences, transposons, and integrons) that provide key vehicles for gene transfer between bacteria and contribute directly to their evolution and, potentially to bacterial speciation. 193 clinical GBS strains isolated from adults submitted to the UK national reference laboratory (179 invasive; 13 non-invasive; one with no information provided) were tested for antimicrobial susceptibility and whole genome sequences were analyzed to characterize the present mobile genetic elements (MGEs) carrying antibiotic resistance genes (ARGs). All 193 GBS isolates were also positive for five penicillin binding protein types (1a, 1b, 2a, 2b and 2x) with amino-acid substitutions that did not confer penicillin resistance, which were not linked to mobile elements. 100% isolates were ampicillin and vancomycin susceptible, one isolate (0.5%) showed resistance to high level gentamicin while 98.4% (n=190) of strains were susceptible to chloramphenicol. Resistance to erythromycin (39/193, 20.2%) and clindamycin (23/193, 11.8%) were anticipated by the presence of *erm* methylases, *mef*, *lsa* or *lnu* determinants. Most isolates (175/193, 90.6%) carried *tet* gene determinants where the greatest prevalence of ARG (154/175, 88%) was for *tetM* positive isolates whereas only 17/193 (8.8%) of isolates carried no ARGs. Most of the *tet*M genes were carried by Tn*916*/Tn*916*-like elements (n=87, 56.4%) followed by Tn*5801* (n=67, 43.5%). Only one isolate carried the *erm*T gene and this was on a plasmid, while for 4 isolates fluoroquinolone resistance was mediated by non-mobile somatic mutation in gyrases and topoisomerases; for all other isolates, ARGs were acquired by MGEs including five novel MGEs: ICE*Sag*84 and ICE*Sag*100414 carrying *erm*A alone, ICE*Sag*662 containing *erm*B, *tet*S, *ant*(6-Ia) and *aph*(3'-III), ICE*Sag*71 carrying *erm*B and *tet*O ,and ICE*Sag*139 containing *erm*A and the high gentamicin level resistance gene *aac*(6')-*aph*(2''). The genetic basis of antimicrobial resistance in GBS has not been studied in depth until now, and the mobilome of these GBS isolates, combined with their multi-drug resistant phenotype, mirror the transfer and prevalence of MGEs contributing to the spread of antibiotic resistance worldwide and require special

attention. These findings could be useful in understanding the antimicrobial resistance genetics of GBS in clinical patients and the difficulty in tackling these infections.

## 4.2 Introduction

GBS is recognized as a cause of postpartum infections, meningitis, and neonatal sepsis [122]. The highest burden of GBS infections is observed in 0-3 months old infants, hence, this age group is the principal target for GBS infection control [122]; however, recent studies have shown increased invasive GBS (iGBS) infections (pneumonia, meningitis, bacteraemia, and soft tissue infection) in adults in developed countries [76, 313] and importantly mortality is markedly higher in adult systemic GBS infections than among neonates [115].

Ampicillin and penicillin are the therapeutics of choice to treat GBS infections because resistance to penicillin is thought to be rare or absent universally in GBS; however, reduced susceptibility to penicillin and other β-lactam antibiotics in GBS have been described due to mutations in PBP2X and other penicillin binding protein genes [505, 544]. For penicillin-allergic patients with low risk to anaphylaxis, cefazolin is the next alternative drug [545] which has analogous mechanisms of action and pharmacokinetics as ampicillin. Cefazolin rapidly crosses the placenta and is detected in cord blood and amniotic fluid at levels above the GBS MIC within 20 minutes after maternal administration [546-549]. Clindamycin given intravenously had been recommended for penicillin-allergic patients at high risk of anaphylaxis, (an extremely rare severe complication of pregnancy in United Kingdom [172]), however, since 2000 there has been a marked increase in resistance to clindamycin in GBS infected patients of all ages [115]. Within the UK, an increase in resistance to erythromycin and clindamycin in GBS was previously observed as a concerning trend [115] which is also reflected in more recent annual PHE surveillance reports of GBS, where resistance to clindamycin and erythromycin in GBS bacteraemia increased between 2016 and 2019, from 25% to 30% for clindamycin and 31% to 36% for erythromycin [8] which appears to be comparable to increasing GBS macrolide resistance trends observed in European and other countries in the last two decades [78, 389, 505, 550, 551]. For example, significant increase of erythromycin resistance in GBS is noticed in France [550] which was recorded 22% in 2007 and 30% in 2019, in Serbia [312], where erythromycin and clindamycin resistance increased from 19.7% to 29.2% and from 17.3% to 24.1%, respectively during 2015–2020,

in Italy [551], from 25% to 48% and 21% to 36%, respectively over five years investigation period 2015–2019, and in USA [389], where the resistance to erythromycin and clindamycin was predicted to be 55.2% (n=3497) and 43.9% (n=2783), respectively during 2015–2017. In comparison, low erythromycin and clindamycin resistance rates, 2% and 14%, respectively were observed in GBS in Brazil [552] between 2008 and 2015.

The most recent guidance from the Royal College of Obstetricians and Gynaecologists (RCOG), published in 2017, has recommended that clindamycin should no longer be used as the antibiotic of choice for women with severe allergy to penicillin [3] alternatively, vancomycin is now recommended for this purpose [3]. While, vancomycin remains largely effective and has a potential to treat adult iGBS infections [312, 554], there have been two documented reports of vancomycin resistance identified among USA iGBS isolates, first two cases were recovered during 2011-2012 and last two during 2015-2017, all four isolates harboured *van*G gene [389, 555], with the first report identified *van*G on an element inserted within a conserved chromosomal site [555]. Binding of vancomycin and other glycopeptides to the terminal of peptidoglycan precursors, known as D-Alanine–D-Alanine (D-Ala-D-Ala) prevents transglycosylation and transpeptidation reactions of peptidoglycan synthesis [556]. On the 3′ end of functional *van*G elements, situated the *van*G operon, that mediates vancomycin resistance by synthesising peptidoglycan precursors with C-terminal D-Ala–D-Serine (D-Ala-D-Ser) residues that have low vancomycin affinity while in parallel removing precursors ending with D-Ala-D-Ala [557].

Conjugative transfer is the most efficient way of horizontal gene transfer hence contributing in spreading antibiotic resistance and virulence factors among bacteria [558]. Thus, representing a severe problem in antibiotic treatment, especially of immunosuppressed patients and in intensive care units [559]. The key vehicles of this horizontal gene transfer are a group of mobile genetic elements, termed as conjugative plasmids [558], were extensively studied and discovered earliest due to the ease of their identification, wide distribution in bacteria, massive impact on the spread of pathogenesis and commonly carried antibiotic resistance determinants (ARDs) [433]. While conjugation in gram-negative bacteria has been studied in great detail over the last decades, the transfer mechanisms of

antibiotic resistance plasmids in gram-positive bacteria remained ambiguous [558]. Based on the findings of established literature, it is proposed that the major differences between conjugation in gram-negative and gram-positive bacteria lie in the mechanisms that have evolved to establish cell-cell contact in order to initiate conjugal transfer [558] and two fundamentally different plasmid-mediated conjugative mechanisms are found in gram-positive microorganisms, namely, the mechanism taking place in unicellular gram-positive bacteria, which is functionally similar to that in gram-negative bacteria, and a second type that occurs in multicellular gram-positive bacteria, which seems to be characterized by double-stranded DNA transfer [558]. GBS has exceptionally rare events of horizontal gene transfer and that is why GBS are very unlikely to carry plasmids, the first study published in 1976, confirmed the presence of two plasmids isolated from a GBS isolate and conferring resistance to tetracycline and to chloramphenicol, erythromycin, lincomycin, and pristinamycin, respectively [436], latter studies confirmed the presence of macrolide resistance ***erm***T on small, mobilizable, broad-host-range plasmids in GBS and GAS isolates from the United States [437, 438] and in a *Streptococcus dysgalactiae subsp*. *equisimilis* isolate from Italy [560].

Tetracycline is not used to treat GBS infections because of high resistance profiles observed in GBS population across the world [312, 505, 551]. A contemporary study by Da Cunha *et al* suggested that the extensive use of the broad-spectrum antibiotic tetracycline from 1948 onwards led to the selection and subsequent emergence of few tetracycline resistant clones [1]. This evolutionary bottleneck driven by tetracycline usage resulted in the disappearance of human GBS population diversity and the emergence of few dominant lineages [1]. Among these, CC17 was found to be genetically most homogenous and had relatively lower recombination rates compared to the other CC [1]. The most frequent tetracycline resistance determinants are *tet*M and *tet*O, both encoding ribosomal protection proteins [312]. The *tet*M genes are mostly associated with the conjugative transposons like the Tn*916*/Tn*1545* family [561], which are found in an extremely diverse range of bacteria and known as ever expanding family of MGE [561]. The conjugative transposon are integrated DNA elements that excise themselves to form a covalently closed circular intermediate which can either reintegrate in the same cell (intracellular transposition) or transfer by conjugation to a recipient and integrate into the

recipient's genome (intercellular transposition) [562] while the simple transposon do not have mobility to move between bacteria and are limited to sole movement within bacteria [563]. The term "conjugative transposon" was first used for Tn*916,* which was originally identified in the late 1970's on 18kb conjugative transposon from *Enterococcus faecalis* DS16 [564]. Different nomenclature, such as conjugative transposons, integrative 'plasmids', genomic islands and numerous unclassified elements have been used to define these elements that excise and integrate, such as prophages, and transfer by conjugation [565] however globally, the term conjugative transposon is dropped generally and ICE now is used to define these elements [442]. The Tn*916* show modular arrangement, comprising four modules that are involved in conjugation, excision and integration (recombination), regulation and accessory functions which are not involved in mobility or regulation (**Figure 4.1a**) [472]. Albeit with well-known variations, the recombination module mostly consists of an integrase (*int*-$_{916}$) gene and an excisionase (*xis*-$_{916}$) gene, and the accessory gene is typically the tetracycline (TET) resistance determinant *tet*M [472].

In streptococci, resistances to macrolide and tetracycline are frequently associated due to the localization of the *erm*B and *tet*M genes on the same composite transposons, Tn*916* derivatives [449]. A composite transposon consists of two inverted repeats from two separate transposons moving together as one unit and carrying the DNA between them [566]. The composite transposons have identified in previous GBS study where *erm*B gene was found on the chromosome within a Tn*917*/Tn*916*-like transposon analogous to one identified in *S. pneumoniae* [450]. Whereas, in most recent study majority of GBS isolates positive with *erm*B gene was found to be carried on a complex transposon labelled as Tn*3872* which resulted due to Tn*917* carrying *erm*B insertion into the orf9 of Tn*916* [1, 203] (**Figure 4.1a**). Such Tn*916* variants have previously been reported in collection of macrolide-resistant *S. agalactiae* and *S. pyogenes* [451, 567]. Another widespread Tn*916* family element carrying *tet*M gene in streptococci including GBS is Tn*5801* (~25.8 kb) which was first discovered in *Staphylococcus aureus* and shows a modular organization like that of Tn*916* and has several similar open reading frames (orfs) [466, 568]. However, besides the presence of additional orfs, whose functions are largely unknown, DNA identities are rather low except in the case of *tet*M (97.7%) [568]. In particular, the

recombination module differs from that of Tn*916*, as it lacks the *xis* gene and shows very low DNA identity (38.6%) between *int*-$_{5801}$ and *int*-$_{916}$ **(Figure 4.1b)** [568]. Many studies have subsequently identified the Tn*5801* and Tn*5801*-like element in GBS carrying *tet*M genes [568, 569], in particularly isolates belonging to CC17 and CC23 [1]. The importance of ICEs associated with *tet*M was highlighted in a recent study [1] which proposed extensive clinical use of tetracycline removed majority of non-resistant GBS types, leaving behind a less diversified GBS with very few STs and CC compared to other bacteria (for example *E.coli* > 10,000 STs) and with the evidence of horizontal gene transfer of Tn*916* observed today in the majority of GBS strains [1].

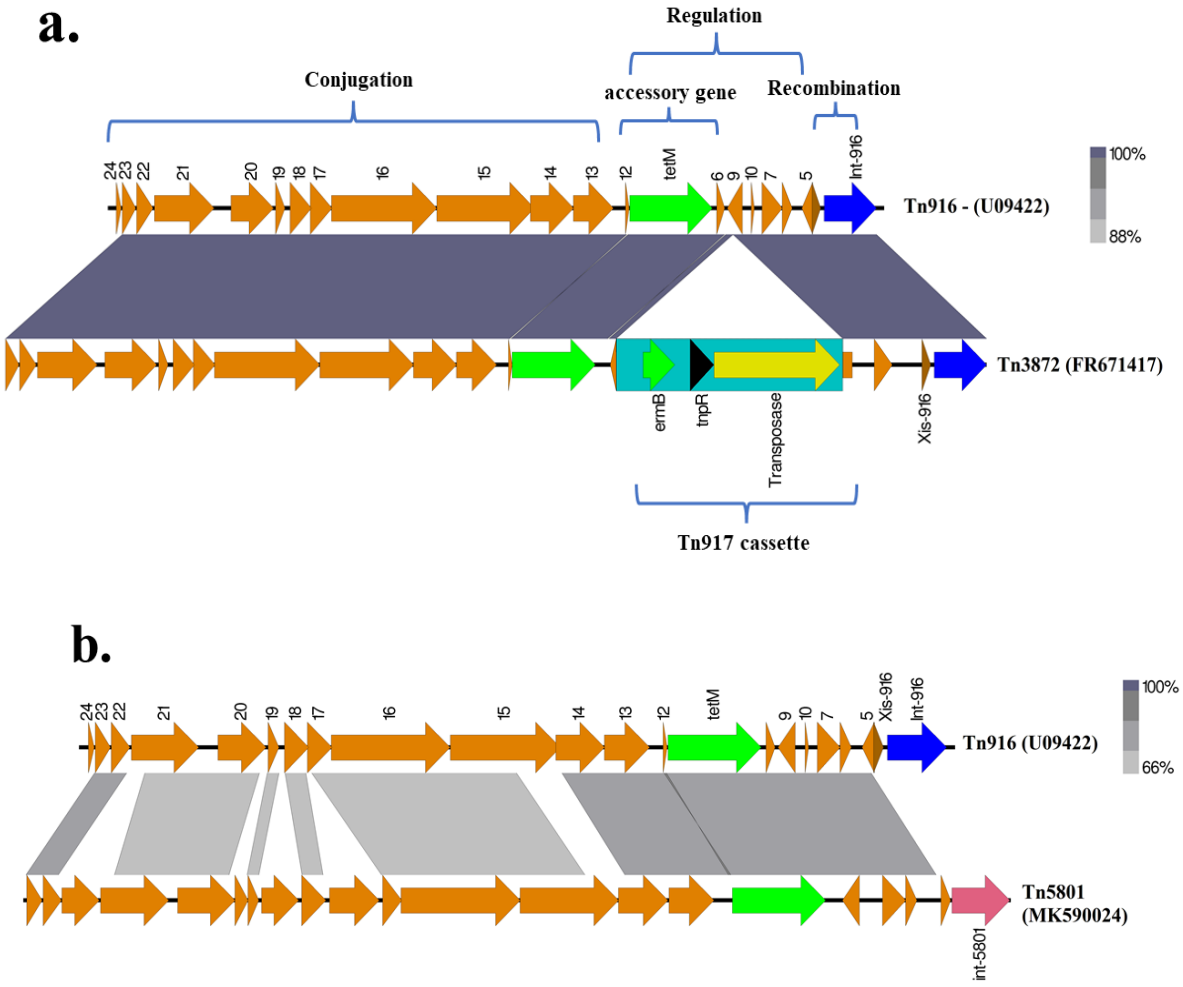



**Figure 4. 1. Modular arrangement of ICE belonging to Tn*916*/Tn*1545* family. a)** Genetic maps and alignment of Tn*916* and Tn*3872*. Blastn is performed using Easyfig v2.2.2 [336] between *Enterococcus faecalis* Tn*916* (NZ_U09422) and *Streptococcus pneumoniae* Tn*3872* (NZ_FR671417), a composite element in which Tn*917* cassette carrying ARG *erm*B is integrated into orf9 of Tn*916*. Sequence

identity is shown in gradient using grey colour and range between 88% and 100%. **b)** Genetic maps and alignment of Tn*916* and Tn*5801*. Blastn is performed using Easyfig v2.2.2 [336] between *Enterococcus faecalis* Tn*916* (NZ_U09422) and *Enterococcus faecalis* Tn*5801*.Sequence identity is shown in gradient using grey colour and range between 66% and 100%. The orfs of all ICEs compared in this figure are represented in different colours arrow where blue arrow shows integrase gene of Tn*916,* brown – excisionase gene of Tn*916,* green -ARG *tet*M and *erm*B, gold – transposase of Tn*917* cassette integrated in Tn*916,* black – *tnp*R gene of Tn*917* cassette and pink – integrase gene of Tn*5801*. The numbers above arrows in Tn*916* represents orfs 5 till 24 and blue horizontal brace represents four modules of Tn*916* namely conjugation, recombination, accessory function, and regulation. The size of Tn*916* is 18kb, Tn*3872* – 22.7kb, and Tn*5801* – 20.3kb, respectively.

In contrast to neonatal GBS infections, the significance of GBS as a cause of severe infections among adults are under investigated, and comprehensive antibiotic susceptibility testing (as opposed to collation of disparate hospital data) and determination of MGEs carrying ARGs in GBS have not been studied in detail. This chapter describes intensive genomic interrogation of WGS for systematic antimicrobial resistance profiling, comparison of phenotypic and genotypic results for antibiotic resistance and identification of MGEs carrying multidrug resistance genes in 193 clinical GBS isolates collected from adult patients in the UK between January 2014 and May 2015.

## 4.3 Results:

### 4.3.1 Phenotypic and genotypic AMR profiling:

All examined GBS isolates (n=193) were found susceptible to ampicillin and vancomycin (**Table 4.1**) which is consistent with a lack of detecting resistance genes for these two antibiotics in this dataset. In total 17 different antibiotic resistance determinants were identified in 193 GBS isolate collection (**Table 4.2a**) among which 40 (20.7%) isolates showed multiple resistance to more than one of the antibiotics including aminoglycosides, macrolides, chloramphenicol, or tetracycline (**Table 4.2b**). By comparison 17 (8.7%) GBS isolates carried no antibiotic resistance genes at all and were phenotypically sensitive to all eight antimicrobials tested.

Ninety percent of GBS isolates (175/193, 90.6%) were predicted to be tetracycline resistant through the presence of *tet*M (79.7%); *tet*O (11.9%); *tet*W (0.5%), *tet*S (0.5%) or *tet*L (0.5%), either alone or in

combination (**Table 4.2a**). However, one of the *tet*M positive isolates - PHEGBS0156 remained tetracycline sensitive on phenotypic and MIC testing, with MIC value of (0.06 µg/mL) for tetracycline. Genomic analysis identified a broken *tet*M present on two contigs in this PHEGBS0156 isolate.

Three isolates (1.5%) were predicted to have chloramphenicol resistance due to the presence of (2/193, 1%) *cat*Q genes (isolates PHEGBS0595 and PHEGBS0608) and one isolate (PHEGBS0738) which carried the *cat*(pC194) gene. Upon phenotypic testing, isolate PHEGBS0608 was found to be sensitive by the Kirby-Bauer disk method; however, chloramphenicol resistance was confirmed when an MIC of 32 µg/mL was determined by antibiotic gradient strip, which is above the threshold concentration of 16 µg/mL as set out in the CLSI 2017 guidelines [485].

Forty-one (21.2%) GBS isolates showed resistance to macrolides mediated by at least one macrolide resistance gene, the most common was *erm*A (n=16; 39%) and *erm*B (n=16; 39%), followed by *mef*A/*msr*D (n=10; 24.3%), *lsa*C (n=3; 7.3%), *erm*T (n=1, 0.5%), and *lnu*C (n=1; 0.5%). Forty of these isolates additionally carried tetracycline resistance genes (1 *tet*S, 13 *tet*O and 29 *tet*M), while only one isolate (PHEGBS0084) was found to carry a macrolide resistance gene (*erm*A) alone with no additional antibiotic resistance gene. Using antibiotic susceptibility testing, all macrolide gene determinant positive isolates except (PHEGBS0207-*erm*B positive) isolate was found resistant to erythromycin and/or clindamycin. The MIC value for clindamycin and erythromycin in isolate PHEGBS0207 were 0.25 µg/mL and 4 µg/mL respectively, supporting clindamycin susceptibility and erythromycin resistance (as per the MIC value observed using erythromycin gradient strip) in PHEGBS0207. Overall, resistance rates to erythromycin and clindamycin were observed in 20.2% (n=39) and 11.8% (n=23) GBS isolates respectively, among which 10.8% (n=21) were resistant to both. Out of 41 macrolide resistant isolates, 33 (80.4%) carried at least one macrolide gene (*erm*A, *erm*B or *erm*T) among which 42.4% (n=14/33) of isolates had a positive D-test, indicating inducible clindamycin resistance (**Table 4.1**).

Aminoglycoside resistance was predicted in (8/193, 4.1%) GBS isolates due to the presence of *aph*(3'-III) and *ant*(6-Ia) (7/8, 87.5%), *aad*E (4/8, 50%) and *aac*(6')-*aph*(2") (1/8 ,12.5%) determinants. On phenotypic testing, only one isolate (PHEGBS0139) with *aac*(6')-*aph*(2") gene (that also carried *lnu*C, *erm*A and *tet*M*)* was found to be resistant to gentamicin (**Table 4.1).**

### 4.3.2 SNP investigation for mechanism of fluoroquinolone resistance using ClustalW:

Four isolates (PHEGBS0139, PHEGBS0559, PHEGBS0595 and PHEGBS0608) were also found resistant to levofloxacin (MIC>256 mg/L) mediated by double point mutations (Ser81Leu for *gyr*A, Ser79Tyr for *par*C) in 2% (n=4/193) GBS isolates (**Table 4.1**) with latter two isolates (PHEGBS0595 and PHEGBS0608) also carrying single point H221Y non-resistance polymorphism in *par*E. This H221Y substitution in *par*E was additionally observed in 28 isolates which remained sensitive to levofloxacin upon AST, indicating this altered amino acid does not contribute to the resistance phenotype. In addition, 65 isolates with different non-resistance polymorphisms in *gyr*B at (V149I; n=58, E93G; n=2, G490A; n=1, I252V; n=1, P559S; n=2, and S536L; n=1) were observed and all these isolates showed sensitivity to levofloxacin with AST (**Table 4.1**).

**Table 4. 1. Summary of 193 GBS isolates used in this study.** For 17 months surveillance by PHE from January 2014 and May 2015, 193 GBS isolates collected from invasive and non-invasive sites of adults in United Kingdom were available for the analysis. X indicates no presence of acquired antibiotic genes or quinolone resistance-determining regions (QRDR) mutations in chromosomal *gyr*A and *par*C genes. Resistance genes of antibiotic classes including Tetracycline, Aminoglycosides, Macrolide Lincosamides and Streptogramin B and Chloramphenicol were predicted by Resfinder v2.1 [504] and phenotypic testing was performed by disk diffusion and MIC test strips. Antibiotic susceptibility testing was performed on all samples against eight antibiotics and D-test was employed on erythromycin and/or clindamycin resistant GBS isolates harbouring *erm*A, *erm*B or *erm*T genes to check for D-factor. Insertion sites of Tn*916* was identified relative to a reference SS1 strain (NZ_ CP010867.1), Tn*5801* relative to DK-NI-005 genome [1] and Tn*3872* relative to the reference CZ-NI-013 genome [1]. Sample ids in olive colour are showing non-iGBS isolates, while red cells showing antibiotic resistance, green showing intermediate resistance and yellow showing discordance results upon AST or MIC testing. N.R.G stands for no resistance gene; "- " means no inhibition zone identified upon AST and NA indicates specific test not performed.

| Sample ID | ERA number | Specimen | CPS type | ST | CC | QRDR mutation | Antibiotic resistance genes | | | | Antibiotic Susceptibility Test | | | | | | | | D-test | ICE carrying *tet*M and its site of insertion |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | TET | MLSB | AMN | CHL | CN 6 / ≥10 | AMP / ≥24 | CHL ≤17 / ≥21 | VAN / ≥17 | LEV ≤13 / ≥17 | CLI ≤15 / ≥19 | TET ≤18 / ≥23 | ERY ≤15 / ≥21 | | |
| 200684 | ERR3531637 | Blood | Ib | ST1215 | CC1 | X | *tet*M | *X* | X | X | 20 | 34 | 28 | 20 | 23 | 26 | 13 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0450 | ERR1741493 | Blood | Ib | ST1 | CC1 | X | *tet*M | *erm*B | X | X | 14 | 32 | 23 | 17 | 21 | - (>256) | 13 | - (64) | No | Tn*3872* - 971kb |
| PHEGBS0670 | ERR1742049 | Blood | Ib | ST1 | CC1 | X | *tet*M | *X* | X | X | 16 | 38 | 28 | 20 | 24 | 26 | 13 | 26 | NA | Tn*916* - 592kb |
| PHEGBS0100 | ERR1741744 | Blood | II | ST3 | CC1 | X | *tet*W | *X* | X | X | 19 | 28 | 27 | 20 | 22 | 25 | 15 | 23 | NA | NA |
| PHEGBS0549 | ERR1741850 | Blood | II | ST1 | CC1 | X | *tet*M | *X* | X | X | 16 | 30 | 25 | 17 | 20 | 23 | 13 | 26 | NA | Tn*916* - 592kb |
| PHEGBS0568 | ERR1741647 | Blood | II | ST1 | CC1 | X | *tet*M | *X* | X | X | 13 | 32 | 30 | 20 | 22 | 34 | 19 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0627 | ERR1741512 | Blood | II | ST2 | CC1 | X | *tet*M | *X* | X | X | 11 | 30 | 32 | 20 | 22 | 24 | 15 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0084 | ERR1741695 | Blood | IV | ST459 | CC1 | X | *X* | *erm*A | X | X | 19 | 40 | 30 | 21 | 26 | -128 | 35 | - (>256) | No | NA |
| PHEGBS0206 | ERR1741687 | Blood | IV | ST196 | CC1 | X | *tet*O | *erm*B | X | X | 17 | 32 | 30 | 20 | 25 | - (>256) | 16 | -128 | No | NA |
| PHEGBS0266 | ERR1741584 | Blood | IV | ST136 | CC1 | X | *tet*M | *erm*A | X | X | 17 | 34 | 26 | 17 | 24 | -128 | 12 | -128 | No | Tn*916* - 592kb |
| PHEGBS0448 | ERR1741564 | Blood | IV | ST196 | CC1 | X | *tet*M | *X* | X | X | 15 | 36 | 29 | 19 | 25 | 26 | 12 | 29 | NA | Tn*916* - 592kb |
| PHEGBS0463 | ERR1742031 | Blood | IV | ST196 | CC1 | X | *tet*M | *X* | X | X | 17 | 36 | 30 | 20 | 25 | 28 | 13 | 29 | NA | Tn*916*-like novel element - unidentified |
| 100414 | ERR2560245 | Blood | V | ST1350 | CC1 | X | *tet*M | *erm*A | X | X | 15 | 34 | 25 | 18 | 20 | -128 | 12 | - (>256) | No | Tn*916* - 592kb |
| 200690 | ERR3531636 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 19 | 34 | 28 | 20 | 21 | 29 | 13 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0041 | ERR1742042 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 18 | 32 | 28 | 19 | 22 | 25 | 14 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0081 | ERR1741497 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 17 | 36 | 26 | 19 | 22 | 26 | 13 | 38 | NA | Tn*916* - 592kb |
| PHEGBS0082 | ERR1741501 | Blood | V | ST1 | CC1 | X | *tet*M | *erm*A | X | X | 17 | 32 | 27 | 19 | 25 | 24 | 15 | 15 (128) | yes | Tn*916* - 592kb |
| PHEGBS0090 | ERR1742102 | CSF | V | ST1314 | CC1 | X | *tet*M | *X* | X | X | 17 | 36 | 32 | 18 | 25 | 30 | 14 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0098 | ERR1741573 | Blood | V | ST1 | CC1 | X | *tet*M | *erm*A | X | X | 18 | 36 | 27 | 20 | 24 | 18 (32) | 14 | - (>256) | yes | Tn*916* - 592kb |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0127 | ERR1741367 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 19 | 32 | 28 | 20 | 23 | 26 | 10 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0128 | ERR1742140 | Blood | V | ST1 | CC1 | X | *tet*M | *erm*B | X | X | 17 | 27 | 25 | 19 | 23 | -128 | 22 | - (>256) | No | Tn*3872* - 971kb |
| PHEGBS0164 | ERR1741749 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 19 | 36 | 30 | 22 | 24 | 28 | 15 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0306 | ERR1741485 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 20 | 34 | 32 | 21 | 28 | 28 | 15 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0360 | ERR1742032 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 15 | 32 | 26 | 19 | 24 | 34 | 18 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0368 | ERR1742016 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 17 | 34 | 26 | 19 | 23 | 28 | 13 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0373 | ERR1741478 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 17 | 32 | 28 | 19 | 23 | 26 | 14 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0378 | ERR1741669 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 17 | 36 | 27 | 18 | 26 | 27 | 12 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0467 | ERR1741981 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 20 | 36 | 32 | 20 | 32 | 34 | 21 | 34 | NA | Tn*916* - 592kb |
| PHEGBS0493 | ERR1741913 | Blood | V | ST1217 | CC1 | X | *tet*M | *erm*B | X | X | 16 | 32 | 27 | 18 | 22 | - (>256) | 12 | -128 | No | Tn*3872* - 971kb |
| PHEGBS0520 | ERR1741361 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 16 | 34 | 27 | 19 | 26 | 29 | 14 | 29 | NA | Tn*916* - 592kb |
| PHEGBS0599 | ERR1741659 | Blood | V | ST1 | CC1 | X | *tet*M | *erm*B | X | X | 15 | 34 | 25 | 17 | 20 | - (>256) | 11 | - (>256) | No | Tn*3872* - 971kb |
| PHEGBS0648 | ERR1741752 | Blood | V | ST1 | CC1 | X | *tet*M | *X* | X | X | 30 | 35 | 32 | 29 | 22 | 34 | 12 | 36 | NA | Tn*916* - 592kb |
| PHEGBS0446 | ERR1741823 | Blood | VI | ST1 | CC1 | X | *X* | *X* | X | X | 15 | 33 | 28 | 18 | 24 | 24 | 28 | 31 | NA | NA |
| PHEGBS0533 | ERR1741445 | Placenta | VI | ST14 | CC1 | X | *tet*M | *X* | X | X | 15 | 30 | 24 | 18 | 19 | 23 | 12 | 34 | NA | Tn*916*-like novel element - unidentified |
| PHEGBS0662 | ERR1741427 | Blood | VI | ST1 | CC1 | X | *tet*S | *erm*B | *ant*(6-Ia), *aph*(*3''-III*) | X | 16 | 34 | 26 | 18 | 24 | -128 | 13 | -32 | No | NA |
| PHEGBS0042 | ERR1742087 | Blood | IX | ST130 | CC130/ CC1216 | X | *X* | *X* | X | X | 17 | 36 | 29 | 21 | 24 | 26 | 31 | 30 | NA | NA |
| PHEGBS0552 | ERR1741992 | Blood | IX | ST1216 | CC130/ CC1216 | X | *X* | *X* | X | X | 15 | 34 | 28 | 20 | 28 | 26 | 31 | 30 | NA | NA |
| PHEGBS0592 | ERR1741448 | Blood | IX | ST130 | CC130/ CC1216 | X | *X* | *X* | X | X | 18 | 34 | 26 | 19 | 23 | 26 | 29 | 27 | NA | NA |
| PHEGBS0575 | ERR1741521 | pus | Ib | ST1220 | CC17 | X | *tet*M | *X* | X | X | 18 | 36 | 28 | 19 | 25 | 25 | 16 | 29 | NA | Tn*916* - 592kb |
| 200693 | ERR3531627 | Blood | III | ST1219 | CC17 | X | *tet*M | *X* | X | X | 18 | 32 | 25 | 19 | 17 | 25 | 14 | 30 | NA | Tn*5801* - 476kb |
| 200694 | ERR3531635 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 13 | 32 | 23 | 17 | 21 | 26 | 13 | 28 | NA | Tn*916* - 592kb |
| 200696 | ERR3531628 | Blood | III | ST550 | CC17 | X | *tet*M | *X* | X | X | 14 | 36 | 32 | 21 | 26 | 29 | 15 | 32 | NA | Tn*5801* - 476kb |
| 200697 | ERR3531629 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 11 | 34 | 24 | 20 | 26 | 23 | 13 | 30 | NA | Tn*5801* - 476kb |
| 200710 | ERR3589625 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 17 | 32 | 28 | 20 | 25 | 28 | 16 | 35 | NA | Tn*916* - 592kb |
| PHEGBS0052 | ERR1741861 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 20 | 23 | 27 | 19 | 25 | 27 | 17 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0066 | ERR1741366 | Blood | III | ST17 | CC17 | X | *tet*M | *erm*A | X | X | 17 | 36 | 30 | 18 | 24 | 20 | 14 | 12 (64) | yes | Tn*916* - 592kb |
| PHEGBS0117 | ERR1741635 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 16 | 36 | 32 | 20 | 26 | 26 | 15 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0144 | ERR1741588 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 15 | 32 | 28 | 19 | 25 | 23 | 14 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0207 | ERR1741997 | Blood | III | ST17 | CC17 | X | *tet*O | *erm*B | *ant*(6-Ia), *aph*(*3''-III*), *aad*E | X | 21 | 34 | 30 | 20 | 23 | 24 (0.25) | 17 | 28 (4) | yes | NA |
| PHEGBS0219 | ERR1741590 | Blood | III | ST550 | CC17 | X | *tet*M | *X* | X | X | 21 | 42 | 32 | 22 | 28 | 22 | 19 | 24 | NA | Tn*5801* - 476kb |
| PHEGBS0237 | ERR1742117 | Placenta | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 16 | 34 | 36 | 19 | 25 | 25 | 16 | 30 | NA | Tn*5801* - 476kb |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0243 | ERR1741633 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 20 | 38 | 32 | 21 | 26 | 28 | 16 | 32 | NA | Tn*5801* - 476kb |
| PHEGBS0275 | ERR1742034 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 17 | 34 | 27 | 20 | 24 | 25 | 16 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0286 | ERR1741880 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 16 | 32 | 38 | 20 | 25 | 26 | 17 | 34 | NA | Tn*5801* - unidentified |
| PHEGBS0288 | ERR1741473 | Blood | III | ST17 | CC17 | X | *tet*O | *mef*A/*msr*D, *erm*B | *ant*(6-Ia), *aph*(*3'-III*), *aad*E | X | 14 | 30 | 28 | 19 | 23 | - (>256) | 15 | 8 (64) | No | NA |
| PHEGBS0318 | ERR1741792 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 16 | 33 | 22 | 17 | 20 | 21 | 15 | 26 | NA | Tn*5801* - 476kb |
| PHEGBS0355 | ERR1741387 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 18 | 32 | 24 | 20 | 24 | 25 | 16 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0380 | ERR1741867 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 18 | 29 | 26 | 19 | 23 | 24 | 16 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0383 | ERR1741449 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 17 | 32 | 30 | 19 | 21 | 24 | 14 | 29 | NA | Tn*5801* - 476kb |
| PHEGBS0389 | ERR1741939 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 15 | 34 | 22 | 18 | 23 | 22 | 13 | 26 | NA | Tn*916* - 592kb |
| PHEGBS0416 | ERR1741402 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 18 | 32 | 25 | 20 | 23 | 26 | 19 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0429 | ERR1741515 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 17 | 34 | 28 | 20 | 25 | 28 | 16 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0480 | ERR1741862 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 16 | 36 | 27 | 18 | 23 | 25 | 13 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0501 | ERR1741911 | Blood | III | ST1221 | CC17 | X | *tet*M | *X* | X | X | 16 | 34 | 27 | 20 | 25 | 25 | 14 | 29 | NA | Tn*5801* - 476kb |
| PHEGBS0513 | ERR1741701 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 20 | 32 | 26 | 20 | 24 | 25 | 18 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0559 | ERR1741492 | Blood | III | ST17 | CC17 | *gyr*A/*par* C | *tet*M, *tet*O | *erm*B | *ant*(6-Ia) , *aph*(*3'-III*) , *aad*E | X | 12 | 32 | 28 | 17 | - (>256) | -128 | 14 | - (>256) | No | Tn*5801* - unidentified |
| PHEGBS0561 | ERR1741542 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 15 | 34 | 26 | 18 | 24 | 26 | 13 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0577 | ERR1741835 | Blood | III | ST17 | CC17 | X | *tet*O | *mef*A/*msr*D, *erm*B | *ant*(6-Ia), *aph*(*3'-III*) , *aad*E | X | 20 | 38 | 26 | 19 | 25 | -128 | 13 | -128 | No | NA |
| PHEGBS0610 | ERR1742130 | Blood | III | ST17 | CC17 | X | *tet*M | *X* | X | X | 19 | 36 | 26 | 18 | 22 | 22 | 13 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0476 | ERR1741458 | Blood | IV | ST1351 | CC17 | X | *tet*M | *X* | X | X | 11 | 31 | 28 | 20 | 24 | 31 | 13 | 36 | NA | Tn*5801* - 476kb |
| PHEGBS0253 | ERR1741828 | Placenta | II | ST19 | CC19 | X | *X* | *X* | X | X | 19 | 32 | 27 | 19 | 22 | 26 | 30 | 30 | NA | NA |
| PHEGBS0399 | ERR1741537 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 15 | 38 | 24 | 17 | 24 | 30 | 15 | 22 | NA | Tn*916* - 592kb |
| PHEGBS0408 | ERR1741526 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 15 | 34 | 27 | 20 | 26 | 26 | 18 | 27 | NA | Tn*916* - 592kb |
| PHEGBS0464 | ERR1741702 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 14 | 32 | 25 | 17 | 23 | 23 | 17 | 34 | NA | Tn*916* - 592kb |
| PHEGBS0532 | ERR1741494 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 12 | 34 | 30 | 20 | 25 | 25 | 18 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0551 | ERR1741634 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 18 | 34 | 32 | 20 | 26 | 26 | 16 | 29 | NA | Tn*916* - 592kb |
| PHEGBS0554 | ERR1741388 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 14 | 32 | 26 | 20 | 21 | 26 | 16 | 26 | NA | Tn*916* - 592kb |
| PHEGBS0556 | ERR1741948 | Blood | II | ST28 | CC19 | X | *tet*M | *X* | X | X | 17 | 36 | 30 | 21 | 25 | 30 | 17 | 34 | NA | Tn*916* - 592kb |
| PHEGBS0624 | ERR1741580 | Blood | II | ST28 | CC19 | X | *tet*M, *tet*O | *erm*B | X | X | 15 | 36 | 32 | 21 | 22 | - (>256) | 16 | -128 | No | Tn*916* - 592kb |
| 200682 | ERR3531626 | Blood | III | ST27 | CC19 | X | *tet*O | *erm*B | X | X | 19 | 34 | 30 | 23 | 24 | -128 | 22 | -128 | No | NA |
| 200699 | ERR3531631 | Vaginal swabs | III | ST861 | CC19 | X | *tet*O | *lsa*C | X | X | 15 | 30 | 28 | 18 | 24 | 16 (64) | 19 | 32 | NA | NA |
| PHEGBS0071 | ERR1742139 | Blood | III | ST19 | CC19 | X | *tet*M, *tet*O | *lsa*C, *erm*B | X | X | 20 | 38 | 28 | 20 | 24 | - (>256) | 17 | - (>256) | No | Tn*916* - 592kb |
| PHEGBS0132 | ERR1741466 | Blood | III | ST529 | CC19 | X | *tet*O | *X* | X | X | 21 | 40 | 32 | 23 | 30 | 30 | 21 | 36 | NA | NA |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0156 | ERR1741857 | Blood | III | ST19 | CC19 | X | *tet*M | *erm*B | *ant*(6-Ia), *aph(3'-III)* | X | 16 | 34 | 30 | 20 | 24 | 9 (8) | 28 (0.06 ) | 11 (32) | No | Tn*916* - 592kb |
| PHEGBS0252 | ERR1741658 | Blood | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 18 | 32 | 30 | 20 | 22 | 28 | 18 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0296 | ERR1741462 | Blood | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 18 | 34 | 29 | 19 | 24 | 26 | 16 | 38 | NA | Tn*916* - 592kb |
| PHEGBS0359 | ERR1741657 | Blood | III | ST19 | CC19 | X | *X* | *X* | X | X | 17 | 40 | 32 | 20 | 27 | 28 | 31 | 32 | NA | NA |
| PHEGBS0372 | ERR1741524 | Blood | III | ST19 | CC19 | X | *tet*M | *lsa*C | X | X | 16 | 38 | 32 | 21 | 26 | 16 (32) | 20 | 26 | NA | Tn*916* - 592kb |
| PHEGBS0377 | ERR1741496 | Blood | III | ST19 | CC19 | X | *X* | *X* | X | X | 16 | 34 | 26 | 18 | 22 | 25 | 28 | 26 | NA | NA |
| PHEGBS0492 | ERR1741922 | Blood | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 16 | 34 | 28 | 18 | 23 | 24 | 14 | 38 | NA | Tn*916* - 592kb |
| PHEGBS0566 | ERR1741860 | Blood | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 17 | 40 | 26 | 21 | 26 | 27 | 18 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0567 | ERR1741438 | Abscess | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 17 | 36 | 26 | 19 | 24 | 23 | 18 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0593 | ERR1741754 | Blood | III | ST19* | CC19 | X | *tet*M | *X* | X | X | 19 | 36 | 30 | 21 | 25 | 23 | 18 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0598 | ERR1741680 | Blood | III | ST1316 | CC19 | X | *tet*M | *X* | X | X | 15 | 32 | 26 | 19 | 22 | 21 | 15 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0626 | ERR1741728 | Blood | III | ST19 | CC19 | X | *X* | *X* | X | X | 15 | 34 | 30 | 21 | 23 | 30 | 32 | 32 | NA | NA |
| PHEGBS0643 | ERR1741746 | Blood | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 17 | 34 | 30 | 20 | 22 | 25 | 17 | 29 | NA | Tn*916* - 592kb |
| PHEGBS0664 | ERR1742011 | Info not provided | III | ST19 | CC19 | X | *tet*M | *X* | X | X | 15 | 31 | 23 | 17 | 20 | 21 | 14 | 25 | NA | Tn*916* - 592kb |
| PHEGBS0139 | ERR1741954 | Blood | V | ST19 | CC19 | *gyr*A/ *par*C | *tet*M | *lnu*C, *erm*A | *aac*(6') - *aph*(2") | X | - (>256) | 40 | 30 | 19 | - (>256) | 22 (0.015) | 16 | 16 (4) | yes | Tn*916* - 592kb |
| PHEGBS0595 | ERR1741483 | Blood | V | ST19 | CC19 | *gyr*A/ *par*C | *tet*M | *mefA/msr*D, *erm*A | X | *cat*Q | 18 | 38 | 11 (48) | 18 | - (>256) | 26 | 15 | 10 (64) | yes | Tn*916* - 592kb |
| PHEGBS0608 | ERR1741534 | Blood | V | ST19 | CC19 | *gyr*A/ *par*C | *tet*M | *mefA/msr*D, *erm*A | X | *cat*Q | 15 | 26 | 25 (32) | 17 | - (>256) | 25 | 14 | 9 (64) | yes | Tn*916* - 592kb |
| PHEGBS0639 | ERR1741887 | Blood | V | ST110 | CC19 | X | *tet*M | *mefA/msr*D | X | X | 21 | 36 | 32 | 20 | 22 | 30 | 18 | 17 (16) | NA | Tn*916* - 592kb |
| PHEGBS0657 | ERR1741517 | Blood | V | ST19 | CC19 | X | *tet*M | *X* | X | X | 15 | 38 | 30 | 20 | 26 | 28 | 15 | 32 | NA | Tn*916* - 592kb |
| PHEGBS0738 | ERR1741902 | Tissue | V | ST19 | CC19 | X | *tet*M | *erm*B | *ant*(6-Ia), *aph(3'-III)* | *cat*(*pC194*) | 19 | 36 | 13 (48) | 19 | 24 | 14 (32) | 18 | 15 (32) | No | Tn*916* - 592kb |
| PHEGBS0080 | ERR1741859 | Blood | II | ST22 | CC22/CC 1213 | X | *tet*M | *X* | X | X | 11 | 36 | 28 | 19 | 22 | 27 | 21 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0398 | ERR1741711 | Blood | II | ST1213 | CC22/CC 1213 | X | *tet*M | *X* | X | X | 15 | 33 | 28 | 19 | 24 | 26 | 15 | 27 | NA | Tn*916* - 592kb |
| 200702 | ERR3531630 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 30 | 26 | 19 | 22 | 25 | 20 | 29 | NA | Tn*5801* - 476kb |
| PHEGBS0044 | ERR1742051 | Blood | Ia | ST23 | CC23 | X | *tet*M | *erm*A | X | X | 19 | 34 | 26 | 20 | 24 | - (>256) | 17 | -128 | No | Tn*5801* - 476kb |
| PHEGBS0048 | ERR1741806 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 21 | 38 | 32 | 21 | 21 | 30 | 21 | 34 | NA | Tn*5801* - 476kb |
| PHEGBS0049 | ERR1741385 | Blood | Ia | ST144 | CC23 | X | *tet*M | *X* | X | X | 20 | 38 | 24 | 20 | 24 | 28 | 15 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0060 | ERR1741459 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 22 | 40 | 27 | 20 | 23 | 26 | 18 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0067 | ERR1741677 | Blood | Ia | ST23 | CC23 | X | *tet*M | *mefA/msr*D | X | X | 21 | 40 | 30 | 23 | 25 | 30 | 18 | 20 (4) | NA | Tn*5801* - 476kb |
| PHEGBS0070 | ERR1742116 | Blood | Ia | ST23 | CC23 | X | *tet*M | *mefA/msr*D | X | X | 21 | 38 | 30 | 20 | 24 | 27 | 16 | 18 (8) | NA | Tn*5801* - 476kb |
| PHEGBS0086 | ERR1741926 | Wound swab | Ia | ST1317 | CC23 | X | *tet*M | *X* | X | X | 19 | 38 | 30 | 19 | 25 | 23 | 13 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0095 | ERR1741918 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 16 | 33 | 30 | 19 | 24 | 26 | 17 | 30 | NA | Tn*5801* - 156kb |
| PHEGBS0106 | ERR1742039 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 20 | 38 | 25 | 19 | 24 | 25 | 17 | 30 | NA | Tn*5801* - 476kb |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0121 | ERR1741801 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 20 | 38 | 22 | 19 | 24 | 26 | 18 | 30 | NA | Tn*5801* - 265kb |
| PHEGBS0122 | ERR1741539 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 32 | 26 | 20 | 25 | 27 | 19 | 25 | NA | Tn*5801* - 156kb |
| PHEGBS0123 | ERR1742074 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 20 | 30 | 24 | 18 | 22 | 23 | 17 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0134 | ERR1741692 | Blood | Ia | ST1218 | CC23 | X | *tet*M | *X* | X | X | 16 | 40 | 27 | 19 | 25 | 26 | 16 | 31 | NA | Tn*5801* - 476kb |
| PHEGBS0170 | ERR1742029 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 18 | 38 | 27 | 19 | 24 | 27 | 18 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0188 | ERR1741906 | Blood | Ia | ST24 | CC23 | X | *tet*M | *X* | X | X | 21 | 40 | 30 | 20 | 23 | 26 | 16 | 30 | NA | Tn*5801* - 265kb |
| PHEGBS0193 | ERR1742119 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 20 | 38 | 29 | 20 | 23 | 28 | 17 | 31 | NA | Tn*5801* - 476kb |
| PHEGBS0230 | ERR1741957 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 32 | 26 | 17 | 22 | 24 | 16 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0246 | ERR1741679 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 18 | 32 | 26 | 20 | 21 | 24 | 15 | 31 | NA | Tn*5801* - 476kb |
| PHEGBS0248 | ERR1742038 | Blood | Ia | ST23 | CC23 | X | *tet*M | *mefA/msr*D | X | X | 17 | 38 | 28 | 19 | 23 | 26 | 15 | 18 (8) | NA | Tn*5801* - 180kb |
| PHEGBS0265 | ERR1741375 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 18 | 36 | 27 | 20 | 24 | 28 | 18 | 31 | NA | Tn*5801* - 476kb |
| PHEGBS0267 | ERR1741548 | Tissue | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 16 | 40 | 30 | 18 | 24 | 30 | 17 | 30 | NA | Tn*5801* - 539kb |
| PHEGBS0270 | ERR1742082 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 14 | 34 | 30 | 19 | 25 | 32 | 16 | 30 | NA | Tn*5801* - 515kb |
| PHEGBS0283 | ERR1741523 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 38 | 32 | 21 | 25 | 26 | 14 | 32 | NA | Tn*5801* - 539kb |
| PHEGBS0287 | ERR1742120 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 21 | 42 | 30 | 18 | 21 | 31 | 20 | 32 | NA | Tn*5801* - 476kb |
| PHEGBS0295 | ERR1741889 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 32 | 26 | 19 | 22 | 24 | 15 | 29 | NA | Tn*5801* - 476kb |
| PHEGBS0300 | ERR1741553 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 36 | 24 | 17 | 23 | 22 | 14 | 27 | NA | Tn*5801* - 476kb |
| PHEGBS0336 | ERR1741747 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 32 | 23 | 17 | 25 | 26 | 14 | 24 | NA | Tn*5801* - 476kb |
| PHEGBS0367 | ERR1742111 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 38 | 28 | 19 | 26 | 26 | 18 | 32 | NA | Tn*5801* - 476kb |
| PHEGBS0401 | ERR1741671 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 16 | 36 | 25 | 18 | 24 | 25 | 15 | 29 | NA | Tn*5801* - 476kb |
| PHEGBS0407 | ERR1741602 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 11 | 34 | 23 | 17 | 22 | 23 | 15 | 27 | NA | Tn*5801* - 476kb |
| PHEGBS0447 | ERR1741785 | Blood | Ia | ST1065 | CC23 | X | *tet*M | *X* | X | X | 20 | 36 | 25 | 18 | 25 | 26 | 17 | 32 | NA | Tn*5801* - 539kb |
| PHEGBS0465 | ERR1741442 | Aortic valve | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 39 | 27 | 20 | 22 | 27 | 17 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0503 | ERR1741579 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 15 | 32 | 32 | 19 | 22 | 34 | 13 | 32 | NA | Tn*5801* - 476kb |
| PHEGBS0512 | ERR1741958 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 18 | 36 | 27 | 19 | 23 | 28 | 19 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0518 | ERR1741606 | Blood | Ia | ST498 | CC23 | X | *tet*M | *mefA/msr*D | X | X | 19 | 34 | 25 | 18 | 24 | 21 | 14 | 19 (128) | NA | Tn*5801* - 156kb |
| PHEGBS0527 | ERR1742059 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 36 | 25 | 18 | 22 | 25 | 14 | 26 | NA | Tn*5801* - 476kb |
| PHEGBS0539 | ERR1741599 | Amniotic membrane | Ia | ST23 | CC23 | X | *tet*M | *erm*T | X | X | 17 | 28 | 22 | 18 | 20 | 20 | 14 | -128 | yes | Tn*5801* - 476kb |
| PHEGBS0547 | ERR1741621 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 40 | 30 | 19 | 24 | 26 | 15 | 29 | NA | Tn*5801* - 539kb |
| PHEGBS0589 | ERR1741794 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 16 | 36 | 27 | 19 | 23 | 29 | 15 | 39 | NA | Tn*5801* - 476kb |
| PHEGBS0616 | ERR1742012 | Blood | Ia | ST24 | CC23 | X | *tet*M | *X* | X | X | 17 | 38 | 28 | 17 | 25 | 24 | 15 | 30 | NA | Tn*5801* - 476kb |
| PHEGBS0618 | ERR1741842 | Blood | Ia | ST1214 | CC23 | X | *tet*O | *X* | X | X | 16 | 34 | 28 | 17 | 28 | 25 | 16 | 34 | NA | NA |
| PHEGBS0623 | ERR1741456 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 13 | 32 | 23 | 17 | 20 | 21 | 14 | 24 | NA | Tn*5801* - 161kb |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0625 | ERR1742108 | Blood | Ia | ST23 | CC23 | X | *tet*M | *mefA/msr*D | X | X | 18 | 40 | 30 | 19 | 24 | 25 | 15 | 19 (16) | NA | Tn*5801* - 180kb |
| PHEGBS0654 | ERR1741549 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 19 | 36 | 34 | 18 | 23 | 26 | 16 | 32 | NA | Tn*5801* - 161kb |
| PHEGBS0658 | ERR1741901 | Blood | Ia | ST23 | CC23 | X | *tet*M | *X* | X | X | 15 | 36 | 23 | 17 | 20 | 21 | 14 | 23 | NA | Tn*5801* - 476kb |
| PHEGBS0663 | ERR1741436 | Blood | Ia | ST144 | CC23 | X | *tet*M | *X* | X | X | 17 | 40 | 30 | 20 | 28 | 26 | 10 | 32 | NA | Tn*5801* - 592kb |
| PHEGBS0047 | ERR1741500 | Blood | Ia | ST24 | CC23 | X | *tet*M | *X* | X | X | 17 | 40 | 34 | 20 | 25 | 27 | 15 | 32 | NA | Tn*5801* - 476kb |
| PHEGBS0152 | ERR1742041 | Blood | Ia | ST24 | CC23 | X | *tet*M | *X* | X | X | 18 | 30 | 25 | 18 | 23 | 25 | 17 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0176 | ERR1742030 | Blood | III | ST23 | CC23 | X | *X* | *X* | X | X | 20 | 36 | 28 | 19 | 25 | 28 | 30 | 30 | NA | NA |
| PHEGBS0524 | ERR1741696 | Blood | III | ST23 | CC23 | X | *tet*M | *X* | X | X | 17 | 36 | 30 | 19 | 29 | 24 | 15 | 34 | NA | Tn*916* - 592kb |
| PHEGBS0337 | ERR1741620 | Blood | V | ST498 | CC23 | X | *tet*M | *X* | X | X | 17 | 34 | 26 | 19 | 24 | 25 | 14 | 28 | NA | Tn*5801* - 476kb |
| PHEGBS0630 | ERR1741525 | Blood | V | ST498 | CC23 | X | *tet*M | *X* | X | X | 16 | 32 | 23 | 18 | 24 | 23 | 13 | 26 | NA | Tn*5801* - 476kb |
| PHEGBS0194 | ERR1741932 | Blood | Ia | ST7 | CC8/CC10 | X | *tet*M | *X* | X | X | 17 | 32 | 30 | 20 | 21 | 27 | 15 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0320 | ERR1741605 | Blood | Ia | ST7 | CC8/CC10 | X | *X* | *X* | X | X | 13 | 32 | 24 | 18 | 22 | 21 | 28 | 23 | NA | NA |
| 141439 | ERR3531632 | Blood | Ib | ST10 | CC8/CC10 | X | *tet*M | *X* | X | X | 15 | 30 | 25 | 19 | 21 | 24 | 14 | 26 | NA | Tn*916* - 592kb |
| 200683 | ERR3531633 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *X* | X | X | 17 | 34 | 29 | 20 | 28 | 30 | 16 | 30 | NA | Tn*916* - 592kb |
| 200698 | ERR3531634 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *X* | X | X | 18 | 36 | 32 | 21 | 24 | 30 | 16 | 32 | NA | Tn*916* - 592kb |
| 200706 | ERR3531625 | Blood | Ib | ST12 | CC8/CC10 | X | *tet*M | *X* | X | X | 15 | 32 | 32 | 22 | 25 | 27 | 14 | 33 | NA | Tn*916* - 956kb |
| PHEGBS0061 | ERR1741474 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M, *tet*O | *X* | X | X | 17 | 36 | 26 | 18 | 26 | 25 | 14 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0068 | ERR1741595 | Blood | Ib | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 12 | 36 | 30 | 20 | 25 | 27 | 15 | 32 | NA | NA |
| PHEGBS0135 | ERR1741673 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *X* | X | X | 16 | 32 | 27 | 17 | 25 | 24 | 14 | 28 | NA | Tn*916* - 592kb |
| PHEGBS0145 | ERR1741818 | Blood | Ib | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 15 | 29 | 27 | 20 | 21 | 24 | 17 | 27 | NA | NA |
| PHEGBS0154 | ERR1741575 | Blood | Ib | ST9 | CC8/CC10 | X | *X* | *X* | X | X | 16 | 32 | 25 | 17 | 23 | 23 | 27 | 27 | NA | NA |
| PHEGBS0382 | ERR1741847 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *X* | X | X | 14 | 32 | 27 | 18 | 25 | 26 | 15 | 26 | NA | Tn*916* - 862kb |
| PHEGBS0393 | ERR1741475 | Blood | Ib | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 15 | 30 | 32 | 21 | 36 | 28 | 19 | 26 | NA | NA |
| PHEGBS0411 | ERR1742019 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *X* | X | X | 20 | 34 | 28 | 20 | 23 | 27 | 15 | 30 | NA | Tn*916* - 592kb |
| PHEGBS0428 | ERR1741389 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *erm*A | X | X | 14 | 33 | 26 | 17 | 24 | 23 | 14 | 13 (32) | yes | Tn*916* - 592kb |
| PHEGBS0555 | ERR1741852 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *erm*A | X | X | 18 | 40 | 29 | 19 | 22 | 19 | 15 | 12 (32) | yes | Tn*916* - 592kb |
| PHEGBS0622 | ERR1741630 | Blood | Ib | ST8 | CC8/CC10 | X | *tet*M | *erm*B | X | X | 15 | 36 | 25 | 19 | 23 | - (>256) | 14 | - (>256) | No | Tn*916* - 592kb |
| PHEGBS0667 | ERR1741656 | Blood | Ib | ST12 | CC8/CC10 | X | *tet*M | *X* | X | X | 16 | 36 | 29 | 19 | 24 | 25 | 15 | 29 | NA | Tn*916* - 592kb |
| PHEGBS0091 | ERR1741407 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *erm*A | X | X | 16 | 32 | 29 | 19 | 26 | 19 | 17 | 12 (64) | yes | NA |
| PHEGBS0092 | ERR1741722 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 16 | 40 | 30 | 21 | 25 | 30 | 16 | 32 | NA | NA |
| PHEGBS0151 | ERR1741944 | Blood | II | ST10 | CC8/CC10 | X | *X* | *X* | X | X | 15 | 32 | 25 | 18 | 23 | 23 | 27 | 26 | NA | NA |
| PHEGBS0171 | ERR1741963 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 21 | 36 | 31 | 21 | 25 | 26 | 16 | 32 | NA | NA |
| PHEGBS0308 | ERR1741488 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *erm*A | X | X | 15 | 30 | 26 | 18 | 20 | 19 | 15 | 14 (32) | yes | NA |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PHEGBS0390 | ERR1741684 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *X* | X | X | 13 | 29 | 21 | 17 | 19 | 20 | 14 | 23 | NA | NA |
| PHEGBS0394 | ERR1741769 | Blood | II | ST9 | CC8/CC10 | X | *X* | *X* | X | X | 15 | 30 | 32 | 18 | 22 | 22 | 30 | 24 | NA | NA |
| PHEGBS0491 | ERR1741535 | Blood | II | ST12 | CC8/CC10 | X | *X* | *X* | X | X | 15 | 36 | 28 | 19 | 20 | 25 | 30 | 28 | NA | NA |
| PHEGBS0509 | ERR1741644 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *erm*A | X | X | 16 | 31 | 28 | 19 | 23 | 19 | 16 | 12 (64) | yes | NA |
| PHEGBS0581 | ERR1741660 | Blood | II | ST12 | CC8/CC10 | X | *tet*O | *erm*A | X | X | 15 | 32 | 28 | 18 | 22 | 14 (8) | 16 | 15 (64) | yes | NA |
| PHEGBS0586 | ERR1741614 | pus | II | ST652 | CC8/CC10 | X | *tet*M, *tet*L | *X* | X | X | 16 | 40 | 28 | 19 | 22 | 26 | 9 | 30 | NA | Tn*916*-like (carrying tetL) - unidentified |
| PHEGBS0072 | ERR1741514 | Blood | Ib | ST15 | singleton | X | *X* | *X* | X | X | 16 | 32 | 32 | 21 | 32 | 27 | 32 | 32 | NA | NA |
| PHEGBS0097 | ERR1741888 | Blood | Ib | ST104 | singleton | X | *tet*O | *X* | X | X | 34 | 40 | 40 | 24 | 26 | 36 | 22 | 40 | NA | NA |
| PHEGBS0153 | ERR1741822 | Blood | II | ST104 | singleton | X | *X* | *X* | X | X | 20 | 38 | 29 | 20 | 23 | 26 | 30 | 30 | NA | NA |
| PHEGBS0054 | ERR1742070 | Blood | III | ST283 | singleton | X | *X* | *X* | X | X | 19 | 24 | 26 | 18 | 21 | 24 | 30 | 29 | NA | NA |
| PHEGBS0483 | ERR1741662 | Blood | III | ST1212 | singleton | X | *tet*M | *X* | X | X | 12 | 30 | 21 | 17 | 23 | 20 | 11 | 24 | NA | Tn*916* - 592kb |
| PHEGBS0635 | ERR1741616 | Throat swab | V | ST26 | singleton | X | *tet*M | *X* | X | X | 16 | 34 | 30 | 20 | 23 | 28 | 15 | 32 | NA | Tn*916* - 592kb |

CN stands for gentamicin (120µg); AMP - ampicillin (10µg), CHL - chloramphenicol (30µg), VAN - vancomycin (30µg), LEV - levofloxacin (5µg), CLI - clindamycin (2µg), TET - tetracycline (3µg) and ERY - erythromycin (15µg).

Value in bracket is showing MIC in µg/mL and value without bracket showing zone of inhibition in cm.

**Table 4. 2.a) Summarizing ResFinder v2.1 predicted antibiotic genetic variants in 193 GBS isolates and b) presenting GBS isolates harbouring multiple antibiotic resistance gene determinants.**

| **a. Antibiotic resistance genes predicted by ResFinder v2.1.** | | |
|---|---|---|
| **Antibiotic (no. of positive resistant isolates)** | **Antibiotic Genetic determinants** | **No. of isolates carrying antibiotic resistance genes (%)** |
| Tetracycline | *tet*M | 154 (79.7%) |
| | *tet*O | 23 (11.9%) |
| | *tet*W | 1 (0.5%) |
| | *tet*S | 1 (0.5%) |
| | *tet*L | 1 (0.5%) |
| MLS - Macrolide, Lincosamide and Streptogramin B | *er*mA | 16 (8.2%) |
| | *erm*B | 16 (8.2%) |
| | *er*mT | 1 (0.5%) |
| | *mef*A/*msr*D | 10 (5.1%) |
| | *lnu*C | 1 (0.5%) |
| | *lsa*C | 3 (1.5%) |
| Phenicol | *cat*Q | 2 (1%) |
| | *cat*(pC194) | 1 (0.5%) |
| Aminoglycosides | *aac*(6') -*aph*(2") | 1 (0.5%) |
| | *ant*(6-Ia) | 7 (3.6%) |
| | *aad*E | 4 (2%) |
| | *aph*(3'-III) | 7 (3.6%) |
| **b. Multiple antibiotic resistance genes carrying isolates predicted by ResFinder v2.1.** | | |
| **Cross-resistance genes** | **no. of isolates (%)** | |
| MLS +Tetracycline | 30 (15.5%) | |
| Aminoglycoside + MLS + Tetracycline | 7 (3.6%) | |
| MLS + Phenicol + Tetracycline | 2 (1%) | |
| Aminoglycoside + MLS + Phenicol + Tetracycline | 1 (0.5%) | |

### 4.3.3 Transpeptidase amino acid substitutions in PBPs

All 193 GBS isolates were positive for PBP types (1a, 1b, 2a, 2b and 2x) with few non-resistance polymorphisms as all these strains remained sensitive to ampicillin with AST. No MIC was determined for penicillin in any of the isolates as none of them carried substitution at hotspot regions. The similarity percentage and substitutions information with respect to reference genome for each PBP types determined in 193 GBS isolates is described in **Table 4.3**.

**Table 4. 3. Summary of Penicillin binding proteins (PBPs) found in 193 GBS UK adult isolates.** The reference PBPs of *S. agalactiae* 2603V/R (NC_004116) were used for this analysis.

| Penicillin binding protein types (PBPs) | Reference protein id for PBPs | similarity (%) to reference PBPs | No. of isolates | Substitutions* |
|---|---|---|---|---|
| **1a** | NP_687333.1 | 100 | 6 | NIL |
| | | 99.98 | 1 | S685N + N702D + G739- + N740- + G741- +V746A |
| | | 98.93 | 2 | A27T + A734V + G739- + N740- + G741- + N742- + N743- + V746A |
| | | 98.8 | 1 | (-)54S + -55S + S473N + N702D + N740- + G741- + N742- + N743- + V746A |
| | | 99.73 | 77 | G739- + N740- |
| | | 99.6 | 6 | R662H + G739- +N740- |
| | | 99.47 | 14 | N740- + G741- + N742- + G743- |
| | | | 1 | G739- + N740- + G741- + N742- |
| | | 99.33 | 2 | N740- + G741- + N742- + G743- + F544V |
| | | 99.2 | 3 | S685N + G739- + N740- + G741- + N742- + V746A |
| | | 99.07 | 76 | N740- + G741- + N742- + N743- + V746A |
| | | | 1 | G739- + N740- + G741- + N742- |
| | | 99.02 | 3 | A27T + G739S + N740- + G741- + N742- + N743- + V746A |
| | | **Total** | **193** | |
| **1b** | NP_687194.1 | 100 | 62 | NIL |
| | | 99.87 | 59 | A36D (n=48); L41S (n=12) |
| | | 99.74 | 66 | A95D + D505E (n=4); A36D + l46S (n=59); A95D + L41S (n=3) |
| | | 99.73 | 3 | A36D + I46S |
| | | 99.52% | 3 | A95D |
| | | **Total** | **193** | |
| **2a** | NP_689052.1 | 100 | 80 | NIL |
| | | 99.96 | 1 | A581T |
| | | 99.74 | 43 | E63K + N666K + V761L |
| | | 99.61 | 27 | P69L + K316Q + R345Q |
| | | 99.57 | 37 | E63K |
| | | 99.53 | 1 | E63K + V549 |
| | | 99.48 | 2 | P69L + D313N + K316Q + R345Q |
| | | 99.40% | 1 | 1M + P69L + +K316Q + R345K |
| | | 99.35% | 1 | D18N + L317V |
| | | **Total** | **193** | |
| **2b** | | 100 | 8 | NIL |

| | | | | |
|---|---|---|---|---|
| | NP_687780.1 | 99.87 | 168 | 1M- + L2M |
| | | 99.8 | 13 | 1M- + L2M +V80A |
| | | 99.76 | 3 | V80A |
| | | 99.60 | 1 | A69V + V80A + G470S |
| | | **Total** | **193** | |
| **2x** | NP_687322.1 | 100 | 8 | NIL |
| | | 99.78 | 88 | I377V |
| | | 99.20 | 2 | I377V + V510I |
| | | 99.11% | 1 | I377V+ V510I + K570E |
| | | 95.21 | 11 | 01 to 35 - (first 35 a.a's had gaps) + I377V |
| | | 95.08 | 79 | 01 to 35 - (first 35 a.a's had gaps) + I377V + V510I |
| | | 94.95 | 4 | 01 to 35 - (first 35 a.a's had gaps) + Y366H + I377V |
| | | **Total** | **193** | |

Note that "NIL" represents no substitution observed and "_" represents gap observed.
* None of the identified substitutions has been reported as causing resistance to penicillin in GBS.

### 4.3.4 Mobile genetic elements (MGEs) carrying MDR genes in GBS strains

In this study, we presented the AMR-associated mobilome, their insertion hotspot and association with CCs of 193 GBS isolates. The 182 isolates clearly fell into five main CCs: CC23 (n=53, 27.4%), CC1 (n=35, 18.1%), CC19 (n=33, 17%), CC8/CC10 (n=29, 15%) and CC17 (n=32, 16.5%), while remaining isolates belonged to CC130/CC1216 (n=3, 1.5%), CC22/CC1213 (n=2, 1%) and singletons (n=6, 3.1%) (**Table 4.1**). For 4 fluoroquinolones resistant isolates including three serotype CC19 V/ST19 (PHEGBS0139, PHEGBS0595 and PHEGBS0608) and one CC17 serotype III/ST17 (PHEGBS0559) isolate, resistance was mediated by somatic mutations in *gyr*A/*par*C (described in earlier section; **Table 4.1**); for all other antimicrobial resistance acquisition of MGE carrying ARG was responsible. A total of 41 MGEs were identified in this collection of isolates, either carrying single or multiple drug resistant determinants other than *tet*M which was mainly carried by Tn*916*/Tn*1545* family. The most common target gene (TG) favouring integration of 41 MGEs was the 3' end of *rum*A found in 29 (70.7%) MGEs, followed by integration identified at 15 nucleotides prior to the 3' end of *rpl*L in 9 (21.9%) MGEs and less commonly at the 3' end of *rps*I found in 3 (7.3%) MGEs. The MGEs for seven ARGs found in 14

isolates belonging to different CC (**Figure 4.4 a-e)** and their respective target genes (TGs) were remains unidentified, either due to the presence of ARG at the end of the contig or relatively on a smaller assembled contig.

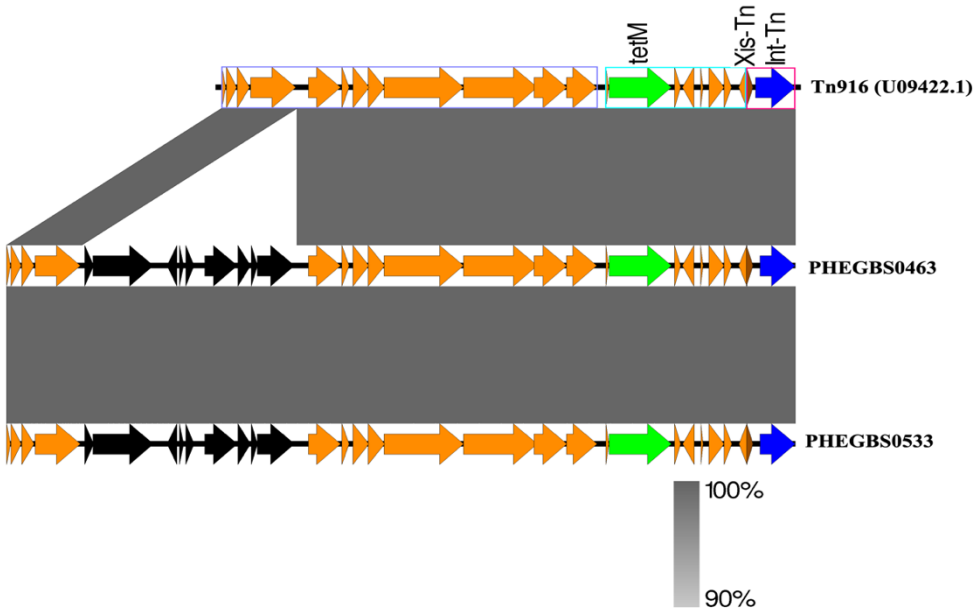



**Figure 4. 2. New Tn*916*-like element identified in two CC1 isolates.** Blastn was performed between reference *Enterococcus faecalis* transposon Tn*916* element (NZ_U09422.1), PHEGBS0463 (IV/ST196) and PHEGBS0533 (VI/ST14) contig carrying Tn*916*-like element using Easyfig v2.2.2 [336]. The different modules of Tn*916* are highlighted using coloured rectangles where red shows recombination module, green – regulation and purple - conjugation module. Novel Tn*916*-like element is 24.3kb long and comprised of 31 open reading frames (orfs) with additional 9 orfs inserted between orf20 and orf21 within conjugation module of typical Tn*916*. Grey areas between orfs denote nucleotide identities with a gradient representing 90% (light grey) to 100% (dark grey) identity. Green colour is used to represent tetracycline antibiotic resistance gene (*tet*M), brown for Tn*916* excisionase and blue for Tn*916* integrase gene.

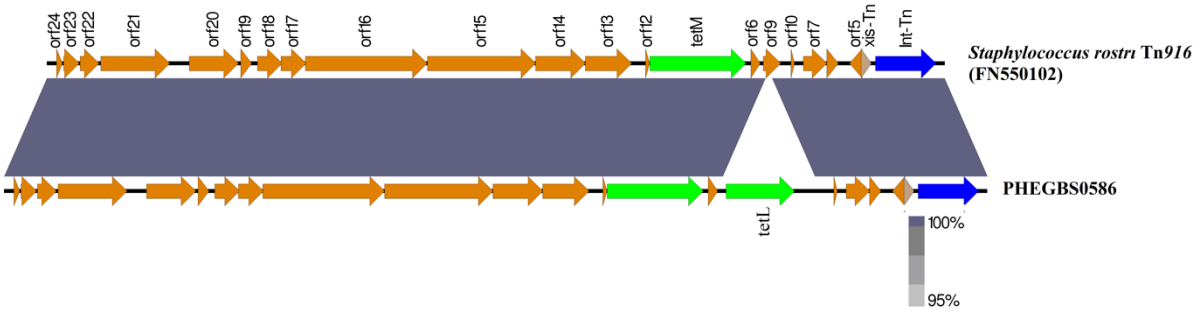
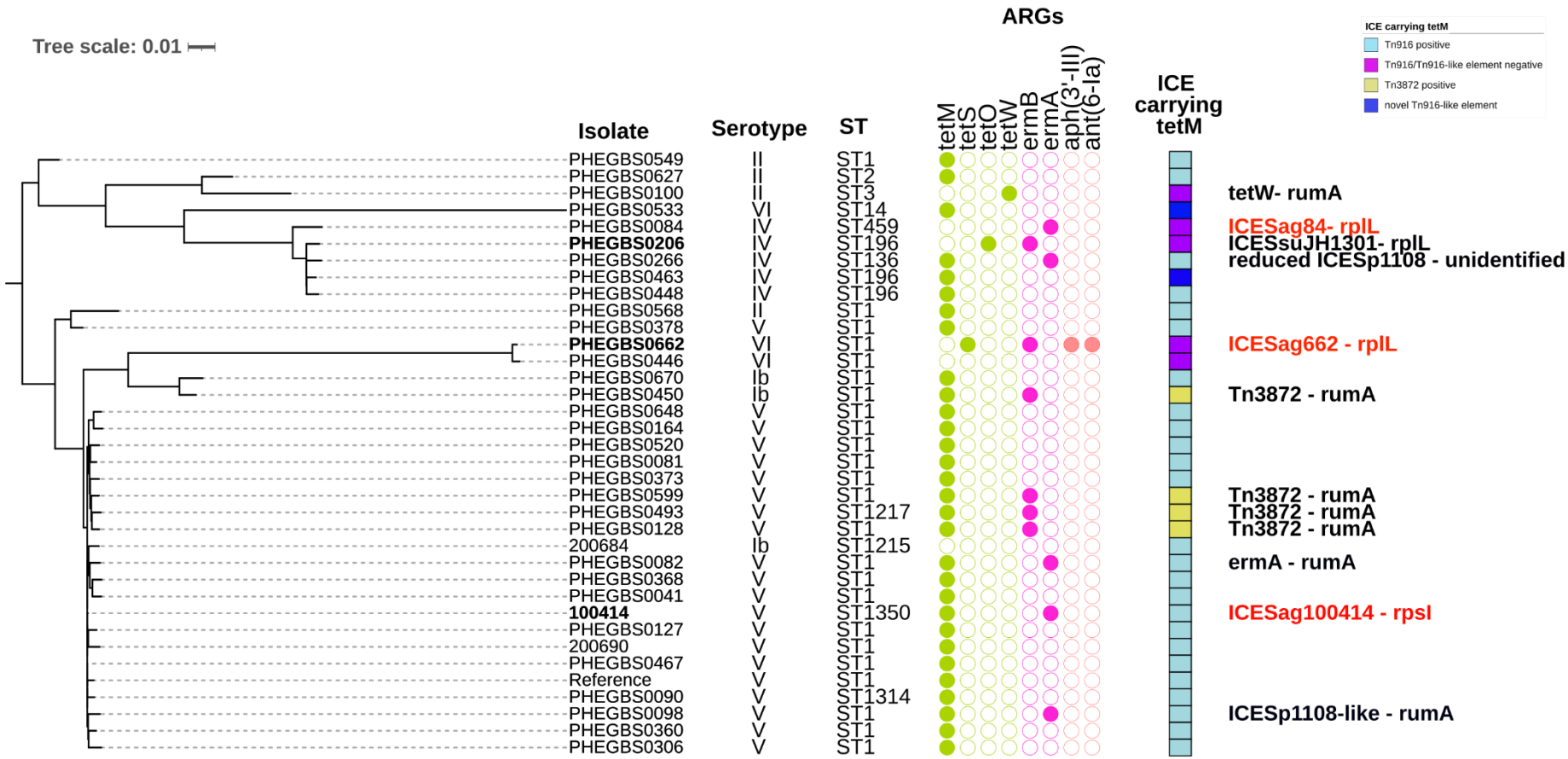
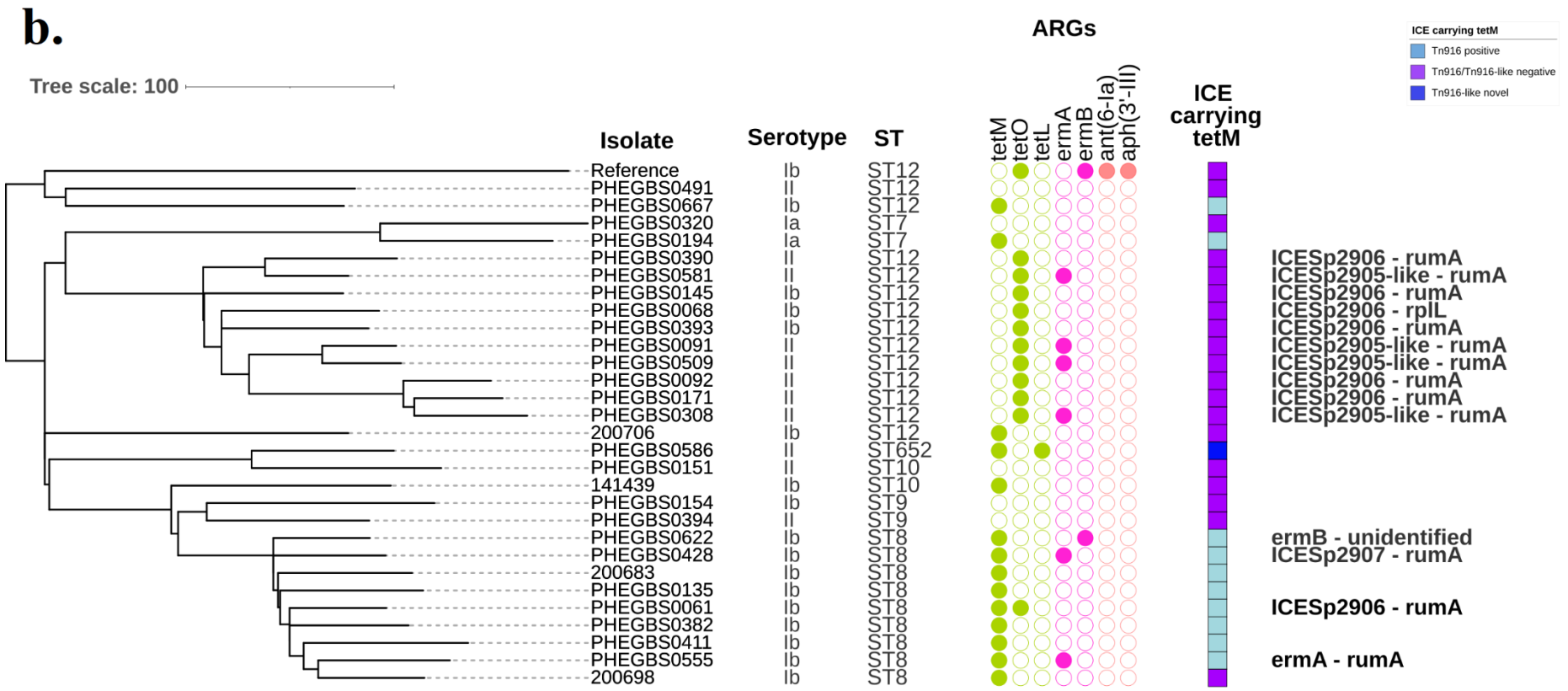
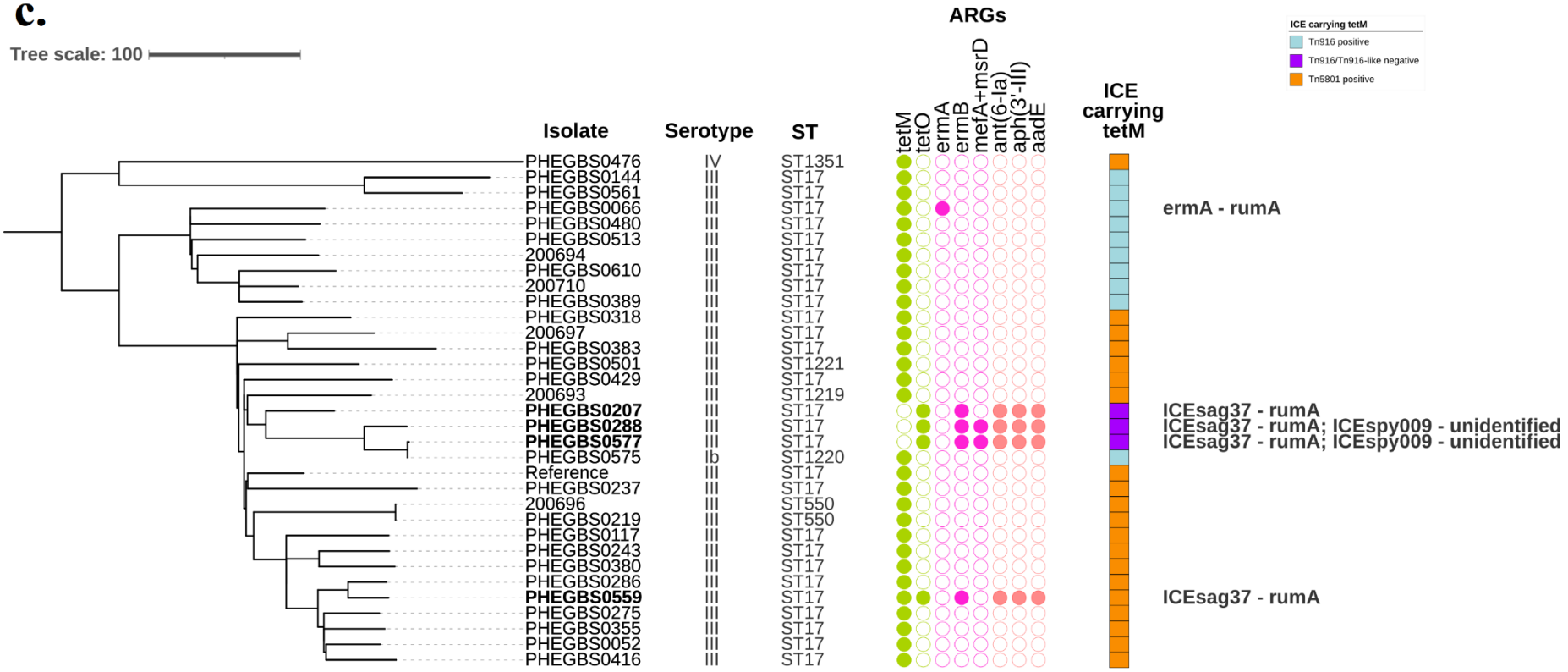



**Figure 4. 3. Tn*916* variant observed in CC8/CC10 isolate - PHEGBS0586 (II/ST652).** Blastn was performed between reference *Staphylococcus rostri* transposon Tn*916* (NZ_FN550102) and PHEGBS0586 contig carrying tetracycline resistance genes (*tet*M and *tet*L) using Easyfig v2.2.2 [336].

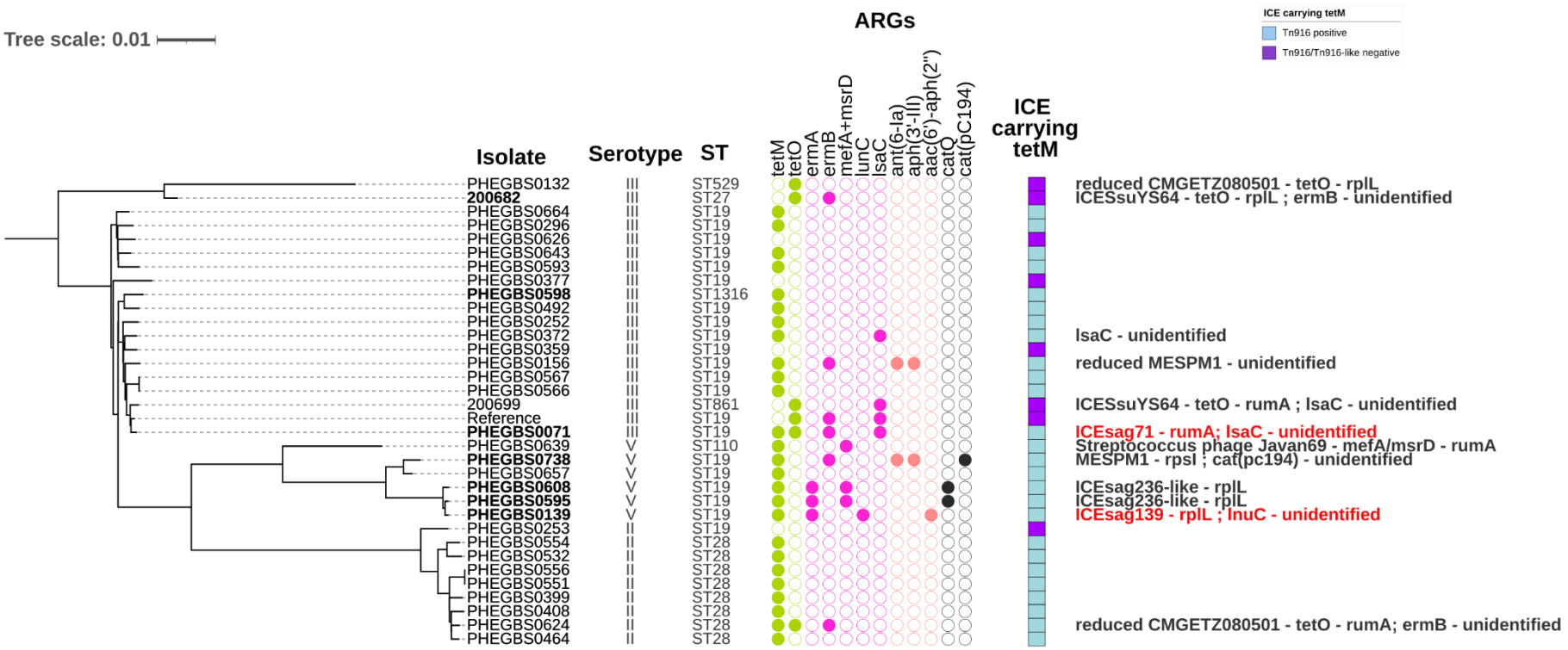
Grey areas between orfs denote nucleotide identities with a gradient representing 95% (light grey) to 100% (dark grey) identity. Green colour is used to represent tetracycline resistance genes (*tet*M and *tet*L), light brown for Tn*916* excisionase and blue for Tn*916* integrase gene.


a.
Tree scale: 0.01
Isolate
Serotype
ST
ARGs
tetM
tetS
tetO
tetW
ermB
ermA
aph(3'-III)
ant(6-Ia)
ICE carrying tetM
ICE carrying tetM
Tn916 positive
Tn916/Tn916-like element negative
Tn3872 positive
novel Tn916-like element
PHEGBS0549 II ST1
PHEGBS0627 II ST2
PHEGBS0100 II ST3
PHEGBS0533 VI ST14
PHEGBS0084 IV ST459
PHEGBS0206 IV ST196
PHEGBS0266 IV ST136
PHEGBS0463 IV ST196
PHEGBS0448 IV ST196
PHEGBS0568 II ST1
PHEGBS0378 V ST1
PHEGBS0662 VI ST1
PHEGBS0446 VI ST1
PHEGBS0670 Ib ST1
PHEGBS0450 Ib ST1
PHEGBS0648 V ST1
PHEGBS0164 V ST1
PHEGBS0520 V ST1
PHEGBS0081 V ST1
PHEGBS0373 V ST1
PHEGBS0599 V ST1
PHEGBS0493 V ST1217
PHEGBS0128 V ST1
200684 Ib ST1215
PHEGBS0082 V ST1
PHEGBS0368 V ST1
PHEGBS0041 V ST1
100414 V ST1350
PHEGBS0127 V ST1
200690 V ST1
PHEGBS0467 V ST1
Reference V ST1
PHEGBS0090 V ST1314
PHEGBS0098 V ST1
PHEGBS0360 V ST1
PHEGBS0306 V ST1
tetW- rumA
ICESag84- rplL
ICESsuJH1301- rplL
reduced ICESp1108 - unidentified
ICESag662 - rplL
Tn3872 - rumA
Tn3872 - rumA
Tn3872 - rumA
Tn3872 - rumA
ermA - rumA
ICESag100414 - rpsI
ICESp1108-like - rumA


b.
Tree scale: 100
Isolate
Serotype
ST
ARGs
tetM
tetO
tetL
ermA
ermB
ant(6-Ia)
aph(3'-III)
ICE carrying tetM
ICE carrying tetM
Tn916 positive
Tn916/Tn916-like negative
Tn916-like novel
Reference Ib ST12
PHEGBS0491 II ST12
PHEGBS0667 Ib ST12
PHEGBS0320 Ia ST7
PHEGBS0194 Ia ST7
PHEGBS0390 II ST12
PHEGBS0581 II ST12
PHEGBS0145 Ib ST12
PHEGBS0068 Ib ST12
PHEGBS0393 Ib ST12
PHEGBS0091 II ST12
PHEGBS0509 II ST12
PHEGBS0092 II ST12
PHEGBS0171 II ST12
PHEGBS0308 II ST12
200706 Ib ST12
PHEGBS0586 II ST652
PHEGBS0151 II ST10
141439 Ib ST10
PHEGBS0154 Ib ST9
PHEGBS0394 II ST9
PHEGBS0622 Ib ST8
PHEGBS0428 Ib ST8
200683 Ib ST8
PHEGBS0135 Ib ST8
PHEGBS0061 Ib ST8
PHEGBS0382 Ib ST8
PHEGBS0411 Ib ST8
PHEGBS0555 Ib ST8
200698 Ib ST8
ICESp2906 - rumA
ICESp2905-like - rumA
ICESp2906 - rumA
ICESp2906 - rplL
ICESp2906 - rumA
ICESp2905-like - rumA
ICESp2905-like - rumA
ICESp2906 - rumA
ICESp2906 - rumA
ICESp2905-like - rumA
ermB - unidentified
ICESp2907 - rumA
ICESp2906 - rumA
ermA - rumA


c.
Tree scale: 100
Isolate
Serotype
ST
ARGs
tetM
tetO
ermA
ermB
mefA+msrD
ant(6-Ia)
aph(3'-III)
aadE
ICE carrying tetM
ICE carrying tetM
Tn916 positive
Tn916/Tn916-like negative
Tn5801 positive
PHEGBS0476 IV ST1351
PHEGBS0144 III ST17
PHEGBS0561 III ST17
PHEGBS0066 III ST17
PHEGBS0480 III ST17
PHEGBS0513 III ST17
200694 III ST17
PHEGBS0610 III ST17
200710 III ST17
PHEGBS0389 III ST17
PHEGBS0318 III ST17
200697 III ST17
PHEGBS0383 III ST17
PHEGBS0501 III ST1221
PHEGBS0429 III ST17
200693 III ST1219
PHEGBS0207 III ST17
PHEGBS0288 III ST17
PHEGBS0577 III ST17
PHEGBS0575 Ib ST1220
Reference III ST17
PHEGBS0237 III ST17
200696 III ST550
PHEGBS0219 III ST550
PHEGBS0117 III ST17
PHEGBS0243 III ST17
PHEGBS0380 III ST17
PHEGBS0286 III ST17
PHEGBS0559 III ST17
PHEGBS0275 III ST17
PHEGBS0355 III ST17
PHEGBS0052 III ST17
PHEGBS0416 III ST17
ermA - rumA
ICEsag37 - rumA
ICEsag37 - rumA; ICEspy009 - unidentified
ICEsag37 - rumA; ICEspy009 - unidentified
ICEsag37 - rumA

**d.**


Tree scale: 0.01
ARGs
ICE carrying tetM
Tn916 positive
Tn916/Tn916-like negative
Isolate
Serotype
ST
tetM
tetO
ermA
ermB
mefA+msrD
lunC
lsaC
ant(6-Ia)
aph(3'-III)
aac(6')-aph(2")
catQ
cat(pC194)
ICE carrying tetM
PHEGBS0132 III ST529
200682 III ST27
PHEGBS0664 III ST19
PHEGBS0296 III ST19
PHEGBS0626 III ST19
PHEGBS0643 III ST19
PHEGBS0593 III ST19
PHEGBS0377 III ST19
PHEGBS0598 III ST1316
PHEGBS0492 III ST19
PHEGBS0252 III ST19
PHEGBS0372 III ST19
PHEGBS0359 III ST19
PHEGBS0156 III ST19
PHEGBS0567 III ST19
PHEGBS0566 III ST19
200699 III ST861
Reference III ST19
PHEGBS0071 III ST19
PHEGBS0639 V ST110
PHEGBS0738 V ST19
PHEGBS0657 V ST19
PHEGBS0608 V ST19
PHEGBS0595 V ST19
PHEGBS0139 V ST19
PHEGBS0253 II ST19
PHEGBS0554 II ST28
PHEGBS0532 II ST28
PHEGBS0556 II ST28
PHEGBS0551 II ST28
PHEGBS0399 II ST28
PHEGBS0408 II ST28
PHEGBS0624 II ST28
PHEGBS0464 II ST28
reduced CMGETZ080501 - tetO - rplL
ICESsuYS64 - tetO - rplL ; ermB - unidentified
lsaC - unidentified
reduced MESPM1 - unidentified
ICESsuYS64 - tetO - rumA ; lsaC - unidentified
ICEsag71 - rumA; lsaC - unidentified
Streptococcus phage Javan69 - mefA/msrD - rumA
MESPM1 - rpsI ; cat(pc194) - unidentified
ICEsag236-like - rplL
ICEsag236-like - rplL
ICEsag139 - rplL ; lnuC - unidentified
reduced CMGETZ080501 - tetO - rumA; ermB - unidentified

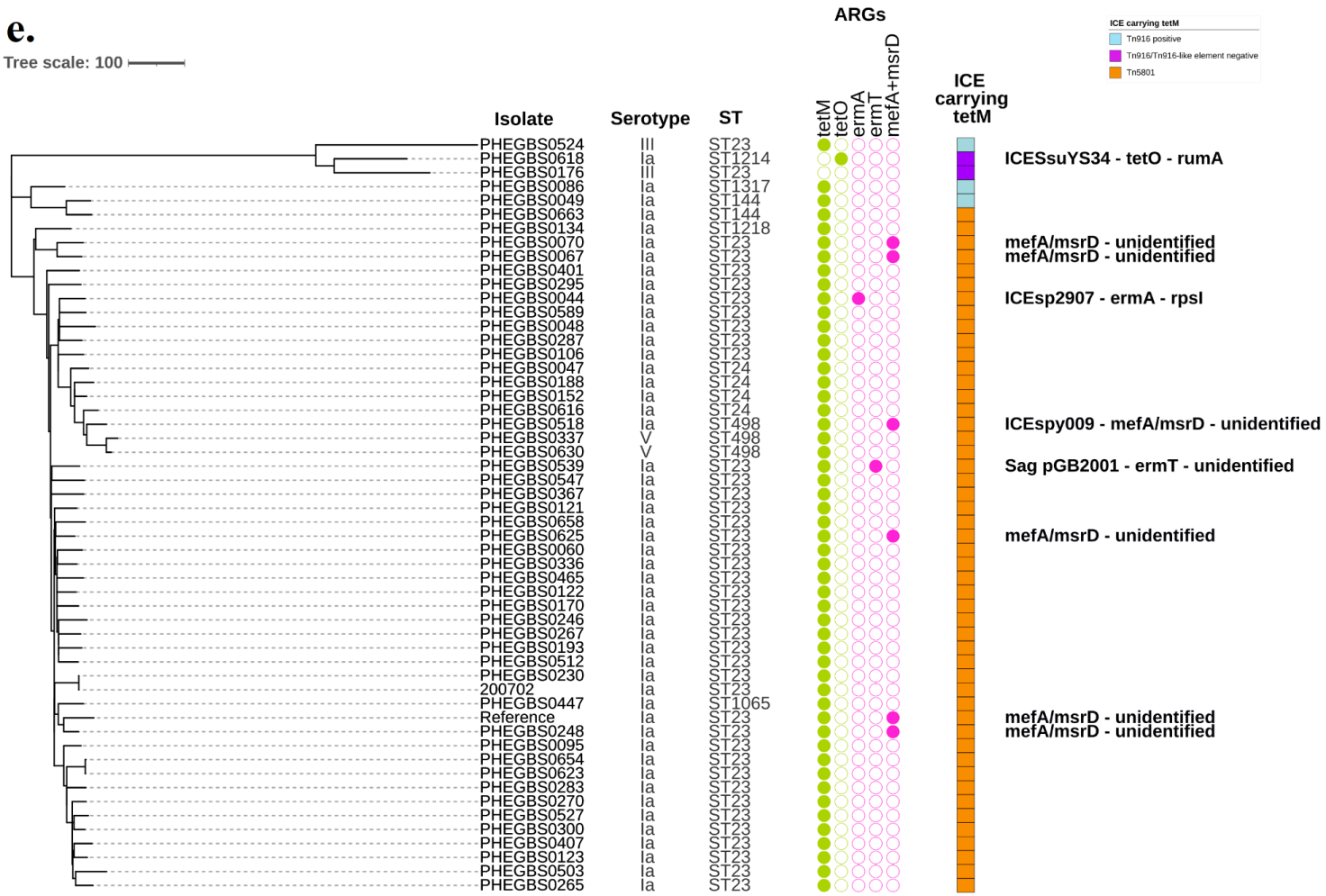

e.
Tree scale: 100
Isolate
Serotype
ST
ARGs
tetM
tetO
ermA
ermT
mefA+msrD
ICE carrying tetM
ICE carrying tetM
Tn916 positive
Tn916/Tn916-like element negative
Tn5801
PHEGBS0524 III ST23
PHEGBS0618 Ia ST1214
PHEGBS0176 III ST23
PHEGBS0086 Ia ST1317
PHEGBS0049 Ia ST144
PHEGBS0663 Ia ST144
PHEGBS0134 Ia ST1218
PHEGBS0070 Ia ST23
PHEGBS0067 Ia ST23
PHEGBS0401 Ia ST23
PHEGBS0295 Ia ST23
PHEGBS0044 Ia ST23
PHEGBS0589 Ia ST23
PHEGBS0048 Ia ST23
PHEGBS0287 Ia ST23
PHEGBS0106 Ia ST23
PHEGBS0047 Ia ST24
PHEGBS0188 Ia ST24
PHEGBS0152 Ia ST24
PHEGBS0616 Ia ST24
PHEGBS0518 Ia ST498
PHEGBS0337 V ST498
PHEGBS0630 V ST498
PHEGBS0539 Ia ST23
PHEGBS0547 Ia ST23
PHEGBS0367 Ia ST23
PHEGBS0121 Ia ST23
PHEGBS0658 Ia ST23
PHEGBS0625 Ia ST23
PHEGBS0060 Ia ST23
PHEGBS0336 Ia ST23
PHEGBS0465 Ia ST23
PHEGBS0122 Ia ST23
PHEGBS0170 Ia ST23
PHEGBS0246 Ia ST23
PHEGBS0267 Ia ST23
PHEGBS0193 Ia ST23
PHEGBS0512 Ia ST23
PHEGBS0230 Ia ST23
200702 Ia ST23
PHEGBS0447 Ia ST1065
Reference Ia ST23
PHEGBS0248 Ia ST23
PHEGBS0095 Ia ST23
PHEGBS0654 Ia ST23
PHEGBS0623 Ia ST23
PHEGBS0283 Ia ST23
PHEGBS0270 Ia ST23
PHEGBS0527 Ia ST23
PHEGBS0300 Ia ST23
PHEGBS0407 Ia ST23
PHEGBS0123 Ia ST23
PHEGBS0503 Ia ST23
PHEGBS0265 Ia ST23
ICESsuYS34 - tetO - rumA
mefA/msrD - unidentified
mefA/msrD - unidentified
ICEsp2907 - ermA - rpsI
ICEspy009 - mefA/msrD - unidentified
Sag pGB2001 - ermT - unidentified
mefA/msrD - unidentified
mefA/msrD - unidentified
mefA/msrD - unidentified

**Figure 4. 4. Phylogenetic relationship between invasive and non-iGBS (n=193) within five major clonal complexes causing disease in adult population in the UK:** Phylogenetic analysis indicates strains belong to the five major clonal complexes identified in this study: **(a) phylogenetic relationship between GBS isolates belonging to CC1:** Thirty-five isolates grouped to CC1 and are represented by a mid-point rooted maximum likelihood (ML) phylogeny tree using 6,299 core SNPs determined using reference sequence SS1 (NZ_CP010867). **(b) phylogenetic relationship between GBS isolates belonging to CC8/CC10.** Twenty-nine isolates belonged to CC8/CC10 and are represented by a mid-point rooted ML phylogeny tree using 9,253 core SNPs called using reference sequence Sag37 (NZ_CP019978.1). **(c) phylogenetic relationship between GBS isolates belonging to CC17.** Thirty-two isolates belonged to CC17 and are represented by a mid-point rooted maximum likelihood phylogeny tree using 2,726 core SNPs relative to reference sequence COH1 (NZ_HG9394456.1). **(d) phylogenetic relationship between GBS isolates belonging to CC19.** Thirty-three isolates belonged to CC19 and are represented by a mid-point rooted maximum likelihood phylogeny tree using 5,560 core SNPs called using reference sequence SG-M25 (NZ_CP021867.1). **(e) phylogenetic relationship between GBS isolates belonging to CC23.** Fifty-three isolates grouped to CC23 and are represented by a mid-point rooted maximum likelihood phylogeny tree using 7,551 core SNPs called from reference sequence FDAARGOS 512 (NZ_CP033822.1) after the removal of recombination blocks. Additional information for isolates presented includes isolate (strain) unique identifier, coloured in black and isolates that are also sequenced by Oxford Nanopore are shown in bold. Next to isolates, the capsular serotyping result, and the sequence type (ST) determined by public MLST database (https://cge.cbs.dtu.dk/services/MLST/). Relevant presence or absence of antimicrobial resistance genes in each CC isolate is shown as a coloured circle, if none of the clonal complex members had a particular gene present, gene name was not added to the description column, otherwise genes were indicated as follows: tetracycline resistance gene (*tet*M, *tet*S, *tet*W, *tet*L or *tet*O gene is present - light green, absent - blank), macrolide resistance genes (*erm*A, *erm*T, *lun*C, *lsa*C or *mef*A+*msr*D; purple, absent - blank), aminoglycoside resistance genes (*ant*(6-Ia), *aph*(3'-III), *aac*(6')-*aph(*2") or *aad*E; gene present – salmon, absent - blank), chloramphenicol resistance genes (*cat*Q or *cat*(pC194) gene present – black, absent – blank, and if not a single isolate in the CC had this gene, that gene column was omitted from visualisation). ICE carrying *tet*M gene in each isolate is shown by a coloured strip where Tn*916* positive isolate is shown by light cyan, Tn*3872* positive isolate – manz*,* Tn*916*/ Tn*916*-like element negative isolate – magenta, and novel Tn*916*-like element identified in this study is shown by blue. MGE carrying ARG(s) other than *tet*M and their target genes (TG) are shown on the right side of the figure in black bold font, whereas novel MGE identified in this study are shown in bold red. In CC1, two isolates did not carry any ARG, whereas among 11 MDR GBS, *rumA* observed as the most common TG for MGE (n=6, including Tn*3872*) insertion followed by *rpl*L (n=3) and *rps*I (n=1). MGE or TG remain undetermined in any isolate are represented as unidentified (n=1). Scale bar, represents a distance of 100 point mutations.

**Table 4. 4. Sumamry of common target genes identified in ARGs associated MGEs.**

| CC | *rum*A insert (n=29) | *rpl*L insert (n=9) | *rps*I insert (n=3) |
|---|---|---|---|
| CC1 | 7 | 3 | 1 |
| CC8/CC10 | 12 | 1 | - |
| CC17 | 5 | - | - |
| CC19 | 4 | 5 | 1 |
| CC23 | 1 | - | 1 |

### 4.3.4.1 ICE carrying *tet*M gene

Tetracycline-resistance was detected in (175/193, 90.6%) GBS isolates, among which the most common tetracycline resistance determinant was *tet*M (154/175, 88%) (**Table 4.2a**). This gene was predominantly carried by Tn*916*/Tn*916*-like elements (n=87, 56.4%) followed by Tn*5801* (n=67,43.5%), found in 4 and 7 different insertion sites, respectively (**Table 4.5**). The Tn*916* (77%) was identified among CC19 (n=26) isolates, CC1 (n=24) and CC8/CC10 (n=13) and, Tn*3872* (n=4) was only observed in CC1 isolates whereas Tn*5801* were exclusively present in CC17 (n=19) and CC23 (n=48) isolates (**Table 4.5**). Specific genomic positions were identified as a hotspot of insertion of these conjugative elements, Tn*916* appear to be inserted at 592kb in majority of isolates (n=78/80), Tn*3872* (n=4) inserted at 971kb, while 476 kb correspond to a hotspot of insertion of Tn*5801* (n=48/67). The target gene for Tn*916* insertion was the 3' end of the *rum*A gene (82/87, 97.6%), and for Tn*5801* it was 5’ end of the *gua*A gene (n=65/67, 97%) (**Table 4.1**).

Among 87 GBS harboured Tn*916*/Tn*916*-like element, 80 isolates carried Tn*916*, four isolates carried Tn*3872* and three isolates carried Tn*916*–like element including two CC1 isolates (PHEGBS0463 and PHEGBS0533) harbouring novel Tn*916*-like element with additional 9 orfs inserted within conjugation module of classical Tn*916* (**Figure 4.2**) and a single CC8/CC10 serotype II ST652 isolate PHEGBS0586 with unique Tn*916* variant harbouring *tet*L and *tet*M gene combinedly in close proximity to each other (**Figure 4.3**). Blastn showed three Tn*916* closest matched references to PHEGBS0586 Tn*916*-like contig, this include *Staphylococcus rostri* transposon Tn*916* (NZ_FN550102) with a query

coverage of 89% and blast identity of 99% followed by *Streptococcus parauberis* Tn*916*-like element (NZ_AB468159.1) and *Enterococcus faecalis* transposon Tn*916* (NZ_U09422.1). Mauve alignment between PHEGBS0586 Tn*916*-like contig and the three blast Tn*916* closest matched references showed similar alignment between all Tn*916* elements however suggested a pickup of *tet*L gene adjacent to *tet*M in PHEGBS0586 Tn*916* variant, which was further verified by mapping PHEGBS0586 reads onto the PHEGBS0586 Tn*916*-like contig.

**Table 4. 5. Characteristics of clones deriving from Tn*916*/ Tn*916*-like element and Tn*5801* insertions.** Insertion genomic positions of Tn*916*, Tn*3872* and Tn*5801* are shown in kb, and were identified using reference SS1 genome (NZ_ CP010867.1), CZ-NI-013 [1] and DK-NI-005 genome [1], respectively.

| CC | Serotype | Tn*916* | Tn*916*-like element | | Tn*5801* | *Tn916*/Tn*916*-like absence |
|---|---|---|---|---|---|---|
| | | | Tn*3872* | novel variants | | |
| **CC1** | V | (n=17) - 592 kb | (n=3) - 971 kb | - | - | - |
| | Ib | (n=2) - 592 kb | (n=1) - 971 kb | - | - | - |
| | II | (n=3) - 592 kb | - | - | - | (n=1) |
| | IV | (n=2) - 592 kb | - | (n=1) novel Tn*916*-like – unidentified | - | (n=2) |
| | VI | - | - | (n=1) novel Tn*916*-like – unidentified | - | (n=2) |
| **CC17** | Ib | (n=1) - 592 kb | - | - | - | - |
| | III | (n=9) - 592 kb | - | - | (n=16) - 476 kb | (n=3) |
| | III | - | - | - | (n=2) - unidentified | - |
| | IV | - | - | - | (n=1) - 476 kb | - |
| **CC19** | II | (N=8) - 592 kb | - | - | - | (n=1) |
| | III | (n=12) - 592 kb | - | - | - | (n=6) |
| | V | (n=6) - 592 kb | - | - | - | - |
| **CC8/CC10** | Ia | (n=1) - 592 kb | - | - | - | (n=1) |
| | Ib | (n=10) - 592 kb | - | - | - | (n=4) |
| | Ib | (n=1) - 862 kb | - | - | - | - |
| | Ib | (n=1) - 956 kb | - | - | - | - |
| | II | - | - | (n=1) Tn*916*-like carrying *tet*L-unidentified | - | (n=10) |
| **CC23** | Ia | (n=2) - 592 kb | - | - | (n=3) - 156 kb | (n=1) |
| | Ia | - | - | - | (n=2) - 161 kb | - |
| | Ia | - | - | - | (n=2) - 180 kb | - |
| | Ia | - | - | - | (n=2) - 265 kb | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Ia | - | - | - | (n=31) - 476 kb | - |
| | Ia | - | - | - | (n=1) - 515 kb | - |
| | Ia | - | - | - | (n=4) - 539 kb | - |
| | Ia | - | - | - | (n=1) - 592 kb | - |
| | III | (n=1) - 592 kb | - | - | - | (n=1) |
| | V | - | - | - | (n=2) - 476 b | - |
| **CC22/CC1213** | II | (n=2) - 592 kb | - | - | - | - |
| **Singleton** | III | (n=1) - 592 kb | - | - | - | - |
| | V | (n=1) - 592 kb | - | - | - | - |

### 4.3.4.2 MGEs carrying ARGs in GBS belonging to different CC

#### 4.3.4.2.1 CC1

A total of (32/35, 91.4%) CC1 isolates were tetracycline resistant including 29 isolates harboured *tet*M gene, and three isolates harboured *tet*O, *tet*S and *tet*W each (**Figure 4.4a**). The *tet*M gene in CC1 isolates was mostly carried by Tn*916* (n=23, 79.3%) followed by Tn*3872* (n= 4, 13.7%; only found in CC1 isolates) and a novel Tn*916*-like element identified in two isolates (PHEGBS0463 and PHEGBS0533) (**Figure 4.4a**). Both Tn*916* (n=23) and Tn*3872* (n=4) were inserted at a single hotspot region 592 kb and 971 kb, respectively across the different genetic background suggesting a horizontal transfer of *tet*M gene (**Table 4.5**). In one isolate – PHEGBS0100, *tet*W was observed on a distinct ICE inserted at the 3' end of *rum*A and showed 99% blastn identity of *tet*W gene to the equivalent gene of *Streptococcus Suis* (NZ_FN396364.1). Three novel ICEs carrying ARG(s) were identified in CC1 isolates (**Figure 4.4a**), while two CC1 isolate didn't carry any ARG at all. ARG (other than *tet*M) associated MGE were mostly inserted at *rum*A (n=7), followed by *rpl*L (n=3) and *rps*I (n=1) (**Figure 4.4a**). The TG for MGE carrying *erm*A in PHEGBS0266 remain unidentified due to the presence of this MGE on a short (14kb) contig.

CC1 had a maximum number of macrolide and tetracycline resistant isolates (11/41, 25.5%), including a single isolate - PHEGBS0662 which also harboured aminoglycoside resistance determinants (**Figure 4.4a**). The *erm*A macrolide gene was found in (5/11, 45.4%) macrolide resistant CC1 isolates including four isolates (100414, PHEGBS0082, PHEGBS0098 and PHEGBS0266) that carried *erm*A and *tet*M in combination and a single isolate – PHEGBS0082 carried *erm*A alone (**Figure 4.4a**). Isolate PHEGBS0098 carried *erm*A by a variant *Streptococcus pyogenes* ICE*Sp*1108 (NZ_FR691054.1; ~48kb long) which was inserted adjacent to *rum*A and had a blastn identity of 95% and query coverage of 70%. Another isolate PHEGBS0266 carried reduced ICE*Sp*1108, terminating at IS1216 on a 35.6kb long contig with a blast identity of 95% and query coverage of 85%, respectively (**Figure 4.5**). In isolate – 100414, *erm*A was observed on a novel ICE*Sag*100414 of ~76kb size, inserted at the 3' end of *rps*I,

this ICE also carried three copies of site-specific recombinase at the right flanking region (**Figure 4.6**). The ICE*Sag*100414 resulted from the fusion of ICE*Sp*1108 (~37.7kb; NZ_FR691054.1) carrying *erm*A and *Erysipelothrix rhusiopathiae* ICE*Er*0106 (~8.7kb; NZ_MG812141.1) showing a (blastn identity: 86% and query coverage: 52%) and (blastn identity: 95% and query coverage: 46%), respectively. However, no putative core sites were identified for ICE*Sag*100414. Another novel ICE (ICE*Sag*84; 71.9kb long) harbouring a single *erm*A macrolide gene was observed in isolate – PHEGBS0084 resulted from the fusion of *Streptococcus Suis* ICE*Ssu*YS209 (NZ_ MK211823.1) (~20.7kb, blastn identity: 95% and query coverage: 32%) and ICE*Sp*1108 (NZ_FR691054.1) (~43.7kb, blastn identity 9% and query coverage: 55%) (**Figure 4.7**). The 15bp putative core sites (DR: GTTACTCTTAAATAA) were identified at flanking regions of the ICE*Sag*84 due to its insertion 15 nt prior to the 3' end of *rpl*L gene.

Six out of eleven macrolide resistant CC1 isolates carried *erm*B gene including four isolates - (PHEGBS0128, PHEGBS0493, PHEGBS450 and PHEGBS0599) carried *erm*B and *tet*M genes in combination via Tn*3872* (**Figure 4.4a**), one isolate – PHEGBS0206 carried *erm*B and *tet*O together by a variant *Streptococcus Suis* ICE*Ssu*JH1301 (NZ_KX077887.1) (blastn identity: 95%, query coverage: 84%, *rpl*L insert) and an isolate-PHEGBS0662 carried *erm*B, *tet*S, *ant*(6-Ia) and *aph*(3'-III) on a novel ICE*Sag*662 **(Figure 4.8)**. This novel 73.6 kb composite ICE*Sag*662 has homology suggesting recombination of two ICEs originally described in *Streptococcus Suis* including ~58.6kb region from ICESsuJH1308-1 (NZ_KX077886; blast identity: 95% and query coverage: 78%) and carrying resistance genes to multiple antibiotics *(tetS, ermB, aadE and aph*(3′-III)) and heavy metal resistance genes encoding cadmium efflux system accessory protein, cadmium-transporting ATPase, arsenical pump-driving ATPase, arsenical resistance operon repressor and arsenical resistance operon trans-acting protein **(Figure 4.8)**. While the remainder of the ICE*Sag*662 (~14kb) shows homology to ICE*Ssu*JH1308-2 (NZ_KX077884) including an integrase gene, relaxase and a potential virulence factor (*NisK*) - a two-component signal transduction system observed earlier in ICE*Sag*37 [477]. The ICE*Sag*662 is inserted 15 nt prior to *rpl*L and flanked by identical *att*L and *att*R recombination sites of 15 bp (TTATTTAAGAGTAAC).

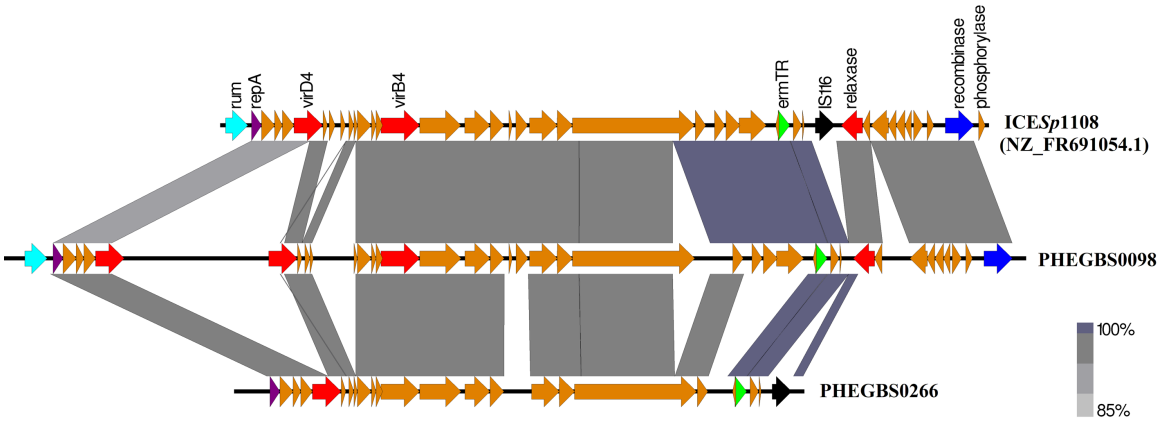
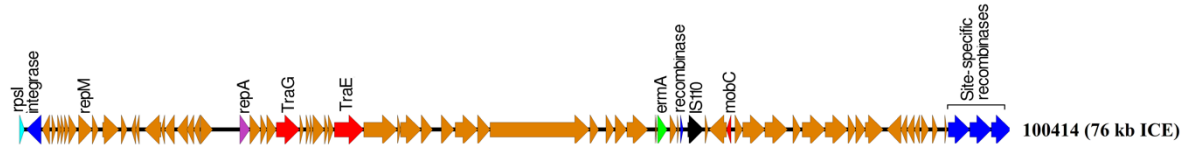



**Figure 4. 5. Comparison of ICE*Sp*1108-like element found in two CC1 isolates:** Blastn was performed using Easyfig v2.2.2 [336] between reference *Streptococcus pyogenes* ICE*Sp*1108 (NZ_FR691054.1) and PHEGBS0098 - serotype V/ST1 and PHEGBS0266 serotype IV/ST136 isolates. ICE orfs are represented in different colours where turquoise shows target gene – *rum*, purple – replication initiator protein – *rep*A, red – conjugal proteins, green – antibiotic resistance gene - *erm*A, black – insertion sequence - IS116 and blue - integrase / recombinase genes. Sequence identity is shown a s a gradient with a range between 85% (light grey) to 100% (dark grey). The size of the ICE*Sp*1108 is 48kb.




`

**Figure 4. 6. Novel ICE*Sag*100414 identified in CC1 isolate.** A novel ICE carrying *erm*A gene is identified in serotype V ST1350 GBS isolate belonging to CC1. Easyfig v2.2.2 [336] was used to generate this figure. Different colours are used to emphasize important orfs of ICE where turquoise shows integration site – *rps*I, blue - integrase / recombinase genes, purple – replication initiator protein – *rep*A, red – conjugal proteins, green – antibiotic resistance gene and black - insertion sequence.

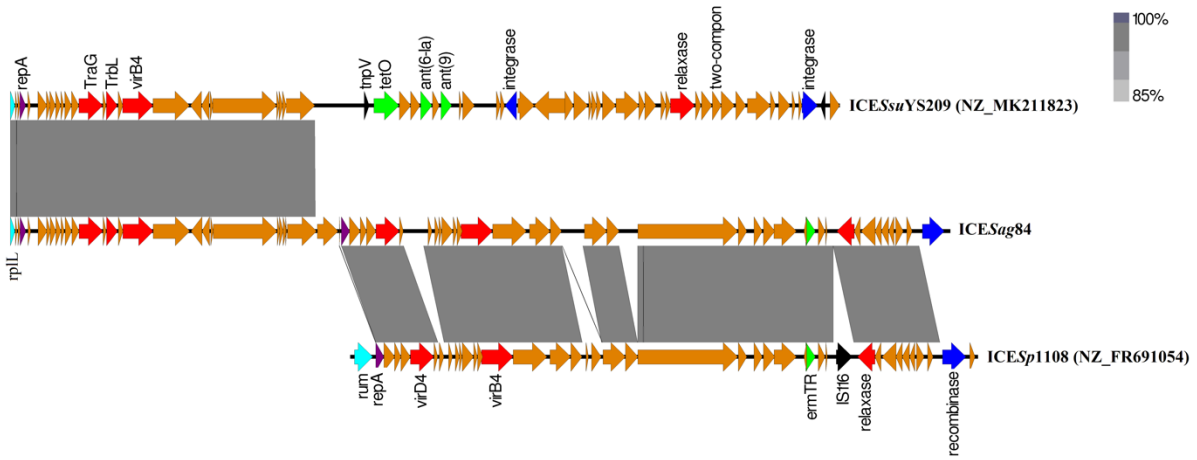



**Figure 4. 7. Newly identified ICE*Sag*87 in CC1 isolate.** Nucleotide alignment is performed between *Streptococcus Suis* ICE*Ssu*YS209 (NZ_MK211823.1), *Streptococcus pyogenes* ICE*Sp*1108 (NZ_FR691054.1) and the serotype IV/ST459 isolate – PHEGBS0084 using Easyfig v2.2.2 [336]. The ICE*Ssu*YS209 shared (~20.7kb) region and ICE*Sp*1108 shared (~43.7kb) region carrying macrolide resistance gene (*erm*TR) to form a resultant ICE*Sag*84 that is 71.9kb in size and comprised 65 orfs in total. Coding sequences (CDSs) of the ICE*Sag*84 are depicted by arrows using different colours where antibiotic resistance determinants are coloured in green, integrase in blue, integration TG (*rpl*L) in turquoise, replication initiator gene in purple – *rep*A, conjugative genes in red, insertion sequence – IS116 in black and all other backbone orfs of ICE*Sag*84 in orange.

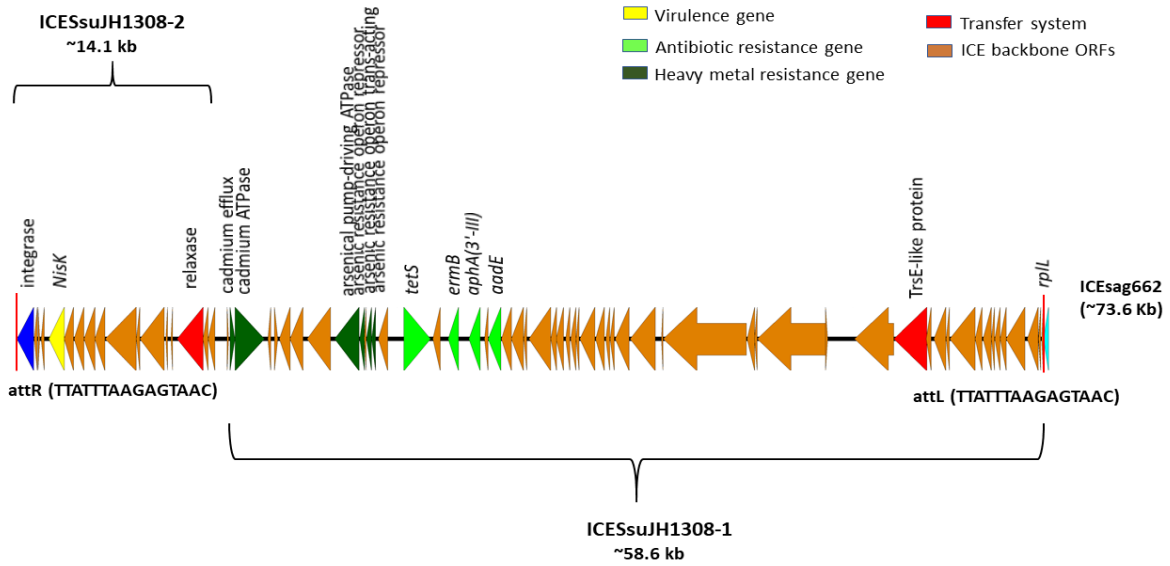



**Figure 4. 8. Schematic representation of the newly identified integrative and conjugative element (ICE*Sag*662) in CC1 isolate.** Novel ICE was identified in serotype VI ST1 isolate - PHEGBS0662, carrying multidrug resistance genes. Coding sequences (CDSs) of the ICE*Sag*662 are depicted by arrows using different colours where antibiotic resistance determinants are coloured in green, integrase in blue, virulence gene in yellow, integration TG (*rpl*L) in turquoise, conjugative elements in red, heavy metal resistance system in dark green and all other backbone orfs of ICE*Sag*662 in orange. The genomic size of ICEsag662 is ~73.6 kb. Easyfig v2.2.2 [336] was used to generate this figure.

#### 4.3.4.2.2 CC8/CC10

Out of 29 CC8/CC10 isolates, (n=25, 86.2%) were tetracycline resistant among which (n=11, 44%) isolates carried *tet*O (4 isolates in combination with *erm*A gene and remaining 7 isolates alone), and (n=14, 56%) isolates carried *tet*M by Tn*916* (**Figure 4.4b**), inserted at three different insertion sites (**Table 4.5**). A unique Tn*916* variant was observed in PHEGBS0586 due to insertion of *tet*L adjacent to the *tet*M gene of Tn*916* (**Figure 4.3**). MGE carrying ARGs in all six MDR CC8/CC10 isolates were inserted at the 3' end of *rum*A (**Figure 4.4b)**. The ICE*Sp*2905 (*NZ*_FR691055.1) is a widespread *ermTR*- and *tet*O carrying genetic element of *S. pyogenes*, resulted from one ICE (ICE*Sp*2907 carrying *erm*TR*)* being integrated into another (ICE*Sp*2906 carrying *tet*O), however both ICE*Sp*2906 and ICE*Sp*2907 fragments can move independently between bacteria due to the recombinases these fragments encode, which is possibly responsible to perform both excision and integration [570] .

The second most prevalent tetracycline resistance gene, *tet*O*,* either present alone in isolates (PHEGBS0068, PHEGBS0145, PHEGBS0171, PHEGBS0390 and PHEGBS0393) or in combination with *tet*M gene (PHEGBS0555), was carried by *rum*A-associated MGE - ICE*Sp2906* except in isolate - PHEGBS0068, where it was found as a *rpl*L insert (**Figure 4.4b**). Similarly, the macrolide ARG *erm*A also exists singly as a *rum*A-associated MGE - ICE*Sp*2907 (NZ_FR691055.1) and when combined with *tet*O on a single contig, was carried by ICESp2905-like MGE in four isolates (PHEGBS0091, PHEGBS0308, PHEGBS0509, PHEGBS0581) with an average blast identity >90% and query coverage >85%, respectively (**Figure 4.9**). For a single CC8/CC10 isolate - PHEGBS0555, the MGE carrying *erm*A and its TG remain unidentified due to the presence of ermA on a relatively smaller contig (32 kb).

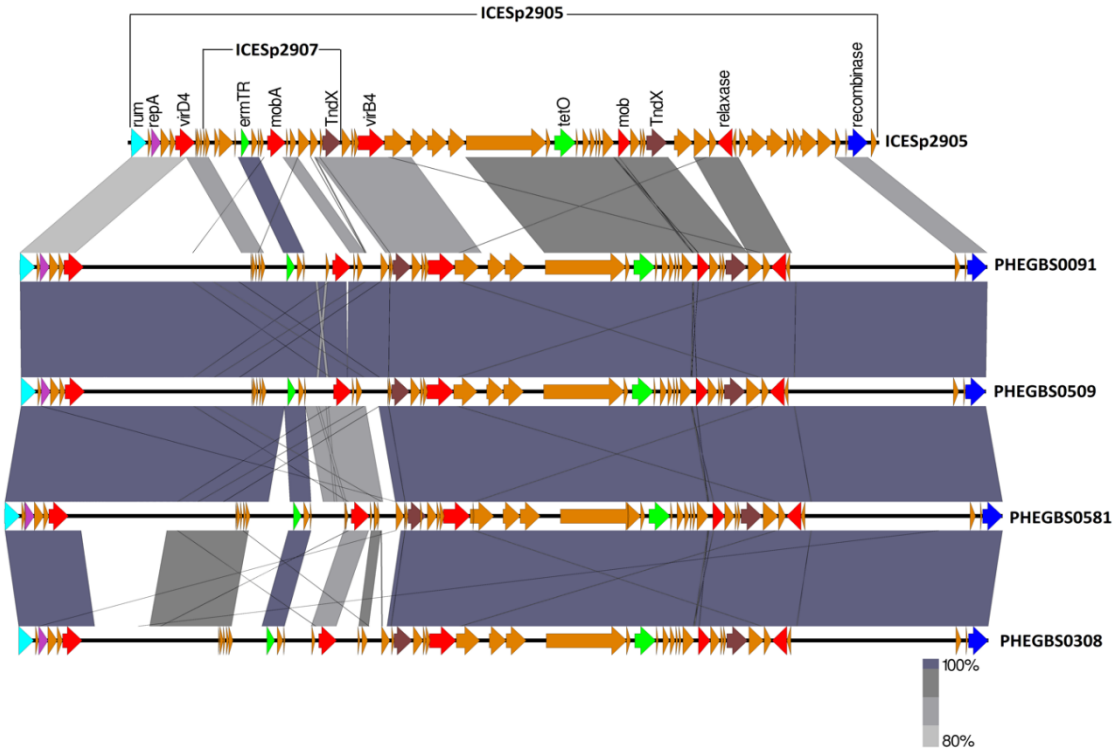



**Figure 4. 9. ICE*Sp*2905-like element (NZ_FR691055.1) identified in four CC8/CC10 serotype II/ST12 isolates.** Different colours are indicating important features where turquoise showing *ru*m (23S rRNA methyltransferase) gene, purple show replication initiator protein – *rep*A, red conjugal proteins, green antibiotic resistance genes, brick red Tn*d*X transpose, dark blue recombinase, and orange other significant genes of ICE*Sp*2905. Areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of ICE*Sp*2905 is 68,252 bp. Easyfig v2.2.2 [336] was used to perform blastn and generate this figure.

### 4.3.4.2.3 CC17

All CC17 isolates (n=32) were tetracycline resistant among which (n=29, 90.6%) isolates carried *tet*M gene, and three (9.3%) carried *tet*O genes in combination with other ARGs (**Figure 4.4c).** Two main clusters based on ICE carrying *tet*M gene in CC17 isolates were observed, first cluster is represented by isolates carrying *tet*M gene by Tn*5801* (n=18) and the second cluster comprising isolates that harboured *tet*M on Tn*916* (n=7). Overall *tet*M genes were mostly carried by Tn*5801* (n=19, 65.5%) followed by Tn*916* (n=10, 34.4%) (**Table 4.5**). All Tn*5801* carrying CC17 isolates were inserted at 472 kb, whereas Tn*916* insertion site corresponds to hotspot region of 592 kb (**Table 4.5**). Only 15.6% (n=5/32) CC17 isolates carried ARGs conferring resistance to multiple antibiotics (**Figure 4.4c).** Previously reported ICE*Sag*37 [477] carrying *erm*B in combination with three aminoglycoside

resistance genes (*ant*(6-Ia), *aph*(3'-III) and *aad*E), were found at the 3' end of *rum*A in isolates (PHEGBS0207, PHEGBS0288, PHEGBS 0559 and PHEGBS0577) and showed a blast identity and query coverage > 95%, respectively (**Figure 4.10**). The ARGs *mef*A/*msr*D in CC17 serotype III/ST17 isolates PHEGBS0288 and PHEGBS0577 were a part of the ICE*Spy*009 MGE [571], also found in CC23 serotype Ia/ST498 isolate PHEGBS0518 (**Figure 4.11)** as a close variant. For a single isolate – PHEGBS0066, MGE carrying *erm*A gene remain unidentified due to presence of *erm*A at the end of a 22kb contig, however *rum*A was found next to *rep*A gene suggesting *erm*A associated MGE was a *rum*A insert (**Figure 4.4c)**.

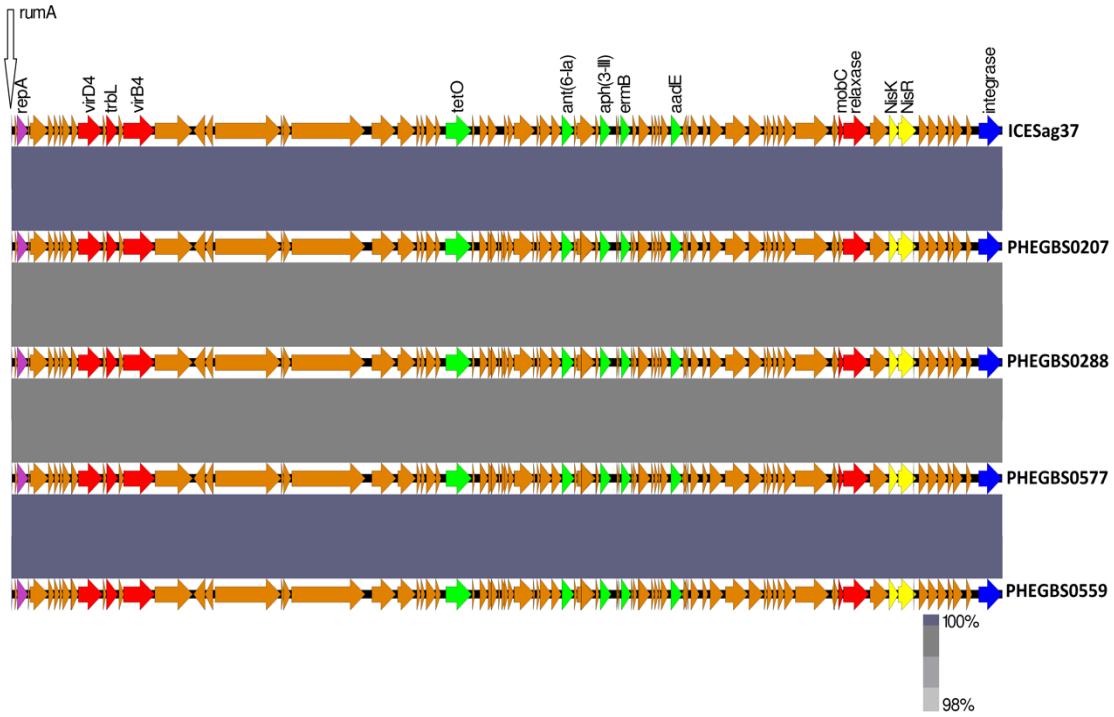



**Figure 4. 10. Comparison of ICE*Sag*37 identified in four CC17 isolates.** ICE*Sag*37 (NZ_CP019978) was identified at the 3' end of *rum*A gene in four serotype III/ST17 isolates (PHEGBS207, PHEGBS0288, PHEGBS0577 and PHEGBS0559). Different colors are indicating important features where, purple shows replication initiator protein - *repA*, red transfer conjugal proteins, green antibiotic resistance genes, yellow potential virulence genes, blue integrase gene and orange other significant genes of ICE*Sag*37. Blastn was performed using Easyfig v2.2.2 [336] and areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of ICE*Sag*37 is 73,429 bp.

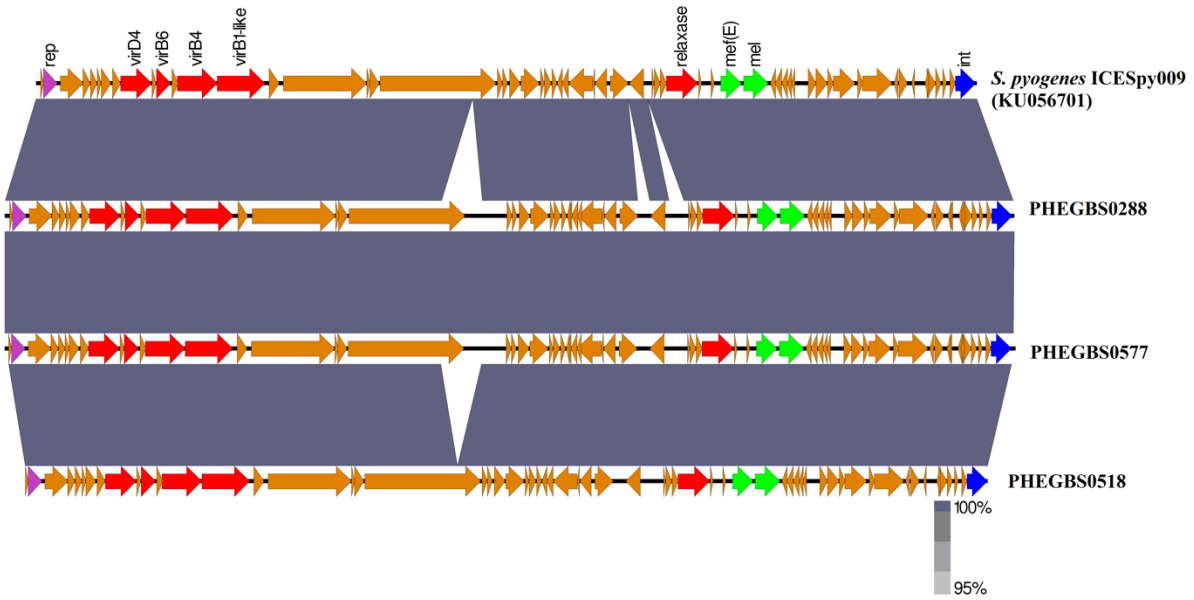



**Figure 4. 11. Comparing *S. pyogenes* ICE*Spy*009 identified in three UK GBS isolates.** Two serotype III/ST17 isolates belonging to CC17 (PHEGBS0288 and PHEGBS0577) and one CC23 serotype Ia/ST498 isolate - PHEGBS0518 found to carry macrolide ARGs *mef*A/*msr*D genes by ICE*Spy*009 (NZ_KU056701)-like element. Different colours are indicating important features where, purple shows replication initiator protein - *rep*A, red transfer conjugal proteins, green macrolide resistance genes, blue integrase gene and orange other significant genes of ICE*Spy*009. Areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of ICE*Spy*009 is 55.5kb. Easyfig v2.2.2 [336] was used to perform blastn and generate this figure.

#### 4.3.4.2.4 CC19

CC19 isolates contained two novel ICEs harbouring multiple ARGs and consist of highest number of MDR GBS isolates (n=11) including all chloramphenicol resistant GBS (n=3) isolates identified in this study (**Figure 4.4d**). The most common target gene for MGE carrying ARGs in multidrug resistant isolates was *rpl*L (n=5), followed by *rum*A (n=4) and *rps*I (n=1), while in remaining isolates (n=2), TG for MGE carrying ARGs remain unknown (**Figure 4.4d).**

All *tet*M positive (n=26,78.6%) CC19 isolates harboured Tn*916* (**Figure 4.4d**), inserted at a single - 592kb genomic position (**Table 4.5**). Six isolates (200682, 200699, PHEGBS0071, PHEGBS372, PHEGBS0624 and PHEGBS0639) were resistant to macrolide and tetracycline both including three isolates (200682, 200699 and PHEGBS0071) that carried macrolide resistance genes (*erm*B and/or *lsa*C) in combination with *tet*O. The *tet*O gene in isolates (200682 and 200699) was carried by variant

*S.s suis* ICE*Ssu*YS64 (NZ_ MN394765.1) and showed a blastn identity of 96% and query coverage of 100%, respectively in both isolates **(Figure 4.12)**. In 200682, ICE*Ssu*YS64 was found as *rpl*L insert, while in 200699, it is inserted at the 3' end of *rum*A. In isolates PHEGBS00624 and PHEGBS0132, *tet*O genes were carried by reduced *S. suis* composite element CMGETZ080501 (NZ_ KX077897.1) and were present on relatively smaller contigs in both isolates (on 74.8kb contig in PHEGBS0132 and 43.3kb contig in PHEGBS0624), respectively while the total genomic size of CMGETZ080501 is 128.7kb. Blastn identity and query coverage between reference CMGETZ080501 and these isolates - PHEGBS0132 and PHEGBS0624 contigs were >95%, respectively. The third isolate – PHEGBS0071 carried a novel ICE*Sag*71 (**Figure 4.13**) and had (~69.9kb) region of *S. suis* composite element CMGEYY060816 (NZ_ KX077898) harbouring both ARGS *tet*O and *erm*B and 34.5kb region of intact *Streptococcus phage* Javan29 (NZ_MK448726.1), respectively integrated into ICE*Sag*71. The ICE*Sag*71 was inserted adjacent to *rum*A in isolates - PHEGBS0071 and was 121.5kb in size.

Five (15.1%) CC19 isolates carried more than two antibiotic class resistance genes including two isolates (PHEGBS0139 and PHEGBS0156) that carried macrolide, tetracycline, and aminoglycoside resistance genes together and three isolates (PHEGBS0595, PHEGBS00608 and PHEGBS0738) that carried macrolide, tetracycline, aminoglycoside and chloramphenicol resistance genes in combination. The macrolide resistance *erm*B gene was also found inserted at *rps*I in isolate PHEGBS0738 carried on a 19.6 kb ICE additionally carrying aminoglycoside ARG *ant*(6-Ia), and *aph*(3'-III) as well as heavy metal resistance genes. This latter ICE appears to be a reduced/partial *Staphylococcus aureus* composite MGE structure ($MES_{PM1}$) (NZ_AB699882.1). A second isolate (PHEGBS0156) also had a related ICE; however, it lacked the IS1216 transposase (**Figure 4.14**) and genomic analysis found the *rps*I region was not disrupted, making it unclear where this last ICE was inserted. PHEGBS0738 was also the only isolate in the cohort to carry the ARG *cat*(pC194) gene, which was found inserted near 16S operon, that starts with a site-specific integrase and ends at an IS1216E transposase. Previously described ICE*Sag*236 [478] (which unfortunately lacks a GenBank submission), reported to be inserted at *rpl*L and carrying multiple ARG *cat*Q, *erm*A and *mefA*/*msr*D was found in isolates PHEGBS0595 and PHEGBS0608 (**Figure 4.15**).

Isolate PHEGBS0639 was also found to carry the macrolide ARG *mef*A/*msr*D adjacent to the *rum*A gene, but also adjacent to the complete genome for *Streptococcus* phage Javan69 (NZ_MK448828.1), which may be responsible for its insertion (**Figure 4.16**).

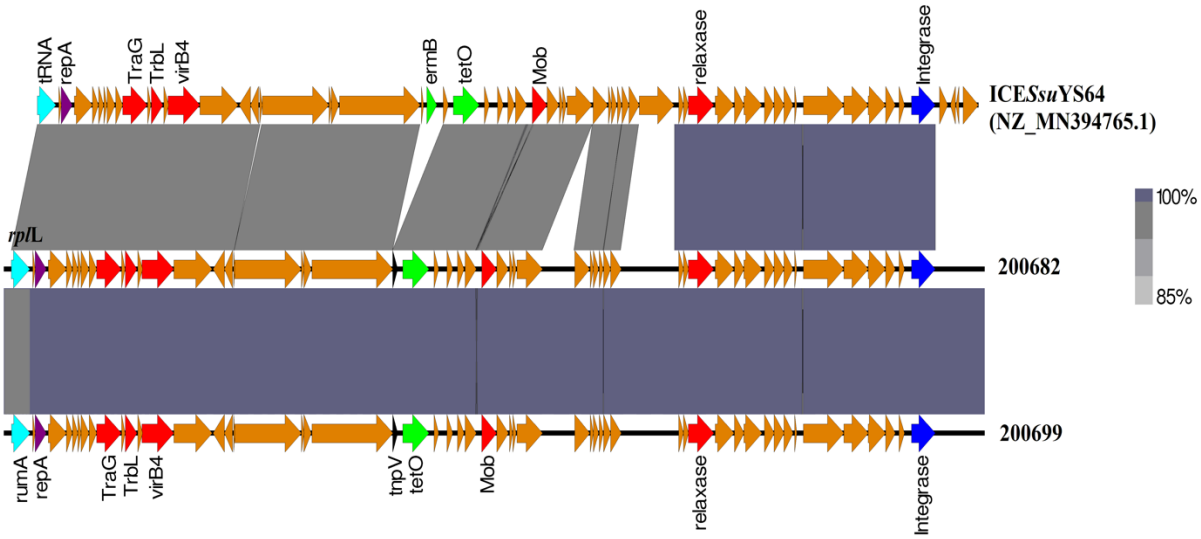



**Figure 4. 12. Comparison of *Streptococcus suis* ICE*Ssu*YS64 with two CC19 isolates.** Two CC19 isolates including serotype III/ST27 isolate – 200682 and serotype III/ST861 isolate – 20069 found to carry *tet*O gene by ICE*Ssu*YS64 (NZ_ MN394765.1) -like element. Different colours are indicating important features where turquoise colour is used to represent target gene - *tRNA* for ICE*Ssu*YS64, *rum*A for ICE*Ssu*YS64-like element insertion in 200699 and *rpl*L for 200682, purple shows replication initiator protein - *rep*A, black shows transposon-encoded protein - *tnp*V gene, green show antibiotic resistance genes, red conjugal proteins, blue - integrase and orange other significant genes of ICE*Ssu*YS64. Areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of ICE*Ssu*YS64 is 69.1kb. Easyfig v2.2.2 [336] was used to generate this figure.

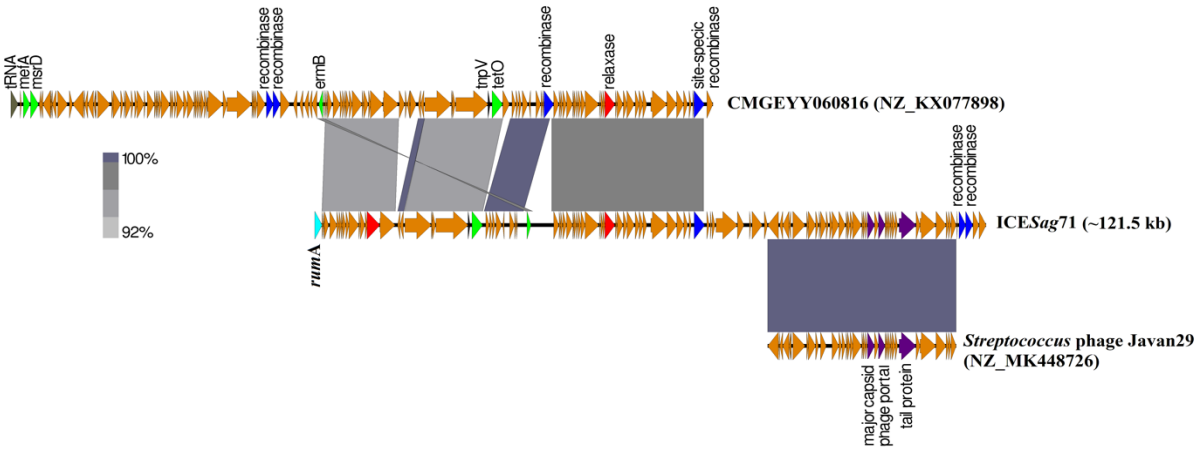



**Figure 4. 13. Novel ICE*Sag*71 identified in CC19 isolate.** A novel ICE*Sag*71 weas identified in CC19 serotype III/ST19 – PHEGBS0071 carrying *erm*B and *tet*O gene, and was a resultant of *Streptococcus suis* composite element CMGEYY060816 (NZ_KX077898; ~69.9kb) and *Streptococcus phage* Javan29 (NZ_MK448726.1; ~ 34.5kb), blastn was performed between the three ICEs using Easyfig v2.2.2 [336]. Different colours are indicating important features of corresponding mobile element where grey and turquoise represents ICE target genes – *tRNA* and *rum*A, respectively, blue- recombinases, red - conjugal proteins, green – antibiotic resistance genes, purple – major capsid, phage portal and tail proteins, respectively. The genomic size of ICE*Sag*71 is ~121.5kb.

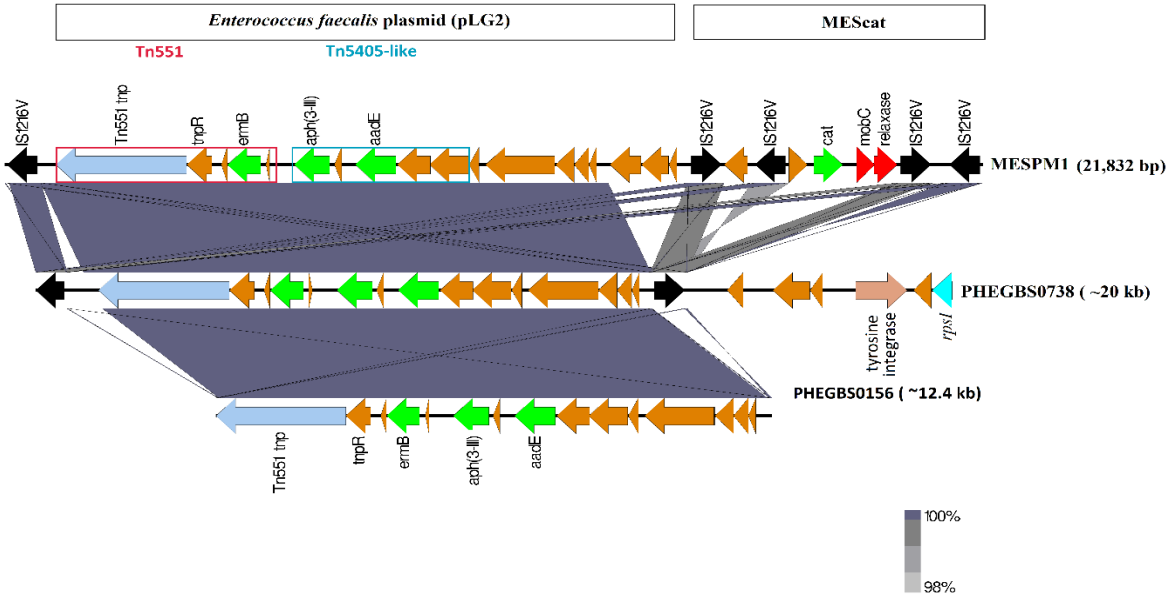



**Figure 4. 14. Comparison of *Staphylococcus aureus* composite mobile element structure MES$_{PM1}$ to CC19 GBS isolates**. Two isolates PHEGBS0738 – V/ST19 and PHEGBS0156 – III/ST19 found to carry ARGs by MES$_{PM1}$ (NZ_AB699882) and reduced MES$_{PM1}$, respectively. Blastn was performed using Easyfig v2.2.2 [336]. Different colours are indicating important features where teal blue showing Tn*551* transposon (tnp) gene, dark blue insertion sequences - IS1216V, green antibiotic resistance genes, red conjugal proteins. salmon-tyrosine integrase and orange other significant genes of MES$_{PM1}$. Areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of MES$_{PM1}$ is 21,8kb.

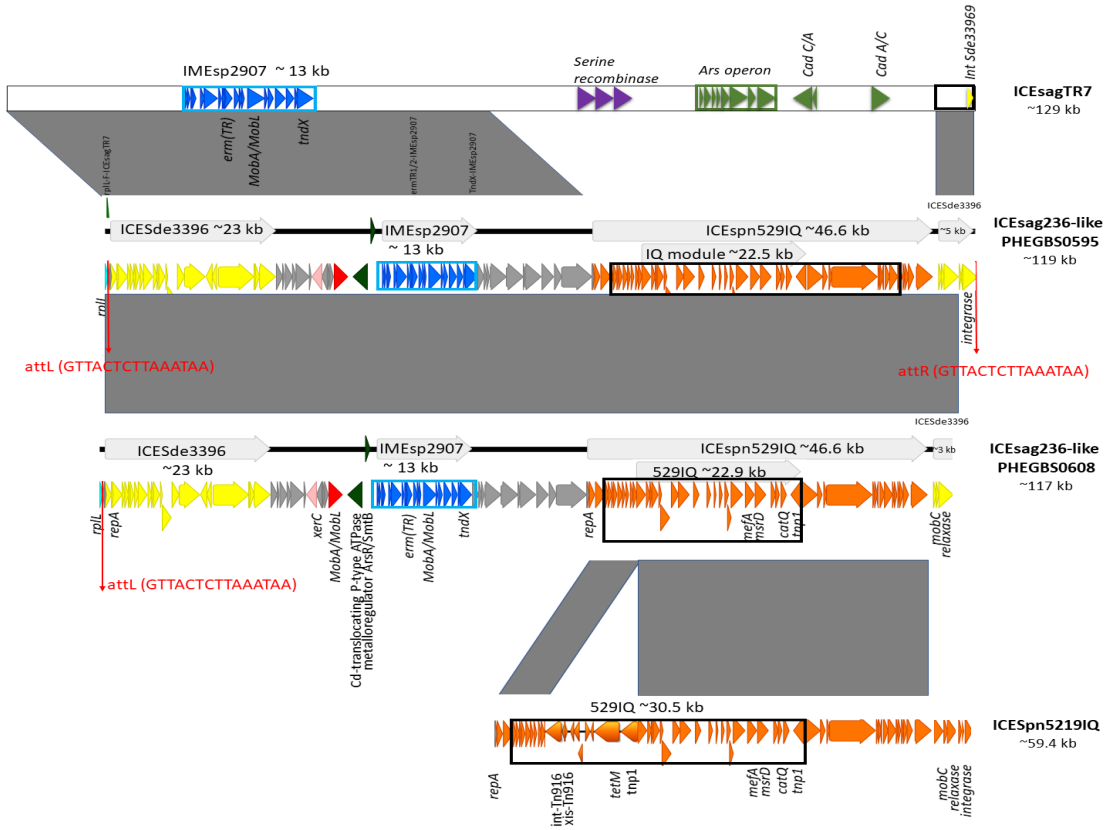



**Figure 4. 15. Schematic, but to scale, representation of the ICE*Sag*236-like element in two CC19 UK GBS isolates.** The ICE*Sag*236-like element in PHEGBS0595 and PHEGBS0608 isolates resulting from the recombination of *S. agalactiae* ICE*Sag*TR7 (top) and *S. pneumoniae* ICE*Spn*529IQ (NZ_HG965092.1) (bottom). The ICE*Spn*529IQ orfs are coloured in orange, ICE*Sag*TR7 portions are white, whereas important features like antibiotic and heavy metal resistance genes shown in dark green, recombinase genes in purple, IME*Sp*2907 orfs are shown in blue box and ICE*Sde*3396 orfs are presented in yellow. Grey areas between ICEs indicate regions with close correspondence using Blastn.

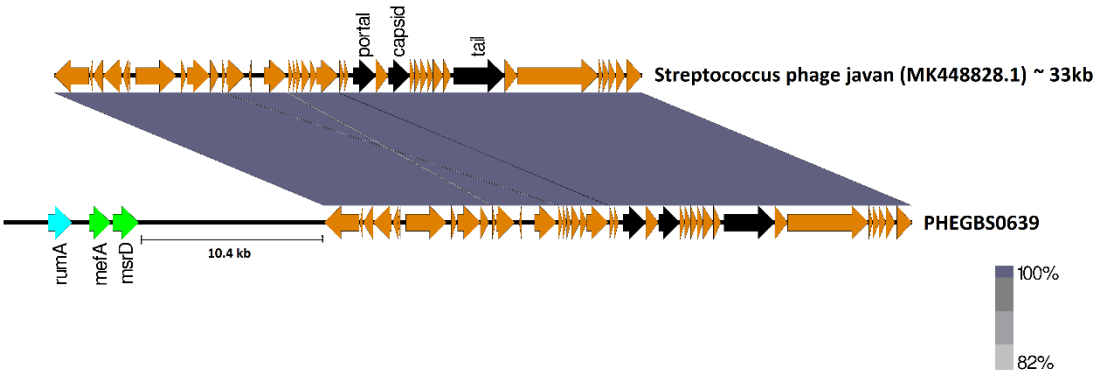



**Figure 4. 16. Comparison of *Streptococcus* phage identified in CC19 isolate.** *Streptococcus* phage Javan (NZ_MK448828.1) was identified in serotype V/ST110 isolate - PHEGBS0639 adjacent to *rum*A and macrolide resistance genes (*mef*A/*msr*D). Different colours are indicating important features where, turquoise shows integration gene - *rum*A, green – antibiotic resistance genes, and black - major portal, major capsid and tail proteins of a phage. Easyfig v2.2.2 [336] and areas with BLAST hits shown by bars between sequences in grey colour. Genomic size of *Streptococcus* phage is 33 kb.

The only other non-specific ARG insertion identified across all the GBS isolates was observed in PHEGBS0139 where the *lnu*C lincomycin resistance gene contained on a 1,724-bp including an IS1 transposase homologue and bounded by 25-bp imperfect inverted repeats flanked by 8-bp direct repeats [572]. The isolate PHEGBS0139 was additionally found to carry the ICE*Sp*2907 *erm*A element (also inserted following the *rpl*L gene); however, in this isolate it was co-located with the aminoglycoside resistance gene *aac*(*6′*)-*aph*(*2″*) which encodes a bifunctional enzyme, AAC(6′)-APH(*2*") that confers resistance to a broad spectrum of aminoglycosides and has to date been detected only in gram-positive bacteria, including *S. agalactiae* (group B), *Enterococcus* species, *Staphylococcus aureus, S. mitis*, and group G *Streptococcus* [573-575]. *Both ermA and aac*(6′)-*aph*(2") genes were carried on a novel ICE designated as ICE*Sag*139 of size 101,616 bp (**Figure 4.17**) and was present on a single contig of sequence length 126,729 bp. The ICE*Sag*139 was inserted 15 nt prior to the end of *rpl*L gene resulting in left (*att*L) and right (*att*R) putative core sites (GTTACTCTTAAATAA) of 15 bp long. The first six orfs (~4.6 kb) of ICE*Sag*139 showed 100% homology to the corresponding orfs of *S. dysgalactiae subsp equisimilis* ICE*Sde*3396 element (NZ_EU142041.1) encoding a site-specific integrase (which mediates the chromosomal integration at the 3′ end of the conserved gene, *rpl*L) and other mobilization proteins (MobC and Relaxase). Next to ICE*Sde*3396 six orfs, present ~12.3kb of ICE*Spn*529IQ (NZ_HG965092.1) (orf7-partial orf18) encoding orf8-zeta toxin (*pez*T), orf9 - *pez*A , orf10 - AtPase with chaperone activity, orf11 - methyl accepting chemotaxis, orf7,12-17 - hypothetical protein and orf18 - partial SNP family protein) with a blast identity of 97% however it does not carry any antibiotic resistance genes. The orf3-orf6 of resultant ICE*Sag*139 in PHEGBS0139 showed higher blast nucleotide identity to the corresponding orfs of ICE*Sde*3396 element (100%) than ICE*Sp*n529IQ elements (96%). The macrolide resistance gene (*erm*A) in ICE*Sag*139 is carried by the ICE*Sp*2907 *erm*(TR) fragment (NZ_FR691055.1) ~12.6 kb insertion with a blast homology of 99% and a query coverage of 100%. The last orf of the ICE*Sp*2907 *erm*(TR) fragment is a Tn*d*X-like transposase (reported as orf24 in ICE*Sp*2905) that is likely to be responsible for site specificity of both excision and integration [570]. Two mosaic genes from *Enterococcus faecium* strain E1077 plasmid pE1077-217

(NZ_MT074686) was observed to insert the aminoglycoside resistance gene - *aac*(6′)-*aph*(2”) and an insertion sequence (IS256-like element) into ICE*Sag*139 (**Figure 4.17**).

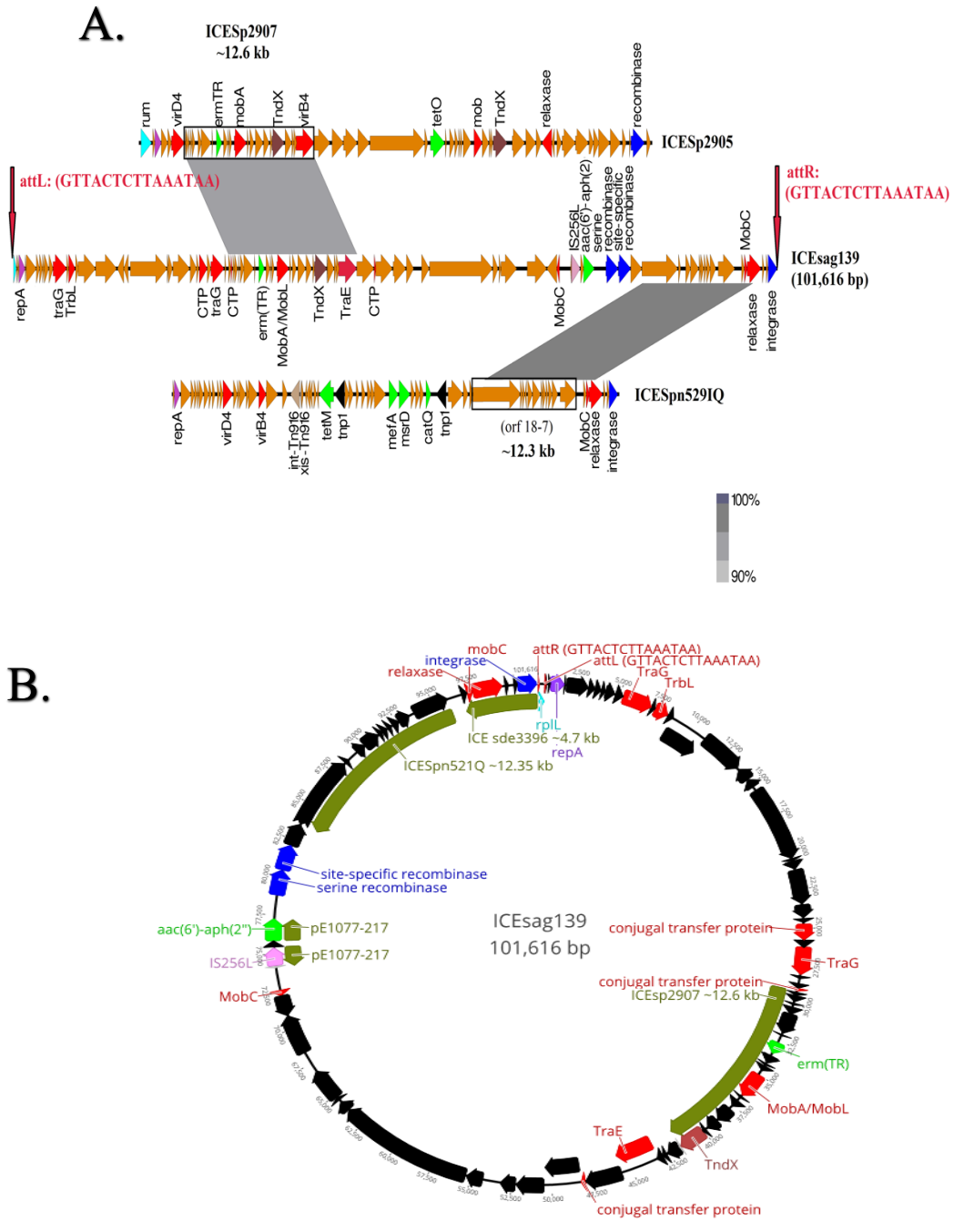



**Figure 4. 17. Representation of newly identified ICE*Sag*139 detected in UK GBS serotype V/ST19 PHEGBS0139: (A)** Blastn between ICE*Sp*2905 (NZ_FR691055.1), ICE*Spn*529IQ (NZ_HG965092.1) and newly identified ICE*Sag*139 in PHEGBS0139. Blastn was performed in Easyfig v2.2.2 [336]. Grey areas between orfs denote nucleotide identities with a gradient representing 90% (light grey) to 100% (dark grey) identity. Turquoise colour is used to represent integration sites *rum* for ICE*Sp*2905 and *rep*A for ICE*Sag*139), purple for replication initiator gene (*rep*A), red for conjugal transfer proteins (CTP), green for antibiotic resistance genes, brick red for Tn*dX* transpose, light pink for insertion sequence (IS256L) in ICE*Sag*139, blue for integrases/recombinases, light brown to show integrase/excisionase of Tn*916* and black to show tnp1 transposases in ICE*Sp*529IQ element respectively. **(B)** Schematic

diagram ICE*Sag*139 in circular form. Open reading frames (orfs) are shown as arrows to indicate the direction of transcription and are coloured in accordance with their predicted gene functions where turquoise used for integration site (*rpl*L), purple for replication initiator gene (*rep*A), red for conjugal transfer proteins (CTP), green for antibiotic resistance genes, brick red for Tn*d*X transpose, light pink for insertion n sequence (IS256L), and olive arrows to represent source/insert of corresponding orfs from *S. dysgalactiae subsp equisimilis* ICE*Sde*3396 element (NZ_EU142041.1), *S. py*ogenes ICE*Spn*529IQ (NZ_HG965092.1), and *E.faecium* plasmid (pE1077-217) (NZ_MT074686).

#### 4.3.4.2.5 CC23

Overall, 13.2% (n=7/53) CC23 isolates were resistant to macrolide and tetracycline both and only one isolate being non-tetracycline resistant (**Figure 4.4e**). Significant number of tetracycline resistant isolates (48/53, 90.5%) carried *tet*M by Tn*5801* (**Figure 4.4e**). Eight different insertion sites were observed for Tn*5801* found in CC23 isolates (**Table 4.5**), the most common of which was 476kb (n=33) genomic region. Five serotype Ia isolates (four ST23 and one ST498) carrying Tn*5801* also had *mef*A/*msr*D macrolide resistance genes; however, in four ST23 serotype Ia isolates (PHEGBS0067, PHEGBS0070, PHEGBS0248 and PHEGBS0625) the TG were unidentified for the macrolide genes, which were inserted disrupting an ABC transporter substrate-binding protein adjacent to a XerS tyrosine recombinase (13kb adjacent to Tn*5801*, rather than within the MGE) and it was unclear if the result was a composite MGE. Whereas in serotype Ia ST498 PHEGBS0518 *mef*A/*msr*D genes were carried by a variant ICE*Spy*009 (NZ_ KU056701.1) (**Figure 4.12**; TG unidentified) and was present on a separate contig.

In only one CC23 isolate (PHEGBS0539; for the *erm*T gene), the resistance gene responsible was carried on a *S. agalactiae* strain GB2001 plasmid pGB2001 (NZ_JF308630.1) with a blast nucleotide identity of 99% (**Figure 4.18**). Presence of *erm*T on a plasmid in invasive human GBS isolates has been reported in earlier studies [183, 205].

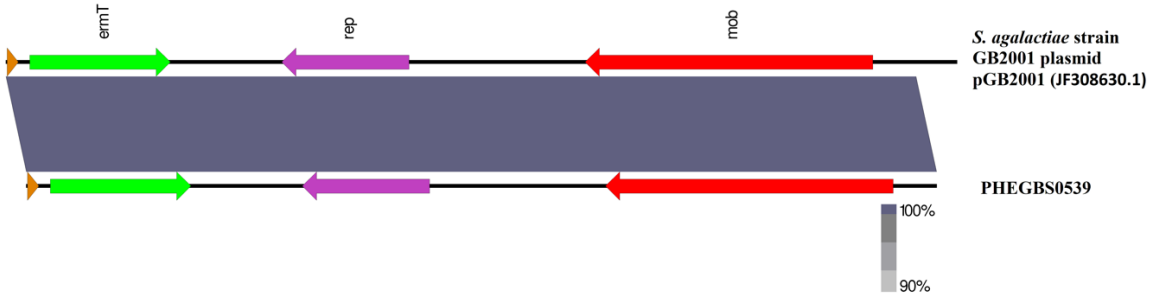



**Figure 4. 18. Identification of *S. agalactiae* strain GB2001 plasmid (pGB2001) carrying *erm*T in CC23 isolate.** Blastn was performed in Easyfig v2.2.2 [336] between reference *S. agalactiae* strain GB2001 plasmid pGB2001 (NZ_JF308630.1) and serotype Ia/ST23 isolate (PHEGBS0539). Green arrow is representing macrolide resistance gene (*erm*T), purple – replication protein, red – conjugal protein and orange - *erm*T leader peptide. Size of the plasmid pGB2001 is 4.9kb.

## 4.4 Discussion

AMR is a progressively critical threat to global public health and MGEs are often responsible to disseminate ARGs between bacteria, thus resulting in the swift increase in MDR strains. In the present study, investigation of the antibiotic resistance profiles and MGEs context of the ARG among 193 UK GBS isolates (collected between January 2014 and May 2015 from England and Wales) were studied for UK adult patients. More than 50% isolates analysed here were from female patients aged between 18-44 years, the bias of this collection was unintentional and do not represent entire collection of UK GBS recovered during that time period. However, as the bias is towards women of childbearing ages, it could have a knock-on effect for AMR in transmission to neonates for the future.

For the treatment of GBS diseases, penicillin is the recommended first line drug [576], our study confirms the uniform susceptibility of all tested GBS isolates to penicillin (100%), also none of the penicillin binding protein (1a, 1b, 2a, 2b and 2x) substitutions in 193 GBS isolates appeared to be associated with decreased β-lactam susceptibility. For PBP2X, similar amino acid substitutions at I377V and V510I were recently reported by Metcalf *et al* [505] but had no impact on decreased β-lactam susceptibility; however, reduced penicillin susceptibility in GBS isolates have been described earlier and emerged through amino acid substitutions particularly at V405A and/or Q557E, in penicillin-

binding protein 2x (PBP2X) [179, 186]. High susceptibility rates to high level gentamicin, and vancomycin [212, 215, 577], whereas moderate chloramphenicol, and levofloxacin resistance rates [214, 238, 505] found in this study are in line with previous studies on GBS disease in adults in different countries .

The prevalence of GBS resistance to clindamycin and erythromycin appear to be high in UK as seen in various countries [78, 215, 312, 505, 578]. A previous study in 2010 in UK showed the prevalence of GBS resistance to clindamycin was 9% and to erythromycin was 15% [115]. The rise of GBS isolates resistant to clindamycin and erythromycin was also observed in the UK between 2014 and 2015 [480, 481]. The overall GBS resistance rate to clindamycin in PHE surveillance reports of 2014 and 2015 was found to be 18% and 23%, based on 631 (37.6%) isolates tested for clindamycin resistance out of 1676 GBS bacteraemia cases reported in UK in 2014 and 997 (53.3%) isolates tested for clindamycin resistance out of 1870 GBS bacteraemia cases reported in UK in 2015 [480, 481]. The overall resistance rates to erythromycin in GBS in PHE reports of 2014 and 2015, was found to be 22% and 28% based on 1062/1676 (63.3%) GBS isolates tested for erythromycin resistance in 2014 and 1231/1870 (65.8%) isolates tested for erythromycin resistance in 2015, respectively [480, 481]. In this analysis, only 193 isolates from adults aged more than 18 years were analysed that represents the isolate submission proportion of less than 11% of total invasive cases reported in UK in 2014 and 2015, with most isolates from neonates (estimated >95%). The erythromycin and clindamycin resistance rates observed in 193 adults GBS were 20.2% and 11.8%, respectively which is not comparable to the overall erythromycin and clindamycin resistance rates observed in GBS isolates recovered from patients of all age groups and reported to PHE in 2014 and 2015 [480, 481]. Providing that resistance to erythromycin and clindamycin is increasing across the globe and presents a challenge for treatment and prophylaxis strategies, vancomycin has been recommended in UK to treat penicillin allergic people as per RCOG 2017 guidelines [3], and since vancomycin has still low resistance rates, its potential to be used for adult iGBS treatment seems promising.

In this study, *erm*A, and *erm*B were the most frequent gene determinants associated with macrolide resistance in our GBS samples which agrees with earlier findings [312, 505, 579]. The prevalence of the *erm* determinant shows that GBS commonly use target methylation as the mechanism of macrolide resistance. This study also confirms that he *mef*A/*msr*D gene are less frequent among GBS isolates, as seen earlier [580], thus efflux pumps mediated by this gene are not a common mechanism of macrolide resistance. Contrary to Bolukaoto *et al* study [215], this current study did not find any *linB* genes 0% (0/23) in any of the clindamycin resistant and intermediate strains. In such cases, where multiple independent resistance genes can cause resistance, the observed phenotypic resistance can be attributed to any of the other known genes or genes that are yet to be discovered. This however limits the usefulness of such diagnostic tests. It was interesting to note that 61.9% (13/21) GBS which were resistant to both erythromycin and clindamycin had the *erm*B gene (**Table 4.1**) and that one GBS strain which was resistant to erythromycin and clindamycin both carried the *erm*T gene, harboured on a plasmid. Such observations are similar to reports from other studies [183, 190, 205, 581, 582].

Tetracycline is relatively cheap antibiotic and its extensive use after its first discovery in 1948 led to the high resistance profile in many bacteria [583]. Since the 1950s, most GBS clones were wiped out by the collateral effect of the widespread use of tetracycline and other antibiotics, even though they were not the treatment target [1]. Da Cunha *et al* showed that in the case of GBS extensive tetracycline usage has exerted strong selective pressure that led to the loss of genetic diversity in GBS and emergence of few tetracycline resistant pathogenic clones that also carried mobile genetic elements carrying a mixture of antibiotic resistance genes, including those of tetracycline resistance [1]. In this study as well, high rate of tetracycline resistance (175/193, 90.6%) was noted, which is in keeping with general consensus [1] and can be attributed to a high presence of the *tetM* gene (154/175, 88%) [580] [505]. This higher rate of *tet*M gene is as a result of the ubiquitous presence of *tet* genes in pathogens, opportunistic pathogens and members of the normal microbiota [228]. The *tet*M gene was mostly carried by two related ICE, Tn*916*/Tn*916* like element (56.4%) and Tn*5801* (43.5%), these elements were inserted at specific hotspot regions but at different insertion sites (**Table 4.4**) between each clonal complexes, similar observations were seen in Da cunha *et al* study [1], suggesting that these ICEs

harbouring *tet*M are clonally related, acquired through limited and rare insertion events and led to expansion of these lineages. A similar analysis of the insertion sites of Tn*916* has also been performed in *Butyrivibrio proteoclasticus* B316 [584] and *Clostridium difficile* [585] demonstrating a preference for Tn*916* insertion in intergenic sites for these genera including GBS. For CC23 Tn*5801* inserted at multiple insertion sites, this could be local variation in DNA topology between strains that influence target site selection by Tn*5801,* further analysis and experiments such as SNP based analysis and mating experiments are required between non-Tn*916* isolates with *gyr*A/*par*C mutation (non-mobile) to see if tetracycline and ciprofloxacin resistant strains is achieved to provide evidence about the Tn*916* movement and to identify if it is inserted at the same insertion site.

Specific associations were identified between ICE carrying *tet*M gene and specific CC these ICE positive isolates belong to, Tn*3872* was limited to CC1 only suggesting a clonal expansion, Tn*916*/Tn*916*-like element was found more prevalent in CC1, /CC10 and CC23 whereas Tn*5801* exclusively found in CC19 and CC23 isolates which supports the earlier findings [1]. Two hotspot TGs were identified for the chromosomal integration of Tn*916*, Tn*3872* and Tn*5801*, respectively that is 3' end of *rum*A for Tn*916 and* Tn*3872* insertion as seen in the reference SS1 genome (NZ_CP010867) and CZ-NI-013 genome [1], respectively and 5' end of *gua*A for Tn*5801* which agrees with previous study findings [1] and demonstrating Tn*916*/Tn*1545* family elements have integration preferences.

Five novel ICEs carrying multiple ARGs were identified in this study including (ICE*Sag*84, ICE*Sag*662 and ICE*Sag*100414) in three CC1 isolates and (ICE*Sag*71 and ICE*Sag*139) identified in CC19 (**Figure 4.4a and d**), respectively. 41 MGEs carried single or multiple antibiotic resistance determinants and three different TGs were determined mediating integration of these 41 MGEs into chromosome, the most frequent of which was 3' end of *rum*A (70.7%), followed by MGEs inserted at 15 nt prior to the 3' end at *rpl*L (21.9%) and less commonly adjacent to *rps*I (7.3%) (**Table 4.4**). A recent study [444] described 15 TGs, as a common hotspot for ICEs integration in *Streptococcus* including sites observed in this study - *rum*A, *rpl*L and *rps*I.

Overall, this study suggests that WGS is an efficient way to characterize horizontal transfer of AMR genes. WGS cannot only identify whether AMR genes are located on the chromosome or on MGEs, which is a critical piece of information, but also help to better characterize those MGEs. WGS can also explore whether an MGE has the molecular machinery for conjugation to occur and therefore predict the mobility profile of an MGE. A diverse group of MGEs carrying ARGs were found accountable for the dissemination of phenotypic resistance in GBS isolated from adult patients in England and Wales, exceptionally those belong to CC1 and CC19 isolates where maximum complex and five novel MGEs were identified. Tn*916*/Tn*1545* family elements carrying *tet*M genes were widely distributed and were highly site specific, however mobility of these elements couldn't identify as no mating experiments were performed. Infections caused by these multi-resistant organisms significantly increase morbidity, mortality, and health care costs hence continued investigation of the basic biology of mobile elements will be needed for meaningful understanding of the properties of known and yet-to-be-discovered elements.

# Chapter 5:

# *Streptococcus agalactiae* in Brazil and Australia: serotype distribution and antimicrobial susceptibility patterns

## 5.1. Summary

The present study aimed to determine the patterns for antimicrobial susceptibility and capsular serotypes in a small dataset of iGBS isolates from Brazilian adults and a dataset of carriage GBS isolates from Australian pregnant women, collected at ≤22 weeks' gestation. A total of 26 iGBS isolates from Brazilian adults and 171 GBS isolates from vaginal and/or rectal e-swabs from Australian pregnant women were available for analysis. Serotype Ia (46.1%) and V (38.4%) in Brazilian adult population and serotype Ia (22.8%), V (22.2%), and III (20.5%) in Australian pregnant women were found to be the most common. All Brazilian and Australian GBS isolates were susceptible to ampicillin, vancomycin, and high-level gentamicin, however within the Australian antenatal cohort 2.9% (n=5) and 1.7% (n=3) of GBS isolates were found to be levofloxacin and chloramphenicol resistant, respectively. Low resistance rates to erythromycin (7.6%, n=2) and clindamycin (3.8%, n=1) were observed in Brazilian GBS strains mediated by *erm*B (n=2) and *mef*A/*msr*D (n=1) macrolide resistance determinants. Likewise, low resistance rates to erythromycin and clindamycin were also observed in GBS isolates from Australian pregnant women (13.4%, n=23) and (8.7%, n=15), respectively. As expected, high resistance rate to tetracycline were detected in Brazilian (80.7%) and Australian isolates (81.2%). This study confirms that hexavalent vaccine targeting serotype Ia, Ib, II, III, IV and V is the most promising strategy in reducing GBS disease. Additionally, the described resistance rate of GBS to clindamycin, a second alternative therapeutic still used in penicillin allergic patients in some geographical regions, supports the need for routine susceptibility testing against recommended antibiotics to ensure effective therapy.

## 5.2. Introduction:

GBS is a clinically crucial pathogen that colonizes 10-40% of pregnant women's gastrointestinal and genital tracts [586]. The colonizing GBS usually remains asymptotic in women but can cause severe infection in neonates and infants manifesting as sepsis and meningitis following vertical transmission from mother to newborns (either during delivery or possibly following ascending infection following membranes rupture) [587]. Other potential adverse effects arising in pregnant women colonized with GBS include preterm labour, urinary tract infection, premature rupture of membranes, postpartum sepsis, chorioamnionitis, pneumonia or meningitis [285, 587]. Some reports also highlight sequelae including osteomyelitis and mastitis to be associated with the presence of GBS in mothers following delivery [588]. While GBS remains the primary cause of neonatal sepsis since the 1970s [589-591], with time, it has also steadily gained recognition as a pathogen causing severe infections in elderly and adults, especially those with certain chronic medical conditions, such as diabetes or liver disease [251,522].

In neonates, based on the age of presentation, GBS infections divides into diseases, early onset and late onset disease [592]. Early onset disease (EOD) is an invasive GBS (iGBS) infection occurring in newborns within first seven days of life, but most neonates (61% to 95%) become ill within the first 24 hours (median, 1 hour) [592]. Infants typically present with respiratory distress such as apnea or tachypnea, grunting respirations, and cyanosis [154]. Other signs include lethargy, poor feeding, abdominal distention, pallor, jaundice, tachycardia, and hypotension [154]. Bacteraemia is the most common form of early-onset GBS disease, accounting for approximately 80% of cases [592]. Pneumonia and meningitis, although not uncommon, are less likely presentations in early-onset disease, accounting for 15% and 5% to 10%, respectively [593]. Late onset disease (LOD) is the second type of iGBS infection in neonates, occurring after the first week of life but prior to 90 days of age, and has a similar clinical presentation to early-onset disease [154, 593]. Although bloodstream infections remain the most common presentation of the late-onset disease, meningitis occurs in about 30% of cases, as opposed to 5% in early-onset disease [593].

Prevalence of maternal GBS colonization vary geographically and estimated highest in Africa (11.9%-31.6**%)** followed by the America (6.5%-36%), Asia (9.1%-25.3%), Australian and New Zealand (15%–21%) and Europe (7.1%-16%) [122, 339, 594], which is supported by a most recent systematic review estimating maternal GBS colonization prevalence in different countries [147]. In addition to prevalence, classification into serotypes based on the capsular polysaccharide has become an important characterization of GBS, mainly expecting the geographical variability of serotypes amongst GBS and because of its potential use as vaccine candidates. The most widespread serotypes of maternal colonization are Ia, Ib, II, III and V, which account for 98 % of serotypes globally [339]. A recent (2019) study of Western Australian (WA) pregnant women [384], from which a subset of isolates used in this analysis has identified predominance of serotypes Ia, III, V and II, representing more than 15 % of the positive isolates combined, and additionally serotype Ib, VI, IV, VIII and IX with no serotype VII being identified [384]. No significant variations were observed in the serotype distribution of GBS colonizing WA pregnant women [384] compared to the previous limited Australian studies except that in early 2008 study, serotype VII was identified instead of serotype VIII [383]. In a study conducted in 2006 [595], higher rate of serotype III (43.6%) was observed compared to (20.8%) serotype III identified in WA pregnant women cohort and lower rate of serotype II (in ≤10 %) of positive isolates were identified in other Australian studies [382, 595-597] compared to higher rate of serotype II (16.3–18.2 %) identified in WA pregnant women cohort. Serotypes IV, VI, VII, VIII and IX, are considered to be rarer among pregnant women in Australia, but this is not true for all countries [598, 599]. For example, a study in Ghana found that serotype VII was the most common GBS serotype colonizing pregnant women, followed by IX [598]. Maternal GBS colonization rate in WA pregnant women cohort was 24% [384], which supports the average maternal colonization rate specified for Australia (15-21%) in an earlier systematic review [594].

Most of the available data from GBS screening in Brazil in the last 10 years is from studies conducted on pregnant women and in the Southeast region, mainly in the states of São Paulo and Rio de Janeiro [48]. A lot of fluctuation in the maternal GBS colonisation rate is observed in different region of Brazil ranging 4.2% to 28.4%, based on reports published between 2008 and 2018 [372-374]. This variation in the maternal colonization rate may be related to the non-existence of an official recommendation from the Federal Brazilian Health Authorities describing guidelines and protocols to screen GBS in pregnant women, in both public and private clinics. Differences are also seen in the distribution of GBS serotype among pregnant women belonging to different regions of Brazil [371, 374-376]. Few studies conducted between 2008 and 2017 in Brazil, evaluated distribution of GBS serotypes, however most of these studies were performed in the South and Southeast Region of Brazil, and reported the most prevalent serotypes to be Ia, Ib, II, III, IV and V [374, 376-380]. In 2002, serotype III was the most abundant serotype found in Rio de Janeiro pregnant women, followed by serotypes V and Ia [375]. In 2018, serotype Ia was the most prevalent serotype observed among pregnant women living in Rio de Janeiro, followed by serotypes II and Ib [371]. In in 2015 and 2016 , in São Paulo, serotype Ia was also observed as the most abundant GBS serotype colonizing pregnant women, but followed by V and II [51]. In the same year study (2015-2016) from Paraná, serotype Ia was observed as the most prevalent serotype among pregnant women; however, there was a surprisingly high number of serotype IV isolates (13.1%) [376]. The most recent global study analysed GBS serotype prevalence, and reported serotypes Ia, Ib, II, III, and V as the most prevalent serotypes in pregnant women worldwide, based on reports published between 1997 and 2020 [147].

Poor quality information on change in the GBS antibiotic resistance profiles is available in Brazil due to fragmented data available from small local studies [48]. However, based on those narrow studies, an increase in erythromycin resistance with no significant trend in clindamycin resistance is observed in Brazil over time. Two studies performed in Rio de Janeiro in 2011 [384, 600], reported GBS isolates resistant to erythromycin was 13.2% and 14%, respectively with later study also reported resistance rate to clindamycin (5%). Later in 2016 [601], a group from Paraná, Brazil, evaluated 136 GBS isolated

from pregnant women, and found resistance rate to erythromycin was 8.1% and to clindamycin, it was 2.2%. Botelho et al. (2018) [371], in Rio de Janeiro, showed that among 592 GBS isolates, 14% isolates were resistant to erythromycin resistance and 2% to clindamycin. To date, based on the studies conducted in Brazil, all isolates presented sensitivity to ceftriaxone, penicillin, and vancomycin [378, 600]. In relation to this, recent studies have described isolates with reduced susceptibility to penicillin in other countries, which raises concerns about routine antibiotic susceptibility testing of GBS for the main antibiotic used in IAP [177, 505].

Overall, limited short term data is available about the occurrence of GBS colonization among pregnant women in Australia, while in Brazil, fragmented data conducted in local regions has so far described circulating GBS serotypes and antibiotic resistance patterns. In this chapter, a subset of Australian pregnant women isolates from earlier study [384] and a small available sample collection from a single site in Brazil were used to study both the serotype distribution and the antibiotic susceptibility patterns.

## 5.3. Results

### 5.3.1. Brazilian GBS isolates:

Of 46 Brazilian samples sent to Cardiff University, only 34 samples (including 27 samples from adults, six from neonates and 1 with no information) were confirmed as GBS isolates after WGS analysis. Neonatal isolates (n=6) due to small number in a given dataset, and one isolates with no information of patient age were excluded from this analysis. Of 27 adult GBS samples, 1 additional sample (BS-2293) did not pass the assembly quality check (**Chapter 2: Table 2.4**), further narrowing down the final Brazil GBS dataset to 26 isolates (**Table 5.1**) obtained from people with age ranging between 23-81 years (median age of 59 years) with 61.5% GBS isolates recovered from male patients (16/26). As expected, due to the high potential of GBS disease in old age people with comorbidities, a higher median age-group was observed for males (age range: 24 to 74 years, with a median age of 59 years old) (**Figure**

**5.1a**), compared to females (age range: 23 to 81 years, with a median age of 45.5 years old) (**Figure 5.1b**). More GBS were isolated from old men aged between 65-84 years, (n=17/16, 43.7%) relative to men in other two age groups; 18-44 yrs (n=4, 25%) and 45-64 yrs (n=5, 31.2%) (**Figure 5.1a)**, whereas half of the female samples were from females in childbearing age group 18-44 yrs (5/10, 50%) (**Figure 5.1b**).

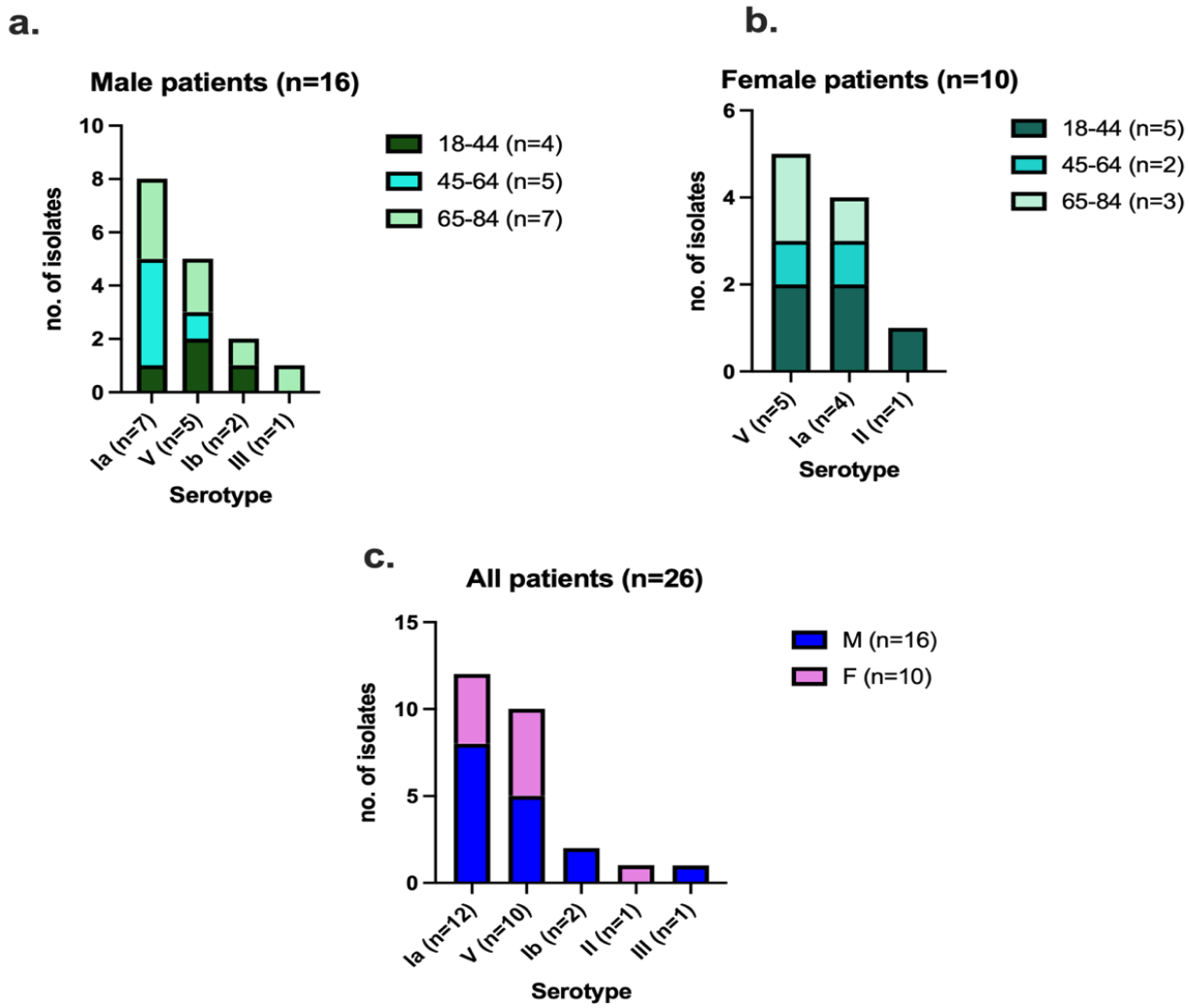



**Figure 5. 1. Summary of demographic information available for 26 iGBS isolates from Southeast region of Brazil. (a)** graphical distribution of GBS serotypes identified in male patients, segregated by age group; **(b)** graphical distribution of GBS serotypes identified in female patients, segregated by age group; **(c)** Overall number of GBS serotypes observed in the dataset segregated by sex.

**Table 5. 1. Description, genotypic and phenotypic antimicrobial resistance profiles of iGBS isolates (n=26) recovered from clinical blood specimens of Brazilian adults between July 2010 to November 2017.**

| Isolates | Serotype | ST | TET | ERY | DA | CHL | AMP | CN | LEV | VA | Age in years | Sex | Year of sample isolation | Specimen | ARG | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | ≤18 / ≥23 | ≤15 / ≥21 | ≤15 / ≥19 | ≤17 / ≥21 | / ≥24 | 6 / ≥10 | ≤13 / ≥17 | / ≥17 | | | | | Tetracycline | | Macrolide |
| BS-S1911 | Ia | 103 | 12 | 24 | 20 | 25 | 30 | 20 | 24 | 20 | 61 | M | 2010 | BC | *tet*M | | |
| BS-S1976 | Ia | 103 | 25 | 25 | 20 | 28 | 32 | 22 | 27 | 21 | 78 | M | 2010 | BC | NRG | | |
| BS-S2039 | Ia | 103 | 13 | 0 | 0 | 26 | 30 | 19 | 24 | 19 | 57 | M | 2011 | BC | *tet*M | | *erm*B |
| BS-S2099 | Ia | 23 | 16 | 25 | 24 | 28 | 28 | 19 | 25 | 20 | 37 | F | 2011 | BC | *tet*M | | |
| BS-S2208 | V | 1 | 11 | 29 | 26 | 26 | 32 | 21 | 25 | 20 | 67 | F | 2012 | BC | *tet*M | | |
| BS-S2309 | V | 1 | 13 | 28 | 24 | 25 | 30 | 20 | 25 | 19 | 63 | F | 2013 | BC | *tet*M | | |
| BS-S2315 | Ia | 23 | 14 | 29 | 26 | 26 | 30 | 24 | 25 | 21 | 55 | M | 2013 | BC | *tet*M | | |
| BS-S2318 | Ia | 23 | 14 | 29 | 26 | 26 | 36 | 21 | 24 | 21 | 72 | M | 2013 | BC | *tet*M | | |
| BS-S2375 | Ib | 12 | 25 | 27 | 26 | 27 | 33 | 20 | 26 | 20 | 66 | M | 2013 | BC | NRG | | |
| BS-S2497 | Ia | 23 | 14 | 22 | 24 | 26 | 28 | 20 | 24 | 20 | 81 | F | 2014 | BC | *tet*M | | |
| BS-S2501 | V | 1 | 14 | 30 | 26 | 28 | 30 | 22 | 26 | 19 | 25 | F | 2014 | BC | *tet*M | | |
| BS-S2509 | V | 1 | 0 | 21 | 20 | 23 | 26 | 15 | 21 | 18 | 69 | F | 2014 | BC | *tet*M | | |
| BS-S2527 | V | 24 | 12 | 25 | 22 | 21 | 26 | 15 | 21 | 19 | 44 | M | 2014 | MD | *tet*M | | |
| BS-S2561 | V | 1 | 0 | 0 | 0 | 26 | 30 | 21 | 26 | 21 | 59 | M | 2014 | BC | *tet*M | | *erm*B |
| BS-S2617 | V | 1 | 10 | 25 | 23 | 26 | 28 | 18 | 22 | 18 | 44 | M | 2014 | BC | *tet*M | | |
| BS-S2633 | Ia | 23 | 0 | 30 | 30 | 30 | 34 | 22 | 22 | 23 | 59 | M | 2014 | BC | *tet*M | | |
| BS-S2709 | III | 529 | 15 | 32 | 27 | 26 | 33 | 17 | 24 | 20 | 72 | M | 2015 | BC | | *tetO* | |
| BS-S2842 | V | 1 | 20 | 34 | 34 | 35 | 37 | 26 | 28 | 23 | 69 | M | 2015 | BC | *tet*M | | |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BS-S2870 | Ia | 103 | 30 | 26 | 23 | 25 | 29 | 19 | 23 | 19 | 68 | M | 2015 | BC | NRG | | |
| BS-S2899 | V | **1349** | 16 | 22 | 26 | 27 | 32 | 19 | 24 | 20 | 33 | F | 2015 | BC | *tet*M | | |
| BS-S3032 | V | **1349** | 14 | 30 | 25 | 26 | 30 | 24 | 25 | 21 | 74 | M | 2015 | BC | *tet*M | | |
| BS-S3149 | Ia | 103 | 25 | 25 | 20 | 25 | 30 | 22 | 26 | 21 | 54 | F | 2016 | BC | NRG | | |
| BS-S3201 | Ia | 144 | 12 | 17 | 26 | 28 | 35 | 22 | 24 | 20 | 28 | F | 2016 | BC | *tet*M | | *mef*A/*msr*D |
| BS-S3255 | Ib | 8 | 28 | 24 | 23 | 24 | 32 | 15 | 24 | 18 | 38 | M | 2016 | BC | NRG | | |
| BS-S3300 | II | 817 | 13 | 24 | 28 | 29 | 36 | 21 | 26 | 20 | 23 | F | 2017 | BC | *tet*M | | |
| BS-S3316 | Ia | 23 | 14 | 28 | 25 | 25 | 28 | 19 | 22 | 18 | 24 | M | 2017 | BC | *tet*M | | |

Pink blocks representing GBS isolate resistant to specific antibiotic, green showing intermediate resistance and yellow showing sensitivity of GBS strain to specific antimicrobial.
Antibiotics short form are used where TET is used for tetracycline; ER - erythromycin; DA - clindamycin; CHL - chloramphenicol; AMP - ampicillin; CN - gentamicin; LEV - levofloxacin; and VA – vancomycin.
BC stands for blood culture

**Table 5. 2. Summary of phenotypic antibiotic resistance patterns of GBS of three geographic locations.**

| **Antibiotic** | **Antibiotic resistance patterns of GBS of three geographic locations** | | |
|---|---|---|---|
| | **Brazilian (n=26)[a]** | **Australian (n=171)[b]** | **UK (n=193)[c]** |
| Tetracycline | 80.7% (n=21) | 81.2% (n=139) | 90.6% (n=175) |
| Levofloxacin | - | 2.9% (n=5) | 2% (n=4) |
| Ampicillin | - | - | - |
| Gentamicin | - | - | 0.5% (n=1) |
| Vancomycin | - | - | - |

| | | | |
|---|---|---|---|
| Chloramphenicol | - | 1.7% (n=3) | 1.5% (n=3) |
| Erythromycin | 7.6% (n=2) | 13.4% (n-=23) | 20.2% (n=39) |
| Clindamycin | 3.8% (n=1) | 8.7% (n=15) | 11.8% (n=23) |

[a] iGBS obtained from adults
[b] GBS colonizing Australian pregnant women at ≤22 weeks' gestation (visit one)
[c] invasive and non-invasive GBS causing disease in adult population of UK

#### 5.3.1.1 Serotype and ST distribution in Brazilian GBS isolates

In total, five serotypes were identified with following prevalences: the most predominant were serotype Ia (12/26, 46.1%) and V (10/26, 38.4%), followed by serotype Ib (2/26, 7.6%), II and III (1/26, 3.8%) (**Figure 5.1c**). Serotype Ia (8/16, 50%) preceding serotype V (5/16, 31.2%) were dominant among male (**Figure 5.1a**) whilst serotype V (5/10, 50%) preceding serotype Ia (4/10, 40%) were more commonly identified in female patients (**Figure 5.1b**). Ten STs were recognised in total, including a novel ST (ST1349, n=2) identified in this study (**Table 5.1**). The most common STs observed was ST1 (7/26, 26.9%), followed by ST23 (6/26, 23%) and ST103 (5/26, 19.2%) (**Table 5.1**).

#### 5.3.1.2 Genotypic Antimicrobial resistance (AMR) profiling

In total, four different antibiotic resistance determinants were identified in the 26 GBS isolate collection among which 3/26 (11.5%) of GBS isolates were found macrolide and tetracycline resistant isolated from two male patients aged >55 yrs and a female of childbearing age (28 yrs) (**Table 5.1**). The *erm*B gene was found in two isolates and *mef*A/*msr*D in one isolate, in combination with *tet*M genes **(Table 5.1**). A total of 21/26 (80.7%) Brazilian GBS isolates conferred tetracycline resistant due to the presence of most commonly observed *tet*M (25/26, 96.1%) antibiotic resistance determinant, followed by *tet*O observed in 1/26 (3.8%) isolate (**Table 5.1**). No chloramphenicol, ampicillin, gentamicin, levofloxacin, and vancomycin resistance genes were observed in any of the isolate examined and 5 out of 26 (19.2%) isolates carried no resistance genes at all (**Table 5.1**).

#### 5.3.1.3 Phenotypic AMR profiling

All 26 GBS isolates (100%) examined were found susceptible to ampicillin, gentamicin, levofloxacin, vancomycin, and chloramphenicol (**Table 5.1 and Table 5.2**) which is consistent with a lack of detecting resistance genes for these five antibiotics. Further the 19.2% (n=5) of GBS isolates that carried no antibiotic resistance genes at all, were confirmed to be phenotypically sensitive to all antimicrobials

tested (**Table 5.1**). A total of 21/26 (80.7%) of GBS isolates were found tetracycline resistant phenotypically including one isolate (BS-S2842) that showed intermediate resistance (**Table 5.1**). Consistent with genotypic results, 3/26 (11.5%) of isolates were found macrolide resistant upon antibiotic susceptibility testing (AST) including two isolate (BS-S2039 and BS-S2561; both *erm*B positive) that showed resistance to both erythromycin and clindamycin, and a single isolate (BS-S3201; *mef*A/*msr*D positive) that showed intermediate resistance to erythromycin only (**Table 5.1**).

### 5.3.2 Australian GBS isolates:

A larger cohort study, Predict1000, was approved by the Women and Newborn Health Service Human Research Ethics Committee (201535EW) and was conducted at King Edward Memorial Hospital, Perth, the largest Tertiary Obstetrics Hospital in Western Australia (WA), over the period 2015–2017 [384]. Specimens were self-collected from two sites: vaginal and rectal on e-swabs (Copan, Italy) by 814 pregnant women at ≤22 weeks (*n*=814) and ≥33 weeks' (*n*=567) gestation. 24.9 % at ≤22 weeks. The GBS colonization rate within this WA pregnant cohort at ≤22 weeks was 24.9 % (n=202/814). A subset of these isolates (n=171/202), collected from pregnant women at ≤22 weeks' gestation (visit one) was used in this analysis. 171 GBS isolates (84 from vaginal site and 87 from rectal) were sent as purified isolates on e-swab by Dr Matthew S. Payne (research lead for this study) to Cardiff University. Metadata included date of collection of specimens, age of pregnant women, weeks of gestation and serotype information. The isolates were cultured and assessed for AST in Cardiff university.

#### 5.3.2.1 Serotype distribution:

Of 171, 84 GBS isolates were recovered from vaginal site and 87 from rectal sites including 21 isolates obtained from either vaginal (n=9) or rectal (n=12) site of pregnant women at ≤22 weeks' gestation (**Figure 5.2a**). Eight serotypes (Ia, Ib, II-VI, VIII, IX) and 6 non-typeable (NT) serotypes were identified in total, with serotype Ia, II, III and V as the most prevalent serotypes identified at both sites (**Figure 5.2a**). GBS colonizing vagina had the following serotype prevalences order: Ia (22/84, 26.1%),

preceding serotype III and V (18/84, 21.4%), II (13/84, 15.4%), Ib (5/84, 5.9%), VI (3/84, 3.5%), and IV (2/84, 2.3%), VIII, IX and NT (1/84, 1.9%), each **(Figure 5.2a**). While GBS colonizing rectal site of pregnant women had following serotype prevalences order: V (20/87, 22.9%), preceding Ia and III (17/87, 19.5%), II (14/87, 16%), VI (6/87, 6.8%), Ib and NT (4/87, 4.5%), IV (3/87, 3.4%), VIII and IX (1/87, 1.1%) (**Figure 5.2a**). 165 isolates recovered from both vaginal and rectal sites were of same serotypes, but three maternal pair of isolates (n=6) from both sites showed different serotype suggesting co-colonization of same mother with different serotypes. The different serotypes involved II and VI, Ib and V, V and III observed for GBS colonizing vaginal and rectal sites of pregnant women, respectively. Overall, at ≤22 weeks' gestation serotype Ia (22.8%), V (22.2%), III (20.5%), II (15.8%), Ib and VI (5.3%), IV and NT (2.9%), VIII and IX (1.2%) were identified in order of highest to lowest prevalence (**Figure 5.2b**).

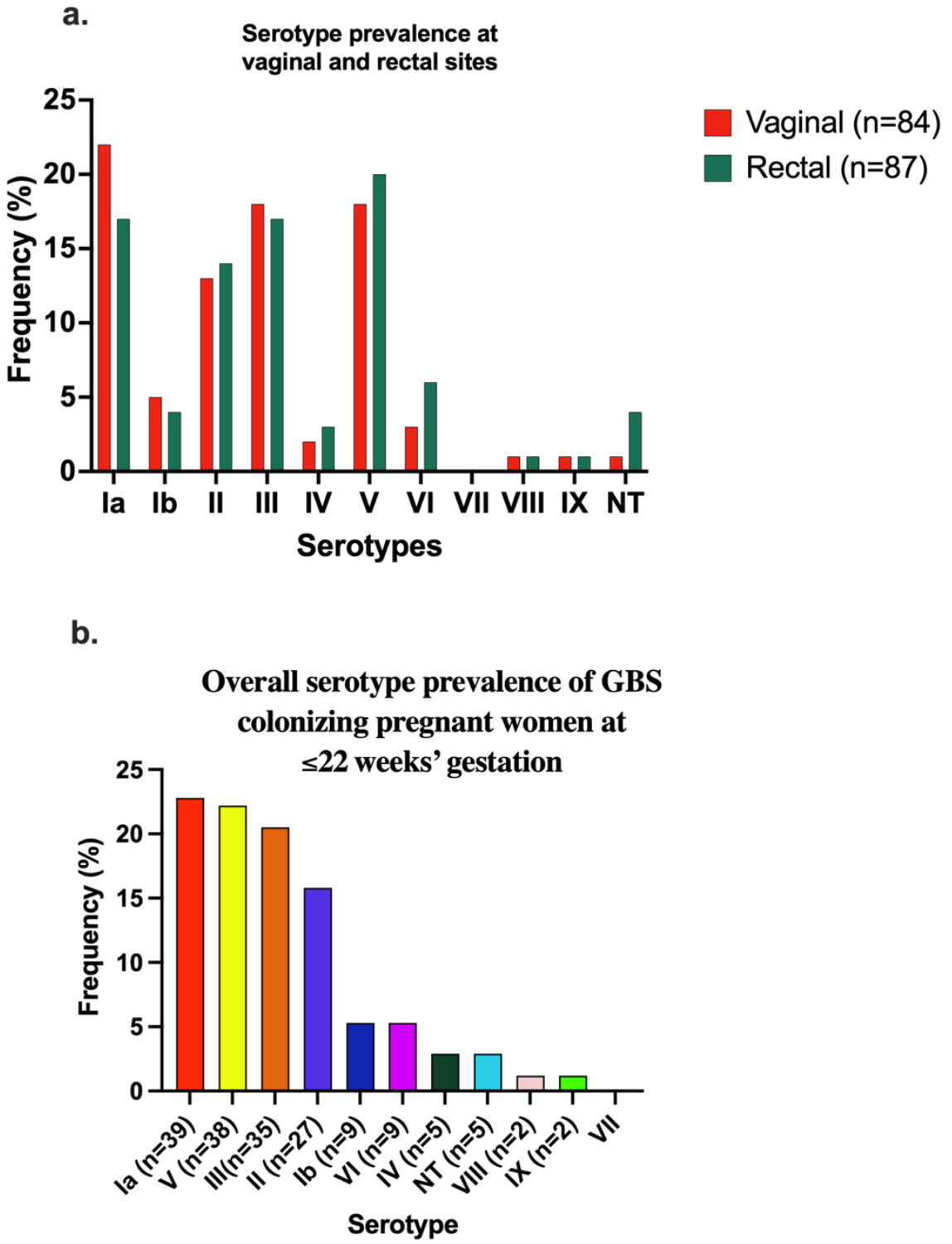



**Figure 5. 2. GBS antenatal serotype distribution at vaginal and rectal sites vs overall prevalence identified. a)** Serotype distribution of 171 GBS colonizing mothers and comparison across two specimens collected at visit one (≤22 weeks' gestation) and their respective site of isolation; **b)** overall serotype prevalence identified among pregnant women at visit one (≤22 weeks' gestation).

#### 5.3.2.2 Phenotypic AMR profiling

All 171 (100%) GBS isolates recovered from Australian pregnant women at ≤22 weeks' gestation were found susceptible to three antibiotics: gentamicin, ampicillin, and vancomycin upon AST **(Table 5.2)**. Levofloxacin resistance was observed in 5/171 (2.9%) of the isolates including three isolates recovered from rectal source (VMB0036_R, VMB0061_R and VMB0183_R) and two of which were from the matching vaginal swab (VMB0061_V and VMB0183_V). Only (3/171, 1.7%) GBS isolates were chloramphenicol resistant (VMB0069_V, VMB00378_R and VMB0422_V). Erythromycin and clindamycin resistance were observed in 23/171 (13.4%) and 15/171 (8.7%) of the GBS isolates respectively, including 12 (7%) isolates mediating combined resistance to erythromycin and clindamycin. Only 32/171 (18.7%) of the GBS isolates from pregnant women were sensitive to tetracycline **(Table 5.2)** and 27/171 (15.7%) of the GBS isolates remained susceptible to all antibiotics tested. No WGS was performed on the Australian GBS dataset as per the original plan due to Covid pandemic in UK that shut down Cardiff University lab for almost a year.

## 5.4 Discussion

In this study, iGBS from Brazilian adults and GBS from Australian colonized pregnant women were assessed to determine the serotype distribution and antibiotic resistance profiles of GBS populations circulating in two geographic regions. Although, limited datasets were available from local region of Brazil (n=26), and from Australia (n=171), this provides detailed molecular epidemiology data for both locations.

In this study, among Brazilian adults, the most prevalent GBS serotypes identified were serotype Ia (46.1%), V (38.4%), and Ib (7.6%). Previous studies conducted in the same southeast region of Brazil, also showed serotype Ia as the most prevalent serotype, but followed by serotypes II and Ib [371], however in the current dataset the second predominant serotype observed was serotype V, which is consistent with the serotype distribution observed in other southeast regional study conducted in São Paulo, Brazil in 2015–2016 [51], and different global studies where serotype Ia and V were identified

as dominant invasive isolates in non-pregnant cases [114, 152, 602]. Serotype III is the most dominant invasive clone accounting for the majority of late-onset meningitis cases in neonates [114, 269, 292, 603, 604], which is consistent with the low occurrence rate of serotype III (3.8%) identified in the current adult dataset, in 2002, serotype III was identified as the most abundant serotype in pregnant women of Rio de Janeiro, Brazil [375] however, in later studies serotype Ia was reported more [376], suggesting an increased prevalence in serotype Ia may be related with a period of time.

Among ≤22 weeks pregnant women of Australia, serotype Ia, III and V were the most common serotypes, identified at both vaginal and rectal sites, however, differences in order of serotype prevalence were identified at both sites. Serotype Ia (26.1%), preceding serotype III and V (21.4%) were identified prevalent in GBS obtained from vaginal sites, supporting the thought of serotype III prevailing and causing severe GBS infection among neonates due to the transmission of mother GBS to new-born during vaginal delivery [593]. At rectal sites, serotype V (22.9%) preceding Ia and III (19.5%) were found prevalent, the difference arises due to the co-colonization of same mother with different serotypes, which has been described earlier in African studies, where two to three serotypes have been detected from the same women [598, 605]. Likewise, a combination of two or three GBS serotypes per participant was observed in pregnant women in an Australian (Brisbane) cohort, where only 10 participants were colonized by the same serotype at both vaginal and rectal sites compared to 22 that showed different serotypes [595]. This contrasts with the findings of this study, where only 3 mothers had the different serotypes occurred at each site. Overall, a predominance of serotypes Ia, III, and V, representing more than 20 % of the positive isolates combined, and additionally Ib, II, VI, IV, VIII and IX were identified. These data are in accordance with the globally predominant serotypes, Ia, Ib, II, III and V [339]. Following comparison to the limited Australian data available, similarities with a study by Zhao *et al* [383] assessing women of child-bearing age in Australia and New Zealand were observed, including Ia (31%), III (30%), V (19%), Ib (16%), II (7%), IV (2%), VI (2%) and VII (1%). Unlike this study, Zhao *et al* identified serotype VII, but not VIII and IX because this study was conducted before the discovery of serotype IX [383]. Other studies showed a predominance of serotype

III [382, 595-597], with one study focusing on sterile site-invasive isolates and not carriage isolates [596]. In comparison to other Australian data, a lower rate of serotype III (20.5%) in our study particularly in comparison to the Brisbane cohort (43.6%) [595], but greater rates of serotype II (15.8%), with other areas around Australia and Australasia having reported this serotype as ≤10 % of positive isolates [382, 383, 595, 597] was found. Serotype IX was identified for the first time in this dataset, which has never been isolated from antenatal samples from other regions of Australia.

Penicillin is the first choice of drug to treat GBS infections and for IAP, however, few cases of reduced penicillin susceptibility have been reported in last two decades in the United States [177], Africa [606], Colombia [607], Central Italy [608], Japan [171], Scotland [176], and Canada [179], hence the question arises if these resistance phenotypes are independently acquired by some GBS isolates due to sporadic mutations, or it is more likely to be an emergence of β-lactam-resistant GBS clones, or even to both events, remains open. In this study, the consistent sensitivity of Brazilian and Australian GBS isolates to beta-lactam antibiotics (penicillin, ampicillin and gentamicin) was observed which supports the findings of previously published data on GBS isolates in Brazil [378] and Australia [384], as well as in several regions in the world [232, 609]. However, among our UK adult GBS cohort, a single isolate (1/193, 0.5%) was resistant to high level gentamicin (**Table 5.2**).

The resistance rate of levofloxacin, an antibiotic of fluoroquinolone class, was also evaluated. Generally, a low level of levofloxacin resistance were reported in GBS from different geographical location, for e.g. in Italy [551], Japan [610] and Taiwan [611]. Similar to these observations, low resistance rates of levofloxacin were also observed in our UK and Australian dataset with a rate of 2 and 2.9%, respectively whereas no GBS isolate resistant to levofloxacin was observed in Brazilian dataset (**Table 5.2**). Fluoroquinolone resistance in GBS is mediated by somatic mutation of genes and therefore cannot be transferred to other isolates through recombination and must exist as sporadic mutation or clonal expansion.

All Brazilian (n=26), Australian (n=171) GBS and UK (n=193) isolates were found susceptible to vancomycin, an alternative to penicillin for penicillin allergic patient in UK, however in Australia lincomycin is recently recommended as a substitute while in Brazil, clindamycin is the first alternative given that a susceptibility test is done, if clindamycin is resistant then vancomycin is the recommended drug of choice [612]. High susceptibility profiles of vancomycin in our three datasets (Australia, Brazil and UK) (**Table 5.2**) are in agreement with the previous reports demonstrating 100% susceptibility of vancomycin in GBS isolates isolated from pregnant women [613], neonates [614] and adults [615] and supporting its use as an alternative drug in penicillin allergic patients.

No chloramphenicol resistance was observed among Brazilian GBS isolates, consistent with the findings of an earlier Brazilian GBS study [48]. However, a few (1.7%) of Australian pregnant women GBS isolates and our UK GBS isolates (1.5%) were found resistant to chloramphenicol (**Table 5.2**) which disagrees to 0% chloramphenicol resistance observed in recent studies on GBS isolates recovered from pregnant women in Ethiopia [616, 617] and Serbia [312]. Despite the rare occurrence of chloramphenicol resistance in GBS, chloramphenicol resistant MDR GBS has also been reported [242, 543, 618]. High resistance rate to tetracycline was detected in our three datasets including GBS colonizing Australian pregnant women (81.2%), and GBS isolates obtained from UK (90.6%), and Brazilian adults (80.7%) (**Table 5.2**). The *tet*M gene was widely disseminated in both Brazilian and UK dataset, consistent with the previous report of Brazil, as well as observed in other countries [232, 380, 397, 609].

The low resistance rate to erythromycin (7.6%, n=2) and clindamycin (3.8%, n=1) among Brazil GBS isolates found in this study, is in agreement with a previous study of Brazil where 8% and 4% GBS isolates were found erythromycin and clindamycin resistant [380]. Compared to this, higher erythromycin and clindamycin resistance rates were observed in our UK GBS isolates (20.2% and 11.8%), and in other studies published from Asia, Europe, United States and Canada [397, 619, 620]. Analogous to Brazilian GBS, Australian pregnant women isolated GBS also showed low resistance to

erythromycin (13.4%) and clindamycin (8.7%) (**Table 5.2**), which is consistent with GBS erythromycin and clindamycin resistance rates published in a recent report from Australia [621] and from some other countries in Latin America [260, 263, 268, 269]. High resistance rates to erythromycin and clindamycin in Asia, Europe, the United States, and Canada [397, 619, 620] could be attributed to the more intense clinical use of these antibiotics. No correlation between erythromycin resistance and GBS serotypes was observed in the Australian dataset. However, the Brazilian GBS serotype Ia (2/3, 66.6%) and V (1/3, 33.3%) isolates were found to contain the macrolide resistance, and the high frequency of these two serotype (Ia and V) associated with macrolide resistance has been previously documented [551].

In general, data on the cases of GBS causing disease and distribution of GBS serotypes and antimicrobial susceptibility profiles among adults living in different Brazilian locations is still largely unknown, as the information available is usually related to small groups of patients mainly pregnant women and short-term observations. In Australia similarly, limited sporadic and regional data with small sample size is available on GBS colonizing pregnant women, their serotype distribution and antibiotic resistance patterns. Addressing data gaps especially by world region and some at-risk populations (notably stillbirths) is fundamental to evidence-based decision-making during vaccine design. Surveillance of circulating, disease-causing serotypes is useful to inform vaccines not targeting surface proteins. The present study used the small size available GBS collections from a single site of Brazilian adults (n=26), and Australian pregnant women (n=171) and demonstrate serotype Ia, Ib, II, III and V are the most prevalent suggesting a hexavalent vaccine (serotypes Ia, Ib, II, III, IV and V), currently under clinical trial could provide comprehensive cover for all at-risk populations. Further the clindamycin resistance rate identified in both Brazilian and Australian GBS datasets corroborate the prerequisite to verify antibiotic susceptibility against clindamycin, the main alternative antibiotics to beta-lactams often suggested for IAP therapy. These data can contribute to help in devising prevention and treatment strategies for GBS infections in these regions.

# Chapter 6:

## Population structure, virulence factors, resistance determinants and pan-genome wide association study (pan-GWAS) of adult isolated *Streptococcus agalactiae* from Brazil, Canada, UK, and United States

## 6.1. Summary

GBS is an opportunistic bacterium that infect infants and elderly people and a normal commensal of the gastrointestinal and vaginal epithelium of a substantial proportion (30%) of healthy women. Out of 447 adult GBS genomes analysed for pan-genome wide association study, Brazilian (n=26), and United Kingdom (UK, n=193) isolates were sequenced in this thesis while the rest of the adult GBS genomes from Canada (n=134), and United States (n=94) were retrieved from public databases. A total of 417 (93.2%) GBS isolates fell into five clonal complexes (CC1, /10, CC17, CC19, and CC23). CC1 was the largest CC (204/417,48.9%) followed by CC17 (96/417, 23%) and CC23 (52/417, 12.4%). Specific associations were identified between surface expressed virulence determinants, STs, serotypes and particular CCs. For example, CC1 lineage was defined by serotype V/ST1/*alp*1+*srr*-1/PI-1+PI-2a; CC17 by serotype III/ST17/*hvg*A+*rib*+*srr*-2/PI-1+PI-2b; CC19 by serotype III/ST19/*srr*-1+*rib*/PI-1+PI-2a; /10 by serotype II+Ib/*srr*-1+*bca*/PI-1+PI-2a; and CC23 by Ia/ST23/*srr*-1+*alp*1/PI-1-2a. Likewise, specific correlation was observed between certain antibiotic resistance and CC such as the majority of the macrolide resistant GBS strains (114/174, 65.6%) carrying *erm*B, *erm*A and *mef*A/*msr*D determinants were clustered in CC1 and were mostly serotype V (108/114, 94.7%), while all *erm*T (n=19) isolates were restricted to serotype III CC17 macrolide resistant strains. MDR GBS strains carrying antimicrobial resistance genes conferring resistance to four classes of antibiotics (including macrolides, tetracycline, chloramphenicol, and aminoglycoside) were confined to CC19. Further each CC was characterized by specific genes that provides selective advantage to GBS for improved colonization, invasion, virulence, and survival within host. In this study, 97 CC-specific genes associated (excluding hypothetical proteins) with virulence, metabolism, and regulation of cellular mechanisms were identified that may explain the differential virulence potential of the isolates confined to specific CCs. Among CC17 and CC23 GBS isolates, micronutrient uptake proteins (iron and manganese), two component systems, accessory secondary proteins, pilus and quorum-sensing genes were identified which were absent in less invasive lineages (CC1, CC8/CC10 and CC19). Metal resistance genes (arsenic, cadmium, and copper) and CRISPR associated genes (cas1/cas2) were confined

to CC8/CC10 whereas the type IV secretory protein (VirD4) was significantly associated to CC19. Collectively this analysis underlines the lineage-specific basis of GBS niche adaptation and virulence.

## 6.2. Introduction:

With the advancement in WGS, Genome-wide association studies (GWAS) have become an increasingly important approach to investigate the statistical link between genotypic variation and bacterial phenotypes [622, 623]. Single-nucleotide polymorphisms (SNPs) are the most studied genetic variants in GWAS [624], although copy-number variants or sequence variations can also be considered as observed in the human genome [625]. GWAS usually report blocks of correlated SNPs that all show a statistically significant association with the trait of interest, known as genomic risk loci [625]. After 15 years of GWAS [626], many replicated genomic risk loci have significantly linked with diseases and traits [626]. Numerous GWAS studies were conducted on a range of bacteria including *S. pneumoniae* [627-629], *Staphylococcus aureus* [630, 631], *Neisseria meningitidis* [627], *Campylobacter* [632], *Mycobacterium tuberculosis* but inadequate studies are available for GBS [362]. These studies have identified genetic variation linked with disease susceptibility [627-631, 633], disease progression [627], nutrient synthesis [632], host adaptation [634], cell adhesion, immune evasion, and carriage duration [635, 636], virulence [634], and antimicrobial resistance [637-641]. Also, studies comparing carriage and disease isolates of different bacteria including GBS have identified variants associated with invasiveness [633, 642, 643] while studies comparing invasive isolates from different tissues, such as blood and CSF, have yielded no differences [629, 630, 643]. These inconsistencies reflect differences in the analytical methods, data set sizes, geographical settings, and control for confounders, such as capsular diversity, geographical origin and strain population structure. The latter is especially problematic in bacterial species with highly structured populations such as GBS [644], but may be less severe in highly recombinogenic species such as *S. pneumoniae* in which the genetic pool is frequently shuffled [645].

Five major Clonal Complexes (CCs) have been identified by MLST in human GBS infections including CC1, CC10, CC17, CC19 and CC23 [1, 536]. Within recent studies, it became evident that GBS strains fit

in to certain CCs that have a greater ability to cause invasive disease, while other CCs were associated with asymptotic carriage [281, 362]. For example, the strains belonging to CC1, CC19 and CC23 are more frequently found in pregnant women, perhaps due to adaptation to vaginal mucosa colonisation and a limited invasive potential for neonates [142]. By contrast, CC17, mostly serotype III, and ST-17, strains have a greater association with invasive neonatal infections whereas serotype V ST-1 strains are predominant in adult GBS invasive diseases [281]. The ability of GBS to adhere to the mucosal epithelium defines the colonisation and persistence of GBS in different host niches [281, 646], using bacterial adhesins including fibrinogen binding protein (Fbs), the C5a peptidase (ScpB) and the immunogenic bacterial adhesin (BibA) [308, 647, 648]. Colonization persists with the formation of biofilm, which plays a significant role in the phenotype switch from commensal to pathogen. Colonization further boosts by bacterial capsule and type IIa pili [299, 649]. In a recent study, deletion of a Biofilm regulatory protein (BrpA) encoding gene shown to impair both the ability of the bacterium to colonise and invade the murine host and the biofilm formation [650]. One characteristic of GBS that renders it particularly virulent during the perinatal period is its ability to invade the chorioamniotic membranes and persist in amniotic fluid, which is nutritionally deplete and rich in fetal immunologic factors such as antimicrobial peptide [651]. A recent genome-wide fitness analysis of GBS in human amniotic fluid revealed a transcription factor (MrvR) that controlled multiple virulence traits and conferred a significant fitness benefit to GBS survival in amniotic fluid [651].

Different CC isolates expresses different virulence factors for eg: Fbs proteins carried by the hypervirulent lineage CC17 and characterised by specific deletions and frameshift mutations that change the sequence or expression rate [652]. In addition to Fbs, CC17 isolates often contain of hypervirulent *hvg*A, *rib* and PI-2b genes [362] known to confer increased virulence and a selective advantage in the human host either by reducing host immune responses and/or increasing their dissemination potential [653]. The comparative analysis of 923 GBS genomes belonging to CC1, CC19, and CC23 revealed that the evolution of CC17 is

distinct from that of other human-adapted lineages and recurrently targets functions related to nucleotide and amino acid metabolism, cell adhesion, regulation, and immune evasion [402]. The most distinctive features of disease-specific CC17 isolates were frequent mutations in the virulence-associated CovS and Stk1 kinases, underscoring the crucial role of the entire CovRS regulatory pathway in modulating the pathogenicity of GBS [402]. Recently, pan-genome wide association study of 1988 GBS genomes, identified genes encoding for pilus, quorum sensing proteins, and proteins for the uptake of ions and micronutrients only limited to more invasive CC17 and CC23 isolates, while absent in less invasive lineages CC1, CC8/CC10 and CC19 [362]. Moreover, functionally important allelic variants in CC17 associated with either carriage or disease [362].

The presence or absence of genes that are inherited or acquired through horizontal gene transfer can be associated to several bacterial phenotypes [654]. The complete set of all genes among a collection of genomes is known as pan-genome [97, 655]. The pan-genome is comprised of the core genome including genes identified in all strains within that collection and the accessory genome containing 'dispensable' genes present in a subset of the strains [97]. Pan-genome can be classified into 4 classes [656]: (1) persistent genome – also called soft and hard core, for the gene families present in almost all genomes within the species [657]; (2) shell genome, for gene families present at intermediate frequencies in the species; (3) cloud genome, for gene families present at low frequency in the species [656]. Hard core genes are found in >99% genomes, soft core genes are found in 95–99% of genomes, shell genes are found in 15–95%, while cloud genes are present in less than 15% of genomes [658]. The influential paper describing the concept of a pan-genome was based on an analysis of eight GBS genome sequences [97]. GBS therefore holds a special place in the early transition to the post-genomic era for bacteria [659]. It was also the first organism described to have an “open” pan-genome; rarefaction analysis predicted that, even with an arbitrarily large number of genome sequences, every new genome sequence would contribute an extra 33 genes that had not previously been seen in any other GBS [97, 655].

The aim of this study was to conduct a pan-genome wide association study (pan-GWAS) and perform phylogenetic analysis of 447 adult GBS genomes, including publicly genomes from Canada [313] and United States [76] together with UK and Brazil GBS genomes that were sequenced in this study.

## 6.3. Results:

### 6.3.1. Geographic distribution of GBS genomes used in this study:

228 publicly available Illumina HiSeq sequences from the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena) for GBS adult isolates originating in Canada (n=134) including serotype V (n=68) and serotype III (n=66) strains; in the United States (n=94) including all serotype V strains; together with 193 UK and 26 Brazilian adult GBS Illumina Hiseq sequences (sequenced earlier as a part of this thesis) were used in this analysis (**Figure 6.1 and Table 6.1**). Serotype V isolates from United Stated [76] and serotype V and III isolates from Canada [313] were deliberately chosen knowing that CC17 serotype III strains are most frequently found both in mothers asymptomatically carrying GBS and in infected new-borns [122]. And are also exceptionally correlated with cases of LOD and meningitis [660]. Whereas serotype V CC1 strains are common causes of GBS infection in adults [578, 661]. Therefore, pan-genome wide association study of these clones was undertaken to decode the distinctive collections of genes conferring increased potential for their persistence and virulence. Among 26 Brazilian adult GBS, five serotypes were identified: serotype Ia (12/26, 46.1%), V (10/26, 38.4%), followed by serotype Ib (2/26, 7.6%), II and III (1/26, 3.8%) (**Figure 6.2A**). While among UK adult GBS population eight serotypes were identified: III (26.9%, 52/193), Ia (26.4%, 51/193), V (15%, 29/193), II (13.9%, 27/193), Ib (11.4%, 22/193), IV (3.1%, 6/193), VI (1.5%, 3/193) and IX (1.5%, 3/193) (**Figure 6.2A**). Meta-data consisting of country of origin, year of isolation, capsular serotype, MLST, and accession number are listed in (**Table 6.1**). In total, 447 invasive GBS genome sequences were used for the analysis.

WORK FLOW of PAN-GWAS

| Country of isolation | GBS sequences used in this study (n=447) | Sequencing technology | Sequenced at Cardiff University | ENA accession |
|---|---|---|---|---|
| Canada | 134 | Illumina HiSeq | No | PRJNA295774 |
| US | 94 | Illumina HiSeq | No | PRJNA274384 |
| Brazil | 26 | Illumina HiSeq | Yes | No** |
| UK | 193 | Illumina HiSeq | Yes | PRJEB18093 |

**The Brazilian GBS sequences are currently under analysis for a multinational collaboration and the foreign collaborating partners have not given permission for these sequences to be submitted to an open database, but their permission (and subsequent submission) will be completed once the analysis has led to a manuscript draft that is nearing publication readiness.

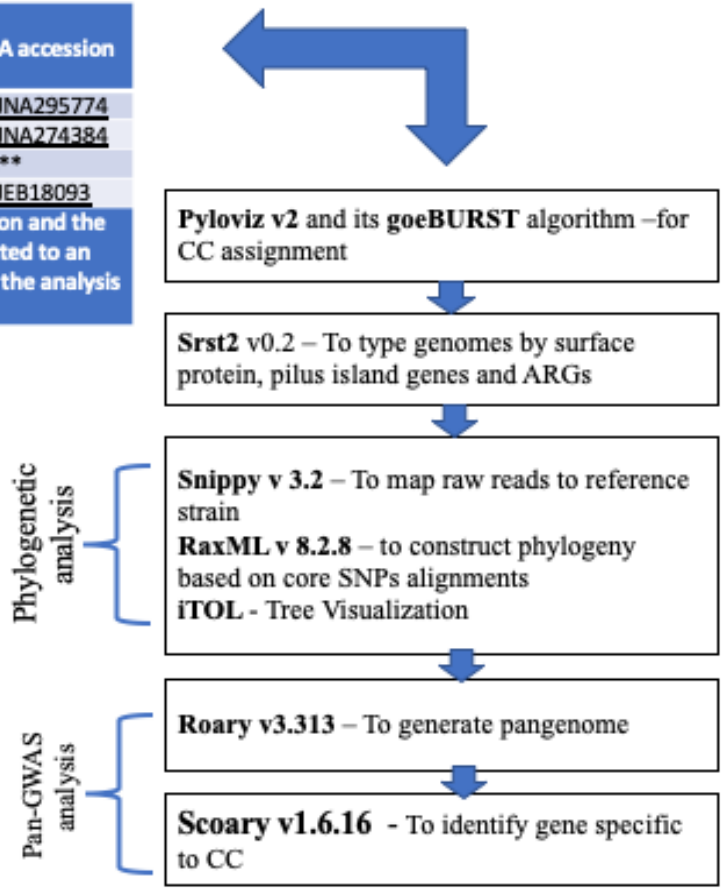



**Figure 6. 1. Workflow diagram of Pan–GWAS analysis performed in this study.**

**Table 6. 1. Description of 228 publicly available GBS genomes used in this study.** Information for each isolate consist of their source (invasive or non-invasive), country of origin, capsular serotype, year of isolation, MLST and accession number. Description of 193 UK and 26 Brazilian GBS are available in **Table 4.1** and **Table 5.1**.

| Isolate | Source | Country | Serotype | Year | ST | ENA-accession number |
|---|---|---|---|---|---|---|
| NGBS10 | invasive | Canada | V | 2009 | 1 | SRR1790740 |
| NGBS107 | invasive | Canada | V | 2010 | 1 | SRR1790741 |
| NGBS117 | invasive | Canada | V | 2010 | 1 | SRR1790743 |
| NGBS171 | invasive | Canada | V | 2010 | 1 | SRR1790746 |
| NGBS172 | invasive | Canada | V | 2010 | 1 | SRR1790747 |
| NGBS177 | invasive | Canada | V | 2010 | 1 | SRR1790748 |
| NGBS180 | invasive | Canada | V | 2010 | 1 | SRR1790749 |
| NGBS200 | invasive | Canada | V | 2010 | 1 | SRR1790750 |
| NGBS21 | invasive | Canada | V | 2009 | 1 | SRR1790751 |
| NGBS210 | invasive | Canada | V | 2011 | 1 | SRR1790752 |
| NGBS22 | invasive | Canada | V | 2009 | 1 | SRR1790753 |
| NGBS246 | invasive | Canada | V | 2011 | 1 | SRR1790758 |
| NGBS25 | invasive | Canada | V | 2009 | 1 | SRR1790759 |
| NGBS267 | invasive | Canada | V | 2010 | 1 | SRR1790760 |
| NGBS272 | invasive | Canada | V | 2011 | 1 | SRR1790761 |
| NGBS273 | invasive | Canada | V | 2011 | 1 | SRR1790762 |
| NGBS275 | invasive | Canada | V | 2010 | 1 | SRR1790763 |
| NGBS279 | invasive | Canada | V | 2010 | 1 | SRR1790764 |
| NGBS283 | invasive | Canada | V | 2010 | 1 | SRR1790766 |
| NGBS287 | invasive | Canada | V | 2010 | 1 | SRR1790767 |
| NGBS288 | invasive | Canada | V | 2010 | 1 | SRR1790768 |
| NGBS298 | invasive | Canada | V | 2011 | 1 | SRR1790769 |
| NGBS30 | invasive | Canada | V | 2010 | 1 | SRR1790770 |
| NGBS303 | invasive | Canada | V | 2010 | 1 | SRR1790771 |
| NGBS321 | invasive | Canada | V | 2011 | 1 | SRR1790772 |
| NGBS325 | invasive | Canada | V | 2011 | 1 | SRR1790774 |
| NGBS330 | invasive | Canada | V | 2011 | 1 | SRR1790775 |
| NGBS331 | invasive | Canada | V | 2011 | 1 | SRR1790776 |
| NGBS332 | invasive | Canada | V | 2011 | 1 | SRR1790777 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NGBS348 | invasive | Canada | V | 2011 | 1 | SRR1790779 |
| NGBS35 | invasive | Canada | V | 2010 | 1 | SRR1790780 |
| NGBS357 | invasive | Canada | V | 2011 | 1 | SRR1790781 |
| NGBS359 | invasive | Canada | V | 2011 | 1 | SRR1790782 |
| NGBS360 | invasive | Canada | V | 2011 | 1 | SRR1790783 |
| NGBS372 | invasive | Canada | V | 2011 | 1 | SRR1790784 |
| NGBS380 | invasive | Canada | V | 2011 | 1 | SRR1790785 |
| NGBS381 | invasive | Canada | V | 2011 | 1 | SRR1790786 |
| NGBS411 | invasive | Canada | V | 2011 | 1 | SRR1790787 |
| NGBS418 | invasive | Canada | V | 2011 | 1 | SRR1790788 |
| NGBS425 | invasive | Canada | V | 2011 | 1 | SRR1790789 |
| NGBS444 | invasive | Canada | V | 2011 | 1 | SRR1790792 |
| NGBS462 | invasive | Canada | V | 2011 | 1 | SRR1790793 |
| NGBS492 | invasive | Canada | V | 2012 | 1 | SRR1790794 |
| NGBS497 | invasive | Canada | V | 2012 | 1 | SRR1790796 |
| NGBS499 | invasive | Canada | V | 2012 | 1 | SRR1790797 |
| NGBS519 | invasive | Canada | V | 2012 | 1 | SRR1790799 |
| NGBS536 | invasive | Canada | V | 2012 | 1 | SRR1790800 |
| NGBS54 | invasive | Canada | V | 2010 | 1 | SRR1790801 |
| NGBS553 | invasive | Canada | V | 2012 | 1 | SRR1790802 |
| NGBS558 | invasive | Canada | V | 2012 | 1 | SRR1790803 |
| NGBS561 | invasive | Canada | V | 2012 | 1 | SRR1790804 |
| NGBS579 | invasive | Canada | V | 2012 | 1 | SRR1790806 |
| NGBS580 | invasive | Canada | V | 2012 | 1 | SRR1790807 |
| NGBS586 | invasive | Canada | V | 2012 | 1 | SRR1790808 |
| NGBS604 | invasive | Canada | V | 2012 | 1 | SRR1790809 |
| NGBS624 | invasive | Canada | V | 2012 | 1 | SRR1790810 |
| NGBS63 | invasive | Canada | V | 2010 | 1 | SRR1790811 |
| NGBS630 | invasive | Canada | V | 2012 | 1 | SRR1790812 |
| NGBS68 | invasive | Canada | V | 2010 | 1 | SRR1790814 |
| NGBS8 | invasive | Canada | V | 2009 | 1 | SRR1790815 |
| NGBS9 | invasive | Canada | V | 2010 | 1 | SRR1790816 |
| NGBS92 | invasive | Canada | V | 2010 | 1 | SRR1790817 |
| NGBS93 | invasive | Canada | V | 2010 | 1 | SRR1790818 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NGBS94 | invasive | Canada | V | 2010 | 1 | SRR1790819 |
| NGBS99 | invasive | Canada | V | 2010 | 1 | SRR1790820 |
| NGBS003 | invasive | Canada | III | 2009 | 2 | SRR2981633 |
| NGBS050 | invasive | Canada | III | 2010 | 17 | SRR2451874 |
| NGBS069 | invasive | Canada | III | 2010 | 17 | SRR2451875 |
| NGBS079 | invasive | Canada | III | 2010 | 17 | SRR2451877 |
| NGBS082 | invasive | Canada | III | 2010 | 17 | SRR2451879 |
| NGBS126 | invasive | Canada | III | 2010 | 17 | SRR2451883 |
| NGBS128 | invasive | Canada | III | 2010 | 17 | SRR2451884 |
| NGBS169 | invasive | Canada | III | 2010 | 17 | SRR2451892 |
| NGBS186 | invasive | Canada | III | 2010 | 17 | SRR2451893 |
| NGBS205 | invasive | Canada | III | 2011 | 17 | SRR2451894 |
| NGBS220 | invasive | Canada | III | 2011 | 17 | SRR2451895 |
| NGBS222 | invasive | Canada | III | 2010 | 17 | SRR2451896 |
| NGBS238 | invasive | Canada | III | 2011 | 17 | SRR2451897 |
| NGBS239 | invasive | Canada | III | 2011 | 17 | SRR2451898 |
| NGBS250 | invasive | Canada | III | 2011 | 17 | SRR2451899 |
| NGBS277 | invasive | Canada | III | 2010 | 17 | SRR2451901 |
| NGBS282 | invasive | Canada | III | 2010 | 17 | SRR2451902 |
| NGBS291 | invasive | Canada | III | 2010 | 17 | SRR2451903 |
| NGBS296 | invasive | Canada | III | 2010 | 17 | SRR2451904 |
| NGBS297 | invasive | Canada | III | 2010 | 17 | SRR2451905 |
| NGBS299 | invasive | Canada | III | 2011 | 17 | SRR2451906 |
| NGBS306 | invasive | Canada | III | 2011 | 17 | SRR2451907 |
| NGBS312 | invasive | Canada | III | 2011 | 17 | SRR2451908 |
| NGBS356 | invasive | Canada | III | 2011 | 17 | SRR2451914 |
| NGBS361 | invasive | Canada | III | 2011 | 17 | SRR2451915 |
| NGBS362 | invasive | Canada | III | 2011 | 17 | SRR2451916 |
| NGBS368 | invasive | Canada | III | 2011 | 17 | SRR2451917 |
| NGBS374 | invasive | Canada | III | 2011 | 17 | SRR2451919 |
| NGBS377 | invasive | Canada | III | 2011 | 17 | SRR2451920 |
| NGBS398 | invasive | Canada | III | 2011 | 17 | SRR2451922 |
| NGBS403 | invasive | Canada | III | 2011 | 17 | SRR2451923 |
| NGBS421 | invasive | Canada | III | 2011 | 17 | SRR2451925 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NGBS422 | invasive | Canada | III | 2011 | 17 | SRR2451926 |
| NGBS431 | invasive | Canada | III | 2011 | 17 | SRR2451927 |
| NGBS456 | invasive | Canada | III | 2011 | 17 | SRR2451929 |
| NGBS464 | invasive | Canada | III | 2011 | 17 | SRR2451930 |
| NGBS469 | invasive | Canada | III | 2011 | 17 | SRR2451931 |
| NGBS470 | invasive | Canada | III | 2011 | 17 | SRR2451932 |
| NGBS483 | invasive | Canada | III | 2011 | 17 | SRR2451933 |
| NGBS485 | invasive | Canada | III | 2011 | 17 | SRR2451934 |
| NGBS486 | invasive | Canada | III | 2011 | 17 | SRR2451935 |
| NGBS500 | invasive | Canada | III | 2012 | 17 | SRR2451936 |
| NGBS501 | invasive | Canada | III | 2012 | 17 | SRR2451937 |
| NGBS515 | invasive | Canada | III | 2012 | 17 | SRR2451939 |
| NGBS534 | invasive | Canada | III | 2012 | 17 | SRR2451942 |
| NGBS551 | invasive | Canada | III | 2012 | 17 | SRR2451943 |
| NGBS583 | invasive | Canada | III | 2012 | 17 | SRR2451945 |
| NGBS593 | invasive | Canada | III | 2012 | 17 | SRR2451946 |
| NGBS594 | invasive | Canada | III | 2012 | 17 | SRR2451947 |
| NGBS596 | invasive | Canada | III | 2012 | 17 | SRR2451948 |
| NGBS607 | invasive | Canada | III | 2012 | 17 | SRR2451949 |
| NGBS608 | invasive | Canada | III | 2012 | 17 | SRR2451950 |
| NGBS609 | invasive | Canada | III | 2012 | 17 | SRR2451951 |
| NGBS613 | invasive | Canada | III | 2012 | 17 | SRR2451952 |
| NGBS618 | invasive | Canada | III | 2012 | 17 | SRR2451954 |
| NGBS632 | invasive | Canada | III | 2012 | 17 | SRR2451958 |
| NGBS636 | invasive | Canada | III | 2010 | 17 | SRR2451959 |
| NGBS641 | invasive | Canada | III | 2010 | 17 | SRR2451960 |
| NGBS644 | invasive | Canada | III | 2011 | 17 | SRR2451961 |
| NGBS650 | invasive | Canada | III | 2011 | 17 | SRR2451962 |
| NGBS502 | invasive | Canada | III | 2012 | 95 | SRR2451938 |
| NGBS531 | invasive | Canada | III | 2012 | 148 | SRR2451941 |
| NGBS622 | invasive | Canada | III | 2012 | 148 | SRR2451955 |
| NGBS318 | invasive | Canada | III | 2011 | 290 | SRR2451910 |
| NGBS241 | invasive | Canada | V | 2011 | 453 | SRR1790756 |
| NGBS327 | invasive | Canada | III | 2011 | 484 | SRR2451911 |

| | | | | | | |
|---|---|---|---|---|---|---|
| NGBS633 | invasive | Canada | V | 2012 | 531 | SRR1790813 |
| NGBS244 | invasive | Canada | V | 2011 | 871 | SRR1790757 |
| NGBS345 | invasive | Canada | III | 2011 | 874 | SRR2451913 |
| SS100 | invasive | USA | V | 2006 | 1 | SRR2981531 |
| SS102 | invasive | USA | V | 2006 | 1 | SRR2981532 |
| SS103 | invasive | USA | V | 2006 | 1 | SRR2981533 |
| SS104 | invasive | USA | V | 2006 | 1 | SRR2981534 |
| SS106 | invasive | USA | V | 2006 | 1 | SRR2981536 |
| SS107 | invasive | USA | V | 2006 | 1 | SRR2981537 |
| SS108 | invasive | USA | V | 2006 | 872 | SRR2981538 |
| SS109 | invasive | USA | V | 2006 | 1 | SRR2981539 |
| SS110 | invasive | USA | V | 2007 | 1 | SRR2981540 |
| SS111 | invasive | USA | V | 2007 | 1 | SRR2981541 |
| SS114 | invasive | USA | V | 2007 | 1 | SRR2981542 |
| SS115 | invasive | USA | V | 2007 | 1 | SRR2981543 |
| SS116 | invasive | USA | V | 2007 | 1 | SRR2981544 |
| SS119 | invasive | USA | V | 2008 | 1 | SRR2981546 |
| SS120 | invasive | USA | V | 2008 | 1 | SRR2981547 |
| SS122 | invasive | USA | V | 2009 | 1 | SRR2981548 |
| SS125 | invasive | USA | V | 2009 | 1 | SRR2981549 |
| SS126 | invasive | USA | V | 2009 | 1 | SRR2981550 |
| SS127 | invasive | USA | V | 2009 | 1 | SRR2981551 |
| SS129 | invasive | USA | V | 2009 | 1 | SRR2981552 |
| SS132 | invasive | USA | V | 2009 | 1 | SRR2981553 |
| SS133 | invasive | USA | V | 2009 | 1 | SRR2981554 |
| SS135 | invasive | USA | V | 2007 | 1 | SRR2981555 |
| SS136 | invasive | USA | V | 2007 | 297 | SRR2981556 |
| SS138 | invasive | USA | V | 2010 | 1 | SRR2981557 |
| SS140 | invasive | USA | V | 2008 | 1 | SRR2981558 |
| SS141 | invasive | USA | V | 2009 | 1 | SRR2981559 |
| SS143 | invasive | USA | V | 2010 | 1 | SRR2981560 |
| SS144 | invasive | USA | V | 2010 | 1 | SRR2981561 |
| SS145 | invasive | USA | V | 2010 | 1 | SRR2981562 |
| SS146 | invasive | USA | V | 2010 | 1 | SRR2981563 |

| | | | | | | |
|---|---|---|---|---|---|---|
| SS147 | invasive | USA | V | 2010 | 1 | SRR2981564 |
| SS148 | invasive | USA | V | 2011 | 1 | SRR2981565 |
| SS150 | invasive | USA | V | 2012 | 1 | SRR2981566 |
| SS151 | invasive | USA | V | 2012 | 1 | SRR2981567 |
| SS152 | invasive | USA | V | 2012 | 1 | SRR2981568 |
| SS31 | invasive | USA | V | 1992 | 1 | SRR2981569 |
| SS32 | invasive | USA | V | 1992 | 1 | SRR2981570 |
| SS33 | invasive | USA | V | 1994 | 1 | SRR2981571 |
| SS34 | invasive | USA | V | 1994 | 1 | SRR2981572 |
| SS35 | invasive | USA | V | 1994 | 1 | SRR2981573 |
| SS36 | invasive | USA | V | 1997 | 1 | SRR2981574 |
| SS37 | invasive | USA | V | 1997 | 1 | SRR2981575 |
| SS38 | invasive | USA | V | 1998 | 1 | SRR2981576 |
| SS39 | invasive | USA | V | 1998 | 1 | SRR2981577 |
| SS40 | invasive | USA | V | 1998 | 1 | SRR2981578 |
| SS41 | invasive | USA | V | 1998 | 1 | SRR2981579 |
| SS42 | invasive | USA | V | 1998 | 1 | SRR2981580 |
| SS43 | invasive | USA | V | 1998 | 1 | SRR2981581 |
| SS44 | invasive | USA | V | 1999 | 1 | SRR2981582 |
| SS45 | invasive | USA | V | 1999 | 1 | SRR2981583 |
| SS46 | invasive | USA | V | 1999 | 1 | SRR2981584 |
| SS47 | invasive | USA | V | 2000 | 1 | SRR2981585 |
| SS48 | invasive | USA | V | 2000 | 1 | SRR2981586 |
| SS49 | invasive | USA | V | 2000 | 873 | SRR2981587 |
| SS50 | invasive | USA | V | 2000 | 1 | SRR2981588 |
| SS51 | invasive | USA | V | 2000 | 1 | SRR2981589 |
| SS52 | invasive | USA | V | 2000 | 1 | SRR2981590 |
| SS53 | invasive | USA | V | 2000 | 1 | SRR2981591 |
| SS54 | invasive | USA | V | 2001 | 1 | SRR2981592 |
| SS56 | invasive | USA | V | 2001 | 1 | SRR2981593 |
| SS57 | invasive | USA | V | 2001 | 1 | SRR2981594 |
| SS58 | invasive | USA | V | 2001 | 1 | SRR2981595 |
| SS59 | invasive | USA | V | 2001 | 1 | SRR2981596 |
| SS60 | invasive | USA | V | 2001 | 1 | SRR2981597 |

| | | | | | | |
|---|---|---|---|---|---|---|
| SS61 | invasive | USA | V | 2002 | 1 | SRR2981598 |
| SS62 | invasive | USA | V | 2002 | 1 | SRR2981599 |
| SS63 | invasive | USA | V | 2002 | 1 | SRR2981600 |
| SS64 | invasive | USA | V | 2002 | 1 | SRR2981601 |
| SS65 | invasive | USA | V | 2002 | 1 | SRR2981602 |
| SS66 | invasive | USA | V | 2002 | 1 | SRR2981603 |
| SS67 | invasive | USA | V | 2002 | 1 | SRR2981604 |
| SS68 | invasive | USA | V | 2003 | 1 | SRR2981605 |
| SS69 | invasive | USA | V | 2003 | 1 | SRR2981606 |
| SS71 | invasive | USA | V | 2003 | 1 | SRR2981608 |
| SS72 | invasive | USA | V | 2003 | 1 | SRR2981609 |
| SS74 | invasive | USA | V | 2004 | 1 | SRR2981610 |
| SS75 | invasive | USA | V | 2004 | 1 | SRR2981611 |
| SS76 | invasive | USA | V | 2004 | 1 | SRR2981612 |
| SS77 | invasive | USA | V | 2004 | 1 | SRR2981613 |
| SS78 | invasive | USA | V | 2004 | 1 | SRR2981614 |
| SS79 | invasive | USA | V | 2004 | 153 | SRR2981615 |
| SS80 | invasive | USA | V | 2004 | 1 | SRR2981616 |
| SS81 | invasive | USA | V | 2004 | 1 | SRR2981617 |
| SS82 | invasive | USA | V | 2004 | 1 | SRR2981618 |
| SS83 | invasive | USA | V | 2004 | 1 | SRR2981619 |
| SS84 | invasive | USA | V | 2005 | 1 | SRR2981620 |
| SS85 | invasive | USA | V | 2005 | 1 | SRR2981621 |
| SS86 | invasive | USA | V | 2005 | 1 | SRR2981622 |
| SS87 | invasive | USA | V | 2005 | 1 | SRR2981623 |
| SS88 | invasive | USA | V | 2005 | 1 | SRR2981624 |
| SS89 | invasive | USA | V | 2005 | 1 | SRR2981625 |
| SS92 | invasive | USA | V | 2005 | 1 | SRR2981626 |
| SS93 | invasive | USA | V | 2005 | 1 | SRR2981627 |

STs in bold presenting novel STs identified in this study

### 6.3.2 Clonal complex assignment and core genome phylogeny

Five CCs (CC1, CC8/CC10, CC19, CC17 and CC23) comprising of 47 different STs were identified through geoBurst and core genome phylogeny (**Table 6.2 and Figure 6.2B**). US and Canadian isolates belonged to CC1 and CC17 (as were chosen purposely) whereas UK and Brazilian isolates had varied STs distributed across all five CCs. The majority of CC1 isolates were serotype V/ST1 (179/204; 86.4%) mostly from the US (n=90) and followed by Canada (n=65), UK (n=16) and Brazil (n=6). A majority of CC17 isolates were III/ST17 (85/96; 88.5%) and consisted of Canadian (n=59) and UK (n=26) strains whereas CC19 strains were mostly type III/ST19 (14/32; 43.7%) and comprised of UK GBS strains only. The third largest CC: CC23 mainly comprised of serotype Ia/ST23 (46/52; 88.4%) isolates from Brazil (n=6) and UK (n=40) while CC8/CC10 (Brazil isolates (n=2) and UK (n=31)) were composed of ST12 (n=14) and ST8 (n=10) strains but represented by two serotypes (II and Ib).

**Table 6. 2. Clonal complexes identified in 417/447 invasive GBS dataset from different countries.** From left to right column present the information of isolate sequence type (ST) and the country from which isolates were originated.

| CC | ST | Brazil | USA | Canada | UK | Total |
|---|---|---|---|---|---|---|
| **CC1 (n=204)** | 1 | 6 | 90 | 65 | 22 | 185 |
| | 2 | 0 | 0 | 1 | 1 | 2 |
| | 14 | 0 | 0 | 0 | 1 | 1 |
| | 136 | 0 | 0 | 0 | 1 | 1 |
| | 153 | 0 | 1 | 0 | 0 | 1 |
| | 196 | 0 | 0 | 0 | 3 | 3 |
| | 297 | 0 | 1 | 0 | 1 | 2 |
| | 453 | 0 | 0 | 1 | 0 | 1 |
| | 459 | 0 | 0 | 0 | 1 | 1 |
| | 531 | 0 | 0 | 1 | 0 | 1 |
| | 817 | 1 | 0 | 0 | 0 | 1 |
| | 871 | 0 | 0 | 1 | 0 | 1 |
| | 872 | 0 | 1 | 0 | 0 | 1 |
| | 873 | 0 | 1 | 0 | 0 | 1 |
| | 1217 | 0 | 0 | 0 | 1 | 1 |
| | 1314 | 0 | 0 | 0 | 1 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 1349 | 2 | 0 | 0 | 0 | 2 |
| | 1350 | 0 | 0 | 0 | 1 | 1 |
| **CC17 (n=96)** | 17 | 0 | 0 | 59 | 26 | 85 |
| | 95 | 0 | 0 | 1 | 0 | 1 |
| | 148 | 0 | 0 | 2 | 0 | 2 |
| | 290 | 0 | 0 | 1 | 0 | 1 |
| | 484 | 0 | 0 | 1 | 0 | 1 |
| | 550 | 0 | 0 | 0 | 2 | 2 |
| | 874 | 0 | 0 | 1 | 0 | 1 |
| | 1219 | 0 | 0 | 0 | 1 | 1 |
| | 1221 | 0 | 0 | 0 | 1 | 1 |
| | 1351 | 0 | 0 | 0 | 1 | 1 |
| **CC8/CC10 (n=33)** | 8 | 1 | 0 | 0 | 9 | 10 |
| | 9 | 0 | 0 | 0 | 2 | 2 |
| | 10 | 0 | 0 | 0 | 2 | 2 |
| | 12 | 1 | 0 | 0 | 13 | 14 |
| | 15 | 0 | 0 | 0 | 1 | 1 |
| | 104 | 0 | 0 | 0 | 2 | 2 |
| | 652 | 0 | 0 | 0 | 1 | 1 |
| | 1220 | 0 | 0 | 0 | 1 | 1 |
| **CC19 (n=32)** | 19 | 0 | 0 | 0 | 20 | 20 |
| | 27 | 0 | 0 | 0 | 1 | 1 |
| | 28 | 0 | 0 | 0 | 8 | 8 |
| | 110 | 0 | 0 | 0 | 1 | 1 |
| | 861 | 0 | 0 | 0 | 1 | 1 |
| | 1316 | 0 | 0 | 0 | 1 | 1 |
| **CC23 (n=52)** | 23 | 6 | 0 | 0 | 40 | 46 |
| | 144 | 1 | 0 | 0 | 2 | 3 |
| | 1065 | 0 | 0 | 0 | 1 | 1 |
| | 1218 | 0 | 0 | 0 | 1 | 1 |
| | 1317 | 0 | 0 | 0 | 1 | 1 |

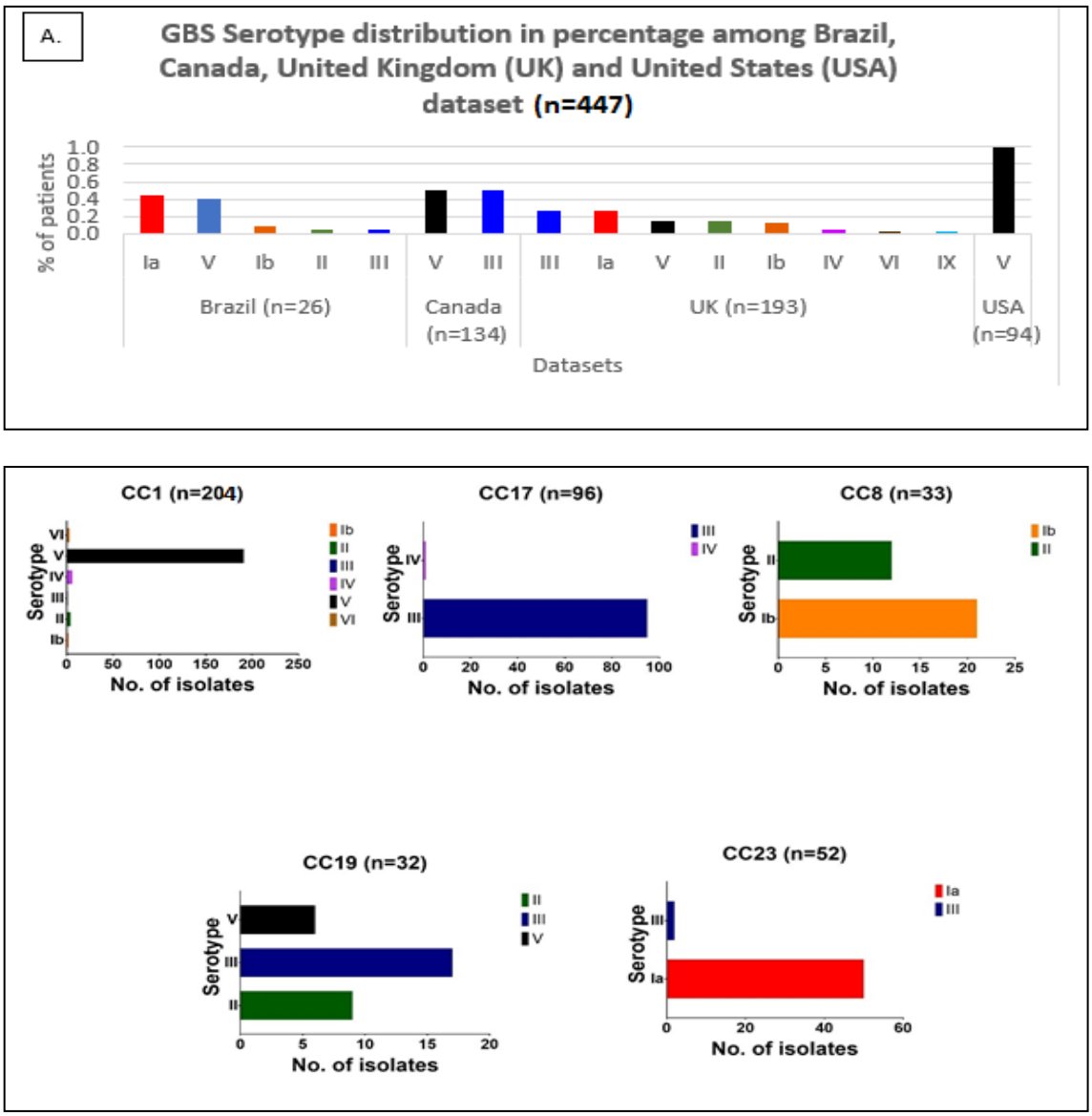



**Figure 6. 2. Serotype distribution from Brazilian, Canadian, the United States and UK datasets and its prevalence among different CCs. A)** Serotypes observed in Brazilian, Canadian, United States and UK datasets. Different colours are used to distinguish different GBS serotypes where Ia is presented by red colour, Ib – orange, II – green, III – blue, IV – magenta, V – black, VI – brown and IX - cyan. **B)** Serotypes observed in five main CCs (CC1, CC17, CC8/CC10, CC19 and CC23) isolates. Different colours are used to present serotypes; Ia -red; Ib - orange; II - green; III - blue; IV - purple; V - black and VI - brown.

#### 6.3.2.1 CC1:

##### 6.3.2.1.1 Surface protein, Pilus Island genes and ARGs:

Ninety-nine percent (201/204) of CC1 isolates had the *srr*-1 gene, except serotype III ST2 (NGBS003) and serotype V ST1 (NGBS177) Canadian strains, while the *srr*-2 gene was absent in 99.1% (202/204) of CC1 strains (**Figure 6.3 and Table 6.3**). Alpha like protein (*alp*1) encoding gene was present in majority of CC1 strains (195/204, 95.5%), of which 191/195 (97.9%) were serotype V; however, its absence was seen in non-serotype V isolates (n=12) only, whereas *alp*2 was present in 10 strains only (including a single UK strain (100414) that was serotype V) (**Figure 6.3 and Table 6.3**). The *bca* gene was present in 4/204 (1.9%) of CC1 strains only, including three UK GBS strains (100414 - serotype V, PHEGBS0446 - serotype VI and PHEGBS0662 - serotype VI) and one Brazil strain (BS-S3300- serotype II). The less common *rib* and *hvg*A genes were observed in 4/204 (1.9%,) strains, including three serotype V CC1 strains (PHEGBS0467, SS116 and SS136) and 3/204 (1.4%,) strains including two serotype V strains (SS116 and SS136), respectively (**Figure 6.3**). In short, *srr*-1 and *alp*1 genes were both found significantly associated with serotype V (190/191, 99.4%) CC1 isolates. The combination of PI-1 and PI-2a predominated among CC1 strains (202/204, 99%), however PI-2b was found in 4/204 (1.9%) of CC1 strains only, including two serotype V strains (SS116 and SS136) from the United States (**Figure 6.3 and Table 6.3**).

Macrolide resistance-mediating genes was observed in 114/204 (55.8%) of CC1 strains dominated by the presence of the *erm*B (74/114, 64.9%) macrolide resistance gene followed by *erm*A (38/114, 33.8%), *mef*A/*msr*D (4/114, 3.5%), and *lsa*C (1/114, 0.8%) (**Figure 6.3 and Table 6.3**). The *tet*M gene was most common in tetracycline resistant CC1 GBS strains (186/188, 98.9%) (including the only chloramphenicol resistant CC1 strain (NGBS273) harbouring *tet*M, *tet*L and *cat*Q genes in combination), followed by *tet*O and *tet*S genes found in one CC1 strain each (**Figure 6.3 and Table 6.3**). Four (1.9%) CC1 strains were positive for aminoglycoside resistance genes including PHEGB50662 and SS120 harbouring a combination

of *ant*(6-Ia) and *aph*(3'-III) genes and two strains (NGBS222 and SS48) carrying *aph*(3'-III) gene alone (**Figure 6.3 and Table 6.3**).

#### 6.3.2.1.2 Phylogenetic analysis:

Among 204 CC1 isolates, the most common strains observed were serotype V ST1 strains (179/204, 87.7%) including 7 from Brazil, 17 from the UK, 65 from Canada and 90 from the United States GBS; followed by serotype II (n=2), Ib (n=2) and VI (n=2) isolates from the UK (**Figure 6.3**). Four main clusters and two outliers (NGBS003 and PHEGBS0511) were observed in CC1 phylogenetic tree; cluster 1, with an average SNP difference of 286 (range: 64 - 12,206) was highly homogenous and consist mainly of serotype V, II and Ib strains from all four regions; cluster 2, with an average of 4,641 SNP difference (range: 4494 - 4753): UK serotype IV ST196, ST136 and ST459 strains (PHEGBS0266, PHEGBS0084, PHEGBS0206, PHEGBS0448 and PHEGBS0463); cluster 3, with an average of 4,380 SNPs difference: serotype II strains from UK and Brazil (PHEGBS0627 and BS-S3300); and cluster 4, with an average of 6,888 SNPs difference (range: 6,544 – 7,481): serotype VI UK strains (PHEGBS0533, PHEGBS0446 and PHEGBS0662) (**Figure 6.3**). Cluster 2 strains were all found positive for *alp*2 surface protein gene, however no clustering based on antibiotic resistance gene profiles was noted.

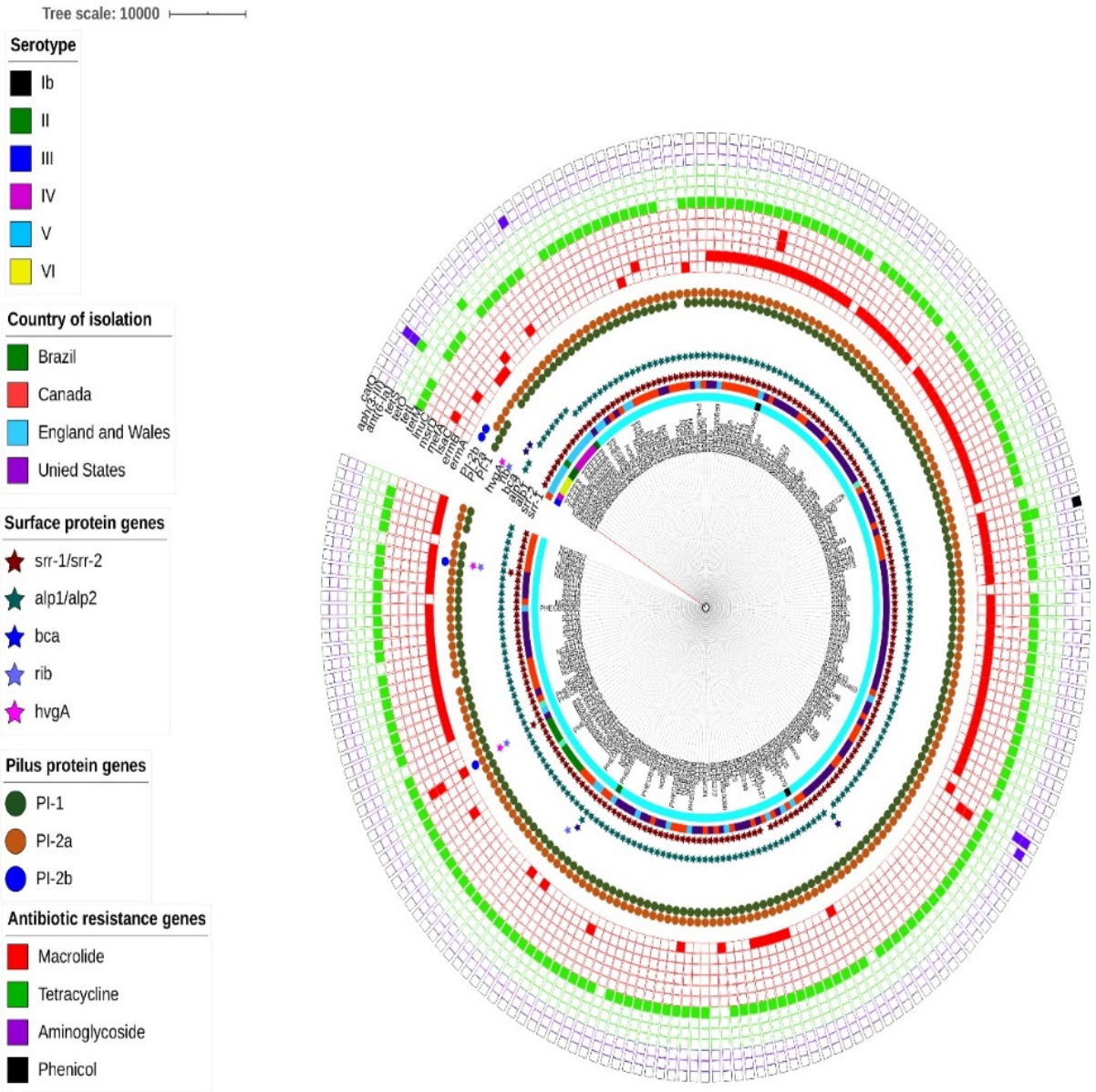



**Figure 6. 3. Phylogenetic relationship of GBS strain population from Brazil, Canada, United States and UK clustered in CC1**. RAxML v8.2.8 [508] was used to generate mid-point rooted maximum likelihood phylogeny tree for 204 isolates clustered in CC1 against reference sequence SS1 (NZ_CP010867) using SNPs based approach (core SNPs: 25,798). Inner to outermost ring represents phylogeny tree of CC1 isolates with a unique identifier, capsular serotype, country of isolation, virulence genes, pilus island genes and antibiotic resistance genes. Different shapes and colours are used to highlight important features i.e. coloured star is used to show virulence genes and *srr*-1/ *srr*-2 are coloured maroon, *alp*1/*alp*2 – green, *bca* – blue, *rib* – light purple, and *hvg*A – magenta; coloured circle was used to show pilus gene positive GBS strains where green circle present PI-1 positive strain, brown – PI-2a positive strain and blue – PI-2b positive strain; while presence of tetracycline resistance genes (*tet*M, *tet*S, or *tet*O; light green), macrolide resistance genes(*erm*A, *erm*B, *mef*A/*msr*D or *lsa*C; red) or aminoglycoside resistance genes (*ant*(6-Ia) or *aph*(3'-III); purple) for each isolate is shown by a coloured square. Scale bar, represents a distance of 10000 point mutations.

#### 6.3.2.2 CC17:

##### 6.3.2.2.1 Surface protein, Pilus Island genes and ARGs:

The *srr*-2, *rib* and *hvg*A genes were found in all CC17 isolates except 11 Canadian serotype III strains (NGBS082, NGBS126, NGBS128, NGBS186, NGBS220, NGBS239, NGBS250, NGBS377, NGBS431, NGBS464 and NGBS485) that lacked the *srr*-2 gene, two Canadian serotype III strains (NGBS239, NGBS250) that lacked the *rib* gene and 4 Canadian strains (NGBS082, NGBS126, NGBS239 and NGBS250) that lacked the *hvg*A gene (**Figure 6.4 and Table 6.3).** None of the CC17 isolates were positive for the *srr*-1, *alp*1, *alp*2 or *bca* genes. The combination of PI-1 and PI-2b genes were found common in the majority of CC17 isolates (78/96, 81.2%) while none of them harboured PI-2a gene (**Figure 6.4 and Table 6.3**). Thirty-four (35.5%) CC17 isolates (5 UK and 29 Canadian strains) were found to carry macrolide resistance genes (*erm*A, *erm*B, *mef*A/*msr*D, and/or *lnu*B), the most common of which was *erm*T gene found in (19/34, 55.8%) GBS strains, followed by *erm*B (14/34, 41.1%), *mef*A/*msr*D (3/34, 8.8) and *lnu*B (2/34, 5.8%) genes (**Figure 6.4 and Table 6.3**). Tetracycline resistance-mediating genes were detected in (95/96, 98.9%) CC17 isolates except a single strain (NGBS583) that carried no antibiotic resistance gene at all. The most prevalent tetracycline resistance gene was the *tetM gene* (84/95, 88.4%) followed by *tet*O found in 15/95 (15.8%) of the CC17 isolates (**Figure 6.4 and Table 6.3**). Whereas aminoglycoside resistant GBS strains had a comparatively higher prevalence in CC17 (14/96, 14.5%) than CC1 (1.9%) due to the presence of combined *ant*(6-Ia) and *aph*(3'-III) genes (**Figure 6.4 and Table 6.3**).

##### 6.3.2.2.2 Phylogenetic analysis:

CC17 consist of 31 UK and 65 Canadian GBS strains, the majority of which were serotype III and ST17 (85/96, 88.5%), while the others were serotype III (either SLV or DLV of ST17), with only one CC17 isolate found as serotype IV in CC17 (**Figure 6.4**). Four clusters were observed in the CC17 phylogeny; cluster 1: red, composed of highly homogenous serotype III and IV strains and had 238 SNP differences on average ranged between 143 and 1,169; cluster 2: blue, with an average SNP difference of 153

(range: 89-296) and cluster 3: green, with an average of SNP difference of 588 (range between 517-710), both of cluster 2 and 3 composed of UK and Canadian serotype III strains. While cluster 4: purple, consisted of only Canadian serotype III strains with an average SNP difference of 181 (range between 100 and 223) (**Figure 6.4**). No clustering based on antibiotic resistance gene profiles was noted.

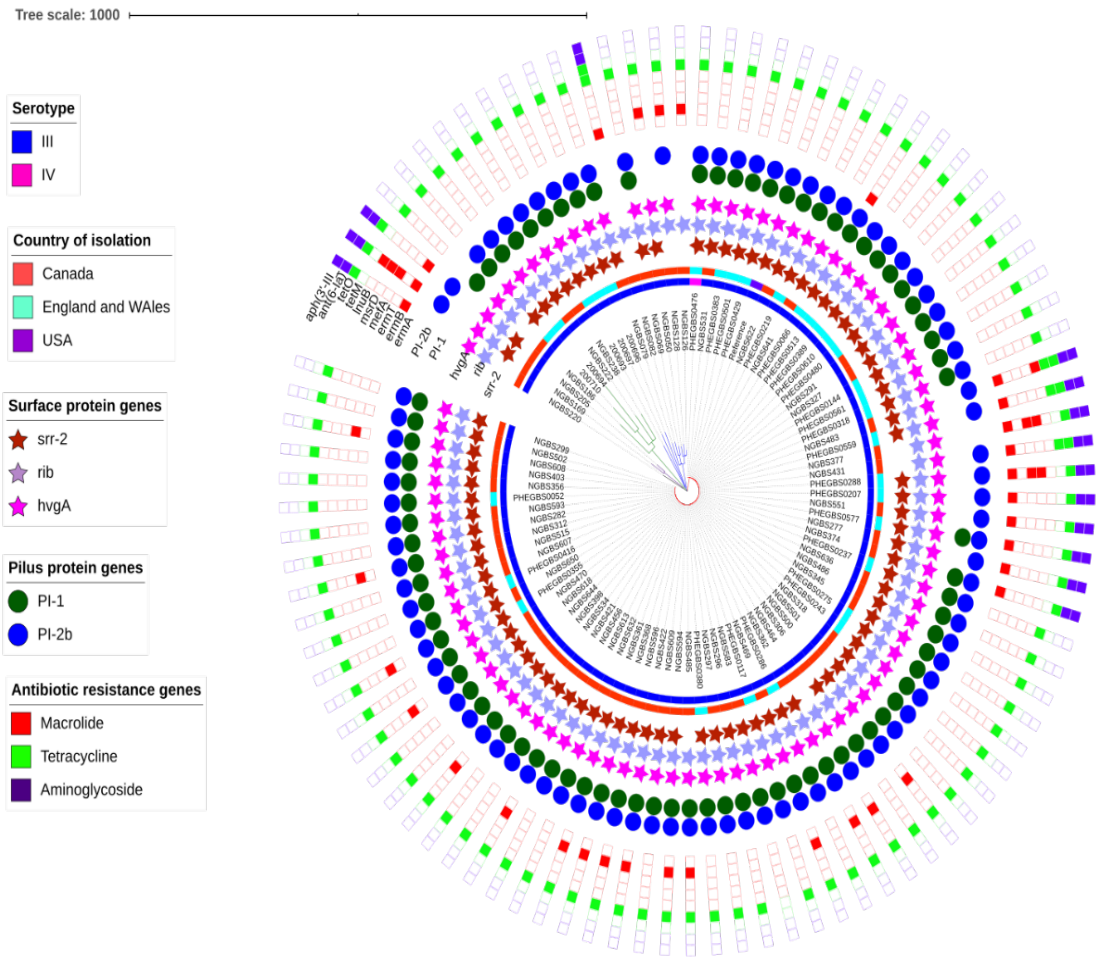



**Figure 6. 4. Phylogenetic relationship of GBS strain population from Canada, United States and UK clustered in CC17**. RAxML v8.2.8 [508] was used to generate mid-point rooted maximum likelihood phylogeny tree for 96 isolates clustered in CC1 against reference sequence COH1 (NZ_ HG939456) using SNPs based approach (core SNPs: 23,477). Inner to outermost ring represents phylogeny tree of CC17 isolates with a unique identifier, capsular serotype, country of isolation, virulence genes, pilus island genes and antibiotic resistance genes. Different shapes and colours are used to highlight important features i.e. coloured star is used to show virulence genes and *srr*-2 is coloured maroon, *rib* – light purple, and *hvgA* – magenta; coloured circle was used to show pilus gene positive GBS strains where green circle present PI-1 positive strain and blue – PI-2b positive strain; while presence of tetracycline resistance genes (*tet*M or *tet*O; light green), macrolide resistance genes (*erm*A, *erm*B, *erm*T, *mef*A/*msr*D or *lnu*B; red) or aminoglycoside resistance genes (*ant*(6-Ia) or *aph*(3'-III); purple) for each isolate is shown by a coloured square. Scale bar, represents a distance of 10000 point mutations.

#### 6.3.2.3 CC8/CC10:

##### 6.3.2.3.1 Surface protein, Pilus Island genes and ARGs:

All CC8/CC10 isolates (32/33, 96.9%) had *srr*-1 and *bca* genes except PHEGBS0575, which was negative for these two genes (**Figure 6.5 and Table 6.3**). However, PHEGBS0575 serotype Ib ST1220 was the only CC8/CC10 isolate positive for *srr*-2, *rib* and *hvgA* genes, which were absent in all the other (32/33, 96.9%) CC8/CC10 isolates (**Figure 6.5 and Table 6.3**). A combination of PI-1 and PI-2a genes were found prevalent in 25/33 (75.8%) of CC8/CC10 isolates, followed by PI-2b alone observed in (6/33, 18.1%) CC8/CC10 isolates (**Figure 6.5**). Seven out of 33 CC8/CC10 isolates (31.2%) were found positive for macrolide resistance genes, the majority of which harboured *erm*A (5/7, 71.4%) and *erm*B (2/7, 28.5%) genes (**Figure 6.5 and Table 6.3**). Tetracycline resistance genes were observed in 25/33 (75.8%) CC8/CC10 strains, consisting of *tet*M (14/25, 56%), followed by *tet*O (12/25, 48%) and *tet*L (1/25, 4%) genes. None of the CC8/CC10 isolates were positive for aminoglycoside and phenicol resistance genes (**Figure 6.5 and Table 6.3**).

##### 6.3.2.3.2 Phylogenetic analysis:

Two Brazil GBS serotype Ib isolates along with 31 UK GBS isolates, defined by two serotypes Ib and II made up CC8/CC10 (**Figure 6.5**). Two clusters were formed on the phylogeny tree, including an outlier isolate (PHEGBS0575) demonstrating significant unrelatedness of this isolate from the other CC8/CC10 strains; cluster 1: green, comprised of 3 serotype Ib and II UK isolates (PHEGBS0072, PHEGBS0153 and PHEGBS0097) with an average SNP difference of 6,930 (range: 6,497 – 7,536) and cluster 2: red, consisting of serotype II and Ib Brazil and UK GBS isolates with an average SNP difference of 1,918 that ranged between 636 and 10,838 SNPs (**Figure 6.5**). Both clusters were found positive for *srr*-1 and *bca* genes however no clustering based on antibiotic resistance gene profiles was noted.

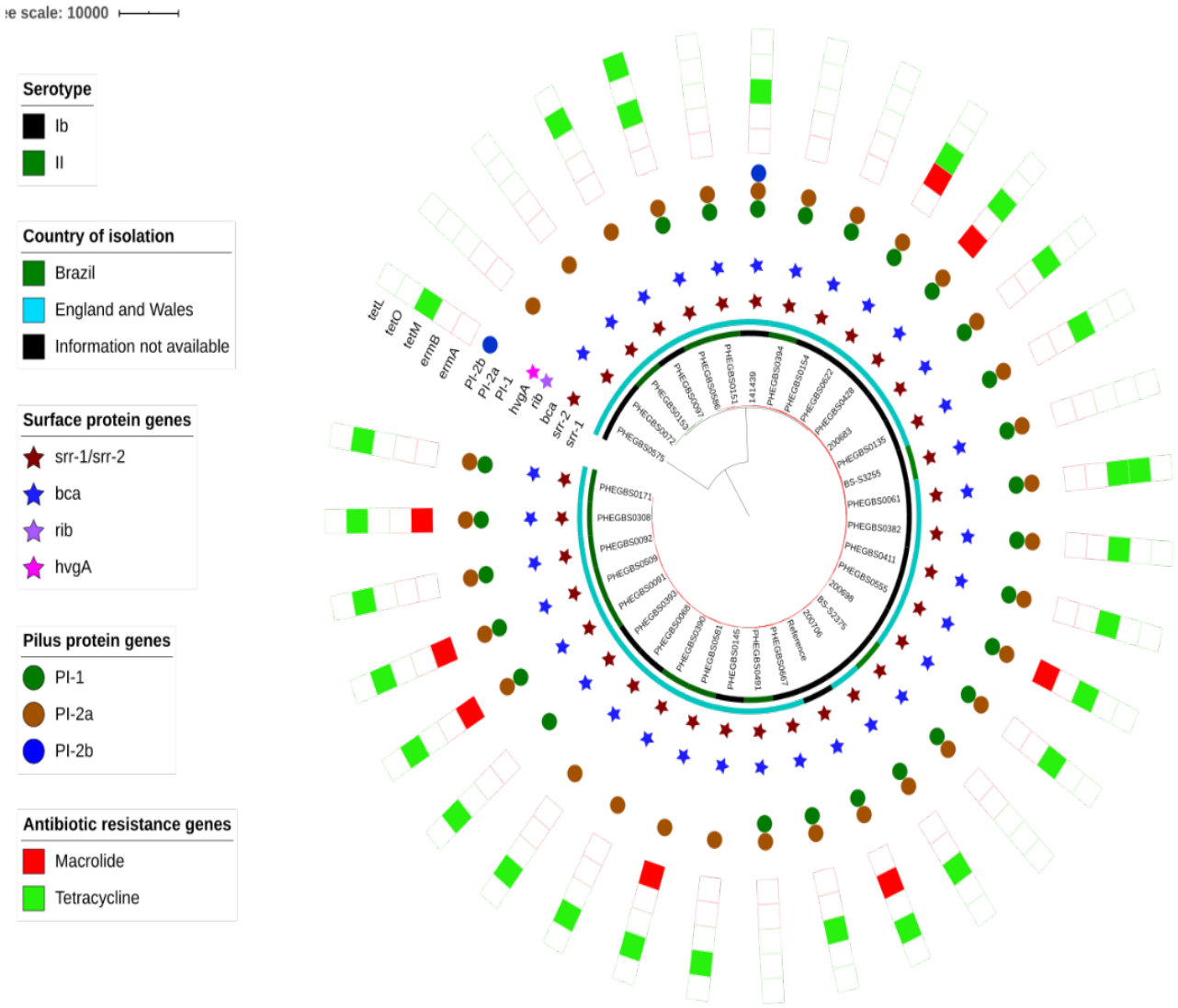



**Figure 6. 5. Phylogenetic relationship of GBS strain population from Brazil and UK clustered in CC8/CC10.** RAxML v8.2.8 [508] was used to generate mid-point rooted maximum likelihood phylogeny tree for 33 isolates clustered in CC8/CC10 against reference sequence Sag37 (NZ_CP019978) using SNPs based approach (core SNPs: 78,332). Inner to outermost ring represents phylogeny tree of CC8 isolates with a unique identifier, capsular serotype, country of isolation, virulence genes, pilus island genes and antibiotic resistance genes. Different shapes and colours are used to highlight important features i.e. coloured star is used to show virulence genes and *srr*-1/*srr*-2 is coloured maroon, *bca* – blue, *rib* – light purple, and *hvgA* – magenta; coloured circle was used to show pilus gene positive GBS strains where green circle present PI-1 positive strain, brown – PI-2a positive strain and blue – PI-2b positive strain; while presence of tetracycline resistance genes (*tet*M, *tet*Lor *tet*O; light green), macrolide resistance genes (*erm*A and *erm*B; red) for each isolate is shown by a coloured square. Scale bar, represents a distance of 10000 point mutations.

#### 6.3.2.4 CC19:

##### 6.3.2.4.1 Surface protein, Pilus Island genes and ARGs:

The *srr*-1 and/or *rib* genes were found in 28/32 (87.5%) of CC19 strains while only a small fraction of CC19 strains (6/32,18.7%), including five serotype V/ST19 isolates (PHEGBS0139, PHEGBS0595, PHEGBS0608, PHEGBS0657 and PHEGBS0738) and a single serotype III/ST1316 harbouring *srr*-1 also had the presence of *alp*1 genes (6/32, 18.75%) (**Figure 6.6 and Table 6.3**). None of the CC19 isolates had the *srr*-2, *alp*2, *bca* or *hvg*A genes. The combination of PI-1 and PI-2a was found most prevalent in 31/32 (96.8%) of CC19 strains, except a single serotype II/ST28 strain (PHEGBS0624), which was found negative for the PI-1 gene and harboured the PI-2a gene alone (**Figure 6.6 and Table 6.3**). CC19 was the only lineage that carried all chloramphenicol resistant GBS strains (4/449, 0.8%) PHEGBS0595, PHEGBS0608 and PHEGBS0598 from the analysed dataset carrying *cat*Q (3/4, 75%) and the single *cat*(pC194) (1/4, 25%) positive strain (**Figure 6.6 and Table 6.3**). Macrolide resistance genes were observed in 12/32, (37.5%) of the CC19 GBS strains due to the presence of *erm*B (5/12, 41.6%), *lsa*C (4/12, 33.3%), *erm*A and/or *mef*A/*msr*D (4/12, 33.3%) and *lnu*C (1/12, 8.3%) (**Table 6.3**). Tetracycline resistance was observed in (25/32, 78.1%) CC19 strain, *tet*M was the most frequent tetracycline resistance determinants were observed in 23/25 (92%) strains, including two strains harbouring *tet*O as well as *tet*M in addition two strains carrying *tet*O alone (8%) (**Figure 6.6 and Table 6.3**). Overall, a low number of CC19 strains (3/32, 9.3%) were positive for aminoglycoside resistance genes including two strains (PHEGBS0156 and PHEGBS0738) harbouring *ant*(6-Ia) and *aph*(3'-III) in combination and a single strain (PHEGBS0139) carrying the high-level gentamicin resistance gene - *aac*(6')-*aph*(2") (**Figure 6.6 and Table 6.3**).

##### 6.3.2.4.2 Phylogenetic analysis:

Thirty-two GBS strains from the UK only were clustered in CC19 with almost half (14/32, 43.7%) the strains identified as serotype III ST19, followed by 8/32 (25%) serotype II ST8 strains (**Figure 6.6**). Three main clusters were observed on phylogeny tree with an outlier, serotype III/ST23 strain -

PHEGBS0377, which carried no ARGs. Cluster 1: red, was made up of 17 serotype III and V ST19, ST110, ST861 and ST110 strains (PHEGBS0639, PHEGBS0492, PHEGBS0598, PHEGBS0252, PHEGBS0156, PHEGBS0566, PHEGBS0567, PHEGBS0071, 200699, PHEGBS0359, PHEGBS0372, PHEGBS0657, PHEGBS0372, PHEGBS0608, PHEGBS0738, PHEGBS0595 and PHEGBS0139), 10/17 (58.8%) of which carried multiple drug resistance genes to macrolides, aminoglycosides and/or chloramphenicols (**Figure 6.6**). Cluster 2: blue, consisting of six serotype III ST19 or ST27 isolates (200682, PHEGBS0664, PHEGBS0296, PHEGBS0626, PHEGBS0593 and PHEGBS0643) only one (200682) of which carried macrolide and tetracycline resistance genes in combination, while the rest of them were tetracycline resistant gene positive only (**Figure 6.6**). Cluster 3: green, comprised of eight serotype II ST28 isolates (PHEGBS0532, PHEGBS0544, PHEGBS0408, PHEGBS0464, PHEGBS0624, PHEGBS0399, PHEGBS0551 and PHEGBS0556) including two isolates (PHEGBS0551 and PHEGBS0624) that harboured macrolide and tetracycline resistance determinants, while the rest carried tetracycline resistance genes only (**Figure 6.6**). The average SNP differences found in each cluster against reference genome was Cluster 1: 2,003 (range between 539 and 4,135), cluster 2: 2,297 (range between 1,688 and 3,712) and cluster 3: 3,156 (range between 1,740 and 5,528).

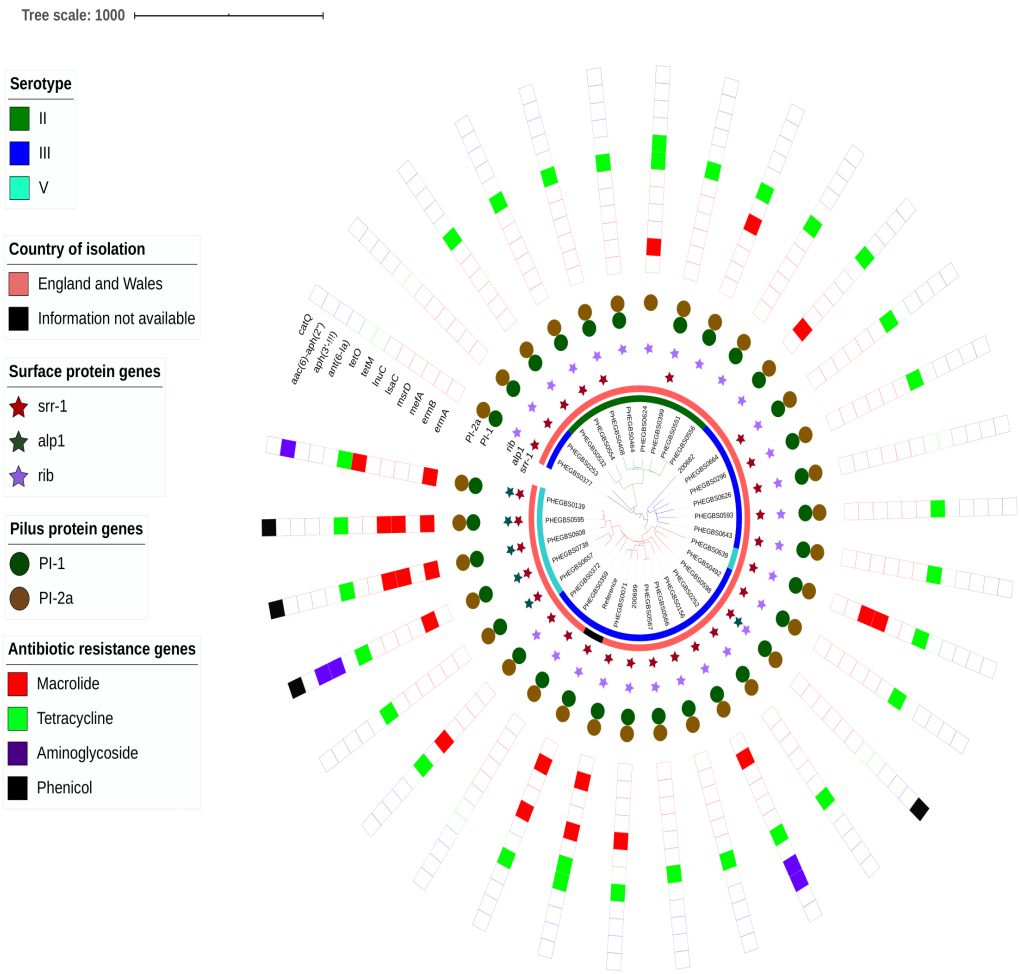



**Figure 6. 6. Phylogenetic relationship of GBS strain population from UK clustered in CC19.** RAxML v8.2.8 [508] was used to generate mid-point rooted maximum likelihood phylogeny tree for 32 isolates clustered in CC19 against reference sequence SG-M25 (NZ_ CP021867) using SNPs based approach (core SNPs: 75,103). Inner to outermost ring represents phylogeny tree of CC19 isolates with a unique identifier, capsular serotype, country of isolation, virulence genes, pilus island genes and antibiotic resistance genes. Different shapes and colours are used to highlight important features i.e. coloured star is used to show virulence genes and *srr*-1 is coloured maroon, *alp*1 – green and *rib* – light purple; coloured circle was used to show pilus gene positive GBS strains where green circle present PI-1 positive strain and brown – PI-2a positive strain; while presence of tetracycline resistance genes (*tet*M or *tet*O; light green), macrolide resistance genes (*erm*A, *erm*B, *mef*A/*msr*D, *lsa*C and *lnu*C; red), aminoglycoside resistance genes (*ant*(6-Ia) or *aph*(3'-III); purple) and chloramphenicol resistance genes (*cat*Q: black) for each isolate is shown by a coloured square. Scale bar, represents a distance of 1000 point mutations.

#### 6.3.2.5 CC23:

##### 6.3.2.5.1 Surface protein, Pilus Island genes and ARGs:

The most common virulence gene found in CC23 isolates was *srr*-1 (51/52, 98%), followed by *alp*1 observed in 46/52 (88.4%) CC23 strains (Figure 6.6). Two strains, (PHEGBS0176 and PHEGBS0524) which were both ST23 serotype III, were found positive for *alp*2 and PI-1 genes while all other CC23 isolates (50/52, 96%) carried the PI-2a gene instead (**Figure 6.7 and Table 6.3**). Interestingly, the *rib* gene was found in three ST144 (a SLV of ST23) serotype Ia isolates (BS-S3201, PHEGBS0044 and PHEGBS0663) only, which is otherwise a characteristic gene for serotype III isolates belonging to either CC17 or CC19 lineages. Macrolide resistance genes were found in 7/52 (13.4%) of CC23 strains (**Table 6.3**), including *mef*A/*msr*D genes found in 5/7 (71.4%) CC23 strains (BS-S3201, PHEGBS0067, PHEGBS0070, PHEGBS0248 and PHEGBS0625), and strains carrying either *erm*A or *erm*T genes observed in PHEGBS0044 and PHEGBS0539 strains, respectively (**Figure 6.7**). About 50/52 (96.1%) of CC23 strains were resistant to tetracycline due to the presence of *tet*M gene alone (50/52, 96.1%) including seven strains that carried macrolide and tetracycline resistance genes in combination (**Figure 6.7 and Table 6.3**). Only 2/52 (3.8%) of CC23 strains (PHEGBS0176 and PHEGBS0193) carried no ARGs at all, while none of the isolates carried aminoglycoside or chloramphenicol resistance genes.

##### 6.3.2.5.2 Phylogenetic analysis:

Most of the CC23 isolates were ST23 serotype Ia (44/52, 84.6%) along with two ST23 serotype III (3.8%) strains (**Figure 6.7**). Three main clusters were found on CC23 phylogeny; cluster 1: red, composed of 45 serotype Ia (ST23, ST1065 and ST1218)- 5/45(11.1%) from the Brazil dataset and 40/45 (88.8%) from the UK strains and had 4,361 SNPS difference on average from the reference, cluster 2: blue, composed of 5 serotype Ia (ST144, ST23 and ST1317) strains, including two strains from Brazil (BS-S3201 and BS-S2318) and the remaining three (PHEGBS0049, PHEGBS0663, and PHEGBS0086) from the UK, with an average SNP difference of 4,510 SNP from the reference (**Figure 6.7**). While cluster 3: purple, is composed of two UK strains (PHEGBS0176 and PHEGBS0524) and

the reference NEM316 strain, with an average of 1,582 SNP difference from the reference (**Figure 6.7**).

No clustering based on antibiotic resistance gene profiles was noted.

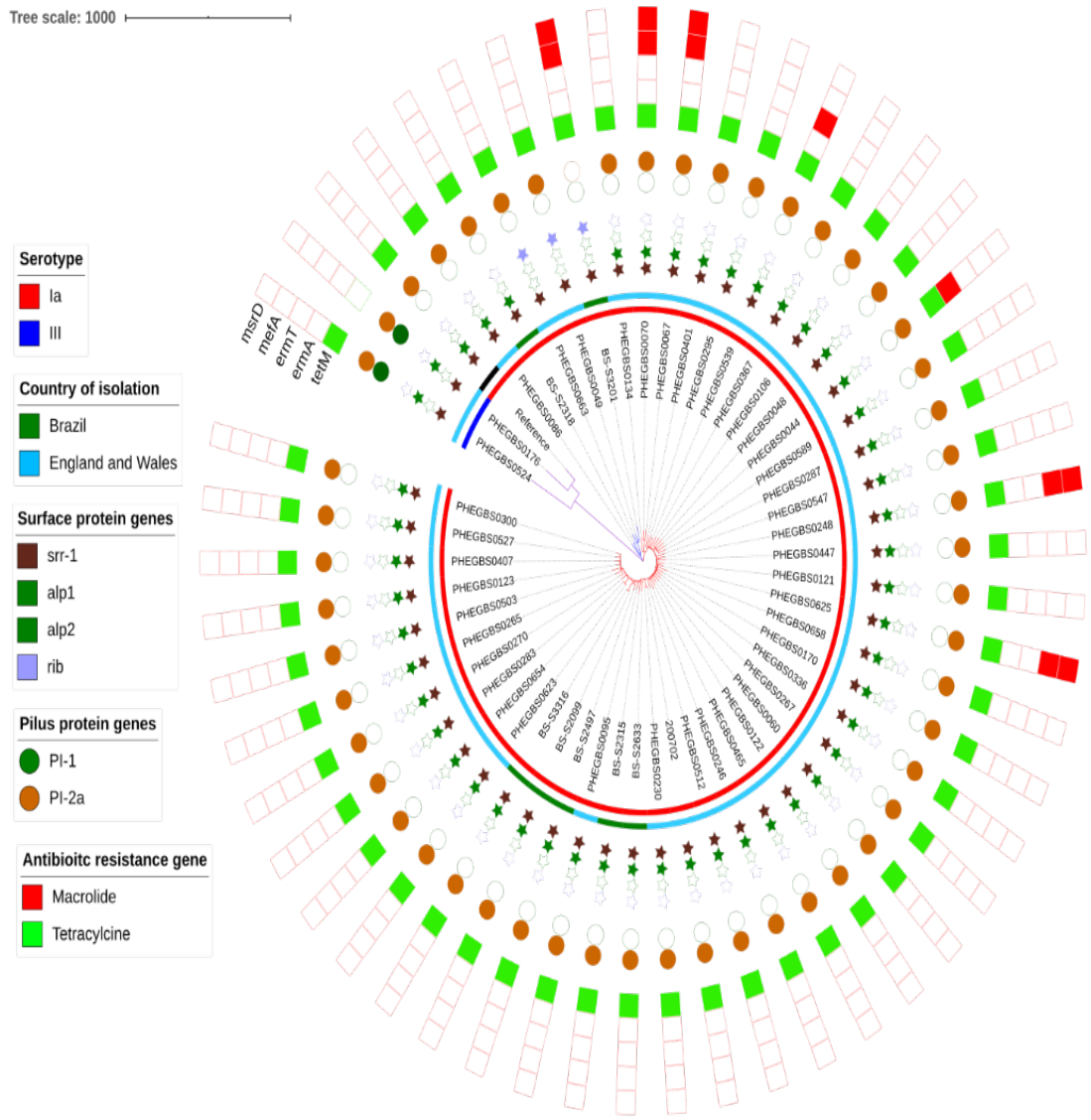



**Figure 6. 7. Phylogenetic relationship of GBS strain population from Brazil and UK clustered in CC23.** RAxML v8.2.8 [508] was used to generate mid-point rooted maximum likelihood phylogeny tree for 52 isolates clustered in CC23 against reference sequence NEM316 (NC_004368) using SNPs based approach (core SNPs: 2,21,983). Inner to outermost ring represents phylogeny tree of CC23 isolates with a unique identifier, capsular serotype, country of isolation, virulence genes, pilus island genes and antibiotic resistance genes. Different shapes and colours are used to highlight important features i.e. coloured star is used to show virulence genes and *srr*-1 is coloured maroon, *alp*1/ *alp*2 – green and *rib* – light purple; coloured circle was used to show pilus gene positive GBS strains where green circle present PI-1 positive strain and brown – PI-2a positive strain; while presence of tetracycline resistance genes (*tet*M; light green), macrolide resistance genes (*erm*A, *erm*t, *mef*A/*msr*D; red) for each isolate is shown by a coloured square. Scale bar, represents a distance of 1000 point mutations.

**Table 6. 3. Prevalence of antibiotic resistance genes, virulence genes and pilus island genes determined in five major CCs.**

| CC | Antibiotic and its resistance determinants | | | | | | | | Virulence genes | | | | | | | Pilus island genes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Tetracycline (tet)** | **Tet resistance determinants** | **Macrolide (mac)** | **Mac resistance determinants** | **Aminoglycoside (amino)** | **Amino resistance determinants** | **Phenicol (phe)** | **phe esistance determinants** | ***srr*-1** | ***srr*-2** | ***alp*1** | ***alp* 2** | ***bca*** | ***rib*** | ***hvg* A** | **PI-1** | **PI-2a** | **PI-2b** |
| CC1 (n=204) | 90.8% (n=188) | *tet*M = 186<br>*tet*O=1<br>*tet*L=1<br>*tet*S=1 | 55% (n=114) | *erm*A = 38<br>*erm*B = 74<br>*mef*A/*msr*D = *40*<br>*lsa*C = 1 | 1.9% (n=4) | *aph*(3'-III) = 4<br>*ant*(6-Ia) = 2 | 0.4% (n=1) | *cat*Q=1 | 99% (n= 205) | 0.9 % (n= 2) | 94.2 % (n= 195) | 4% (n= 10) | 1.9 % (n= 4) | 1.9 % (n= 4) | 1.4 % (n= 3) | 98% (n= 203) | 98% (n= 203) | 1.9 % (n= 4) |
| CC17 (n=96) | 98.9% (n=95) | *tet*M = 84<br>*tet*O=15 | 35.5 % (n=34) | *erm*T=19<br>*erm*B=15 | 14.5% (n=14) | *aph*(3'-III) = 14<br>*ant*(6-Ia) = 14 | NIL | NIL | NIL | 88. 5% (n= 85) | NIL | NI L | NI L | 97. 9% (n= 94) | 95. 8% (n= 92) | 82.2 % (n= 79) | NIL | 91. 6% (n= 88) |
| CC8/ CC10 (n=33) | 75.8% (n=25) | *tet*M = 14<br>*tet*O=12<br>*tet*L=1 | 21.2 % (n=7) | *erm*A = 6<br>*erm*B = 2 | NIL | NIL | NIL | NIL | 96.9 % (n= 32) | 3% (n= 1) | NIL | NI L | 96. 9% (n= 32) | 3% (n= 1) | 3% (n= 1) | 75.8 % (n= 25) | 75.8 % (n= 25) | 18. 1% (n= 6) |
| | 7831% (n=25) | *tet*M = 23 | 37.5 % | *erm*B = 5 | 9.3% (n=3) | *aph*(3'-III) = 2 | | *cat*Q=3 | 87.5 % | NI L | 18.7 % | NI L | NI L | 87. 5% | NI L | 96.8 % | 100 % | NI L |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CC19 (n=32) | | *tet*O=4 | (n=12) | *lsa*C = 4<br>*erm*A =4<br>*mef*A/*msrD = 4*<br>*lnu*C=1 | | *ant*(6-Ia) = 2<br>*aac*(6')-*aph*(2") =1 | 3.1% (n=4) | *cat*pc(194)=1 | (n=31) | | (n=6) | | | (n=31) | | (n=31) | (n=32) | |
| CC23 (n=52) | 96.1% (n=50) | *tet*M = 50 | 13.4 % (n=7) | *erm*A = 1<br>*erm*T = 1<br>*mef*A/*msrD = 5* | NIL | NIL | NIL | NIL | 98% (n=51) | NIL | 88.4 % (n=46) | 3.8 % (n=2) | NIL | 5.8 % (n=3) | NIL | 3.1 % (n=2) | 96.1 % (n=50) | NIL |

NIL – means none of the GBS isolates harboured specific gene.

### 6.3.3 Pan-genome and pan-GWAS

A pan-genome was built on a dataset of 417 invasive GBS genomes, representing 5 different clonal complexes (**Figure 6.2B**). According to the Roary results [512], 1,328 were the core genes (i.e. present in 99%-100% of the isolates), 204 were the soft core genes (present in 95% - 99% of the isolates), 910 were the shell genes (present in 15% - 95% of the isolates), and 11,602 were the cloud genes (present in 0% - 15% of the isolates). Overall, 14,044 genes were identified present in 0% - 100% of the isolates. Gene's characteristic of all CCs was classified into four functional groups (**Table 6.4**) including i) metabolism, ii) environmental information processing, iii) Signalling and cellular processes, and iv) genetic information processing.

**Table 6. 4. Functional class and pathways associated with each CC as per KEGG database** [662] **results.**

| CC | Functional class | KEGG reference | Pathway |
|---|---|---|---|
| CC1 | Genetic Information Processing (09120) | 99976 | Replication and repair |
| | | 03020 | RNA polymerase |
| | | 03036 | Chromosome and associated proteins |
| | Metabolism (09130) | 00510 | N-Glycan biosynthesis |
| | | 01100 | Metabolic pathways |
| CC19 | Genetic Information Processing (09120) | 03000 | Transcription factors |
| | | 99976 | Replication and repair |
| | Environmental Information Processing (09130) | 09132 | Signal transduction |
| | | 03070 | Bacterial secretion system |
| CC8/CC10 | Environmental Information Processing (09130) | 02010 | ABC transporter |
| | Genetic Information Processing (09120) | 02048 | Prokaryotic defence system |
| | | 99976 | Replication and repair |
| | Metabolism (09130) | 00010 | Glycolysis / Gluconeogenesis |
| | | 00052 | Galactose metabolism |
| | | 01100 | Metabolic pathways |
| | | 01110 | Biosynthesis of secondary metabolites |
| | | 01120 | Microbial metabolism in diverse environments |

| | | | |
|---|---|---|---|
| | | 00053 | Ascorbate and aldarate metabolism |
| | | 02060 | Phosphotransferase system (PTS) |
| | | 00520 | Amino sugar and nucleotide sugar metabolism |
| | | 04917 | Prolactin signalling pathway" |
| | | 00040 | Pentose and glucuronate interconversions |
| | | 00051 | Fructose and mannose metabolism |
| | | 00541 | O-Antigen nucleotide sugar biosynthesis |
| CC17 | Environmental Information Processing (09130) | 02010 | ABC transporters |
| | | 02020 | Two-component system |
| | | 03070 | Bacterial secretion system |
| | | 02044 | Secretion system |
| | | 02000 | Transporters |
| | | 02024 | Quorum sensing |
| | Genetic Information Processing (09120) | 99976 | DNA Replication and repair protein |
| | | 03060 | Protein export system |
| | | 03420 | Nucleotide excision repair |
| | Metabolism (09130) | 01001 | Peptidoglycan biosynthesis and degradation |
| | | 01002 | Peptidases |
| | | 01001 | Protein kinases |
| | | 01003 | Glycosyltransferases |
| | | 00680 | Methane metabolism |
| | | 03060 | Protein export |
| CC23 | Metabolism (09130) | 01059 | Biosynthesis of enediyne antibiotics |
| | | 01100 | Metabolic pathways |
| | | 01110 | Biosynthesis of secondary metabolites |
| | Genetic Information Processing (09120) | 03000 | Transcription factors |
| | Environmental Information Processing (09130) | 02010 | ABC transporters |
| | | 02000 | Transporters |
| | Signalling and cellular processes (09183) | 02042 | Bacterial toxins |

Scoary analysis presented that a total of 89 genes were found to be CC1 specific (present in all or at least >95% (n=186) of CC1 isolates but absent in all or present in <5% (n=10) isolates of other CCs), this includes 77 hypothetical proteins, and remaining 12 genes involved in metabolism, genetic information processing and environmental information processing

pathways (**Table 6.5**). CC1 mainly comprised of recombinases (*xer*C/*xer*D), translocase (*fts*K) and transposases (IS5 and Tn3 family).

A total of 134 genes were found specific to CC19 including heavy metal resistance genes (cadmium, copper, and arsenic) and type IV secretion system protein (VirD4) – a large complex protein that can cross the cell envelope of many bacteria [663], found limited to CC19 in addition to 118 hypothetical genes (**Table 6.6**). The CC19 associated genes were involved in environmental information and genetic information processes (**Table 6.6**).

A total of 45 genes were found associated with CC8/CC10 including 19 hypothetical protein and other genes fitting into three functional pathway classes (**Table 6.7**). In particular, prokaryotic defence systems, such as the CRISPR-Cas9 associated genes (*cas*1/*cas*2), two component regulatory protein gene (*cus*S; that facilitates $Cu^{2+}$ response regulation), phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) genes (*gat*A, *gat*B, and *gat*C; a major carbohydrate transport system of bacteria [664]), and carbohydrate metabolism and uptake genes (*ara*Q and *yes*O) were found CC8/CC10 specific (**Table 6.7**).

CC17 found to have 155 associated genes including 125 hypothetical protein genes and other genes playing role in signal transduction through *lyt*S secretion systems, through *asp*1/*asp*2 genes, as transporters through *feu*C genes and quorum sensing through *nis*L genes, hence harbouring genes important for signal transduction, virulence, and detection of environmental stimuli (**Table 6.8**).

CC23 specific genes (n=59) included 46 hypothetical protein encoding genes and 13 genes confined to metabolism, signalling and cellular processes, environmental information

processing and genetic information processing (**Table 6.9**). This mainly includes transcription factor and regulator genes (*mnt*R and *ram*B) and bacterial toxins encoding genes (*cyl*E and *cyl*K).

**Table 6. 5. Genes specific to CC1 as identified by Scoary v1.6.16** [513]**.** Each column shows (from left to right) the name and annotation of the gene as identified by Prokka v1.12 [497] number of CC1 strains positive for a specific gene, number of CC1 strains negative for a specific gene, number of other CC isolates with presence or absence of a specific gene , respectively, followed by the KEGG database id entry name and functional class (where available). 271 hypothetical protein encoding genes specific to CC1 were excluded from the table.

| Gene | Annotation | No. of CC1 strain carrying specific gene | No. of CC1 strain devoid of specific gene | No. of other CC strains carrying specific gene | No. of other CC strains devoid of specific gene | Functional class with class reference in brackets | KEGG Pathway reference |
|---|---|---|---|---|---|---|---|
| *fts*K | DNA translocase FtsK | 171 | 36 | 0 | 213 | | |
| *hin* | DNA-invertase hin | 67 | 140 | 3 | 210 | | |
| *gly*G | Glycosyltransferase GlyG | 185 | 22 | 6 | 207 | | |
| *hip*O | Hippurate hydrolase | 203 | 4 | 0 | 213 | | |
| *ama*A | N-acyl-L-amino acid amidohydrolase | 197 | 10 | 0 | 213 | | |
| group_893 | IS5 family transposase IS1381A | 168 | 39 | 2 | 211 | Genetic Information Processing (09120) | K07498 |
| *deo*B | Phosphopentomutase | 122 | 85 | 2 | 211 | | |
| group_1719 | Ribonuclease | 198 | 9 | 5 | 208 | Genetic Information Processing (09120) | K03013 |
| *pkn*D | Serine/threonine-protein kinase PknD | 201 | 6 | 0 | 213 | Metabolism (09130) | K01001 |
| group_1484 | Tn3 family transposase | 68 | 139 | 3 | 210 | | |
| *xer*C | Tyrosine recombinase XerC | 199 | 8 | 4 | 209 | Genetic Information Processing (09120) | K03733 |
| *xerD* | Tyrosine recombinase XerD | 171 | 36 | 8 | 205 | Genetic Information Processing (09120) | K04763 |

**Table 6. 6. Genes specific to CC19 as identified by Scoary v1.6.16** [513]**.** Each column shows (from left to right) the name and annotation of the gene as identified by Prokka v1.12 [497], number of CC19 strains positive for a specific gene, number of CC19 strains negative for a specific gene, number of other CC isolates with presence or absence of a specific gene, respectively, followed by the KEGG database id entry name and functional class (where available). 118 hypothetical protein encoding genes specific to CC19 were excluded from the table.

| Gene | Annotation | No. of CC19 strain carrying specific gene | No. of CC19 strain devoid of specific gene | No. of other CC strains carrying specific gene | No. of other CC strains devoid of specific gene | Functional class with class reference in brackets | KEGG Pathway reference |
|---|---|---|---|---|---|---|---|
| *ars*A | Arsenical pump-driving ATPase | 28 | 4 | 3 | 385 | Genetic Information Processing (09120) | K21903 |
| *ars*D | Arsenical resistance operon trans-acting repressor ArsD | 28 | 3 | 2 | 385 | Genetic Information Processing (09120) | K21903 |
| *acr*3 | Arsenical-resistance protein Acr3 | 28 | 4 | 3 | 385 | Genetic Information Processing (09120) | K21903 |
| *cad*C | Cadmium resistance transcriptional regulatory protein CadC | 31 | 1 | 3 | 385 | Genetic Information Processing (09120) | K21903 |
| *cad*A | Cadmium-transporting ATPase | 31 | 1 | 3 | 385 | | |
| *cdr* | Coenzyme A disulfide reductase | 28 | 4 | 3 | 385 | | |
| *cop*B | Copper-exporting P-type ATPase B | 31 | 1 | 4 | 384 | Environmental Information Processing (09130) | K17686 |
| *act*P | Copper-transporting P-type ATPase | 29 | 5 | 3 | 385 | | |
| *dna*G | DNA primase | 28 | 4 | 3 | 385 | | |
| *sdp*R | Transcriptional repressor SdpR | 28 | 4 | 3 | 385 | | |
| *yad*H | Inner membrane transport permease YadH | 28 | 4 | 3 | 385 | | |
| group_1647 | IS6 family transposase IS1216E | 31 | 1 | 8 | 380 | Genetic Information Processing (09120) | K07498 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| group_1129 | ISAs1 family transposase ISSeq10 | 31 | 1 | 9 | 379 | Genetic Information Processing (09120) | K07498 |
| *virD*4 | Protein VirD4 | 31 | 1 | 6 | 344 | Environmental Information Processing (09130) | K03205 |
| *yad*G | putative ABC transporter ATP-binding protein YadG | 28 | 4 | 3 | 385 | | |
| *cop*Y | Transcriptional repressor CopY | 31 | 1 | 4 | 384 | | |

**Table 6. 7. Genes specific to CC8/CC10 as identified by Scoary v1.6.16** [513]**.** Each column shows (from left to right) the name and annotation of the gene as identified by Prokka v1.12 [497], number of CC8/CC10 strains positive for a specific gene, number of CC8/CC10 strains negative for a specific gene, number

of other CC isolates with presence or absence of a specific gene, respectively, followed by the KEGG database id entry name and functional class (where available). 19 hypothetical protein encoding genes specific to CC8/CC10 were excluded from the table.

| **Gene** | **Annotation** | **No. of CC/CC10 strain carrying specific gene** | **No. of CC8/CC10 strain devoid of specific gene** | **No. of other CC strains carrying specific gene** | **No. of other CC strains devoid of specific gene** | **Functional class with class reference in brackets** | **KEGG Pathway reference** |
|---|---|---|---|---|---|---|---|
| *cus*S | Sensor histidine kinase CusS | 33 | 0 | 1 | 385 | Two-component system (02020 ) | K07644 |
| *gal*M | Aldose 1-epimerase | 33 | 0 | 2 | 384 | Metabolism (09130) | K01785 |
| *aga*A_2 | Alpha-galactosidase | 33 | 0 | 3 | 383 | | K07407 |
| *ula*C | Ascorbate-specific PTS system EIIA component;putative licABCH operon regulator | 33 | 0 | 4 | 382 | Metabolism (09130) | K02821 |
| *cas*1 | CRISPR-associated endonuclease Cas1 | 31 | 2 | 1 | 385 | Genetic Information Processing (09120) | K15342 |
| *cas*2 | CRISPR-associated endoribonuclease Cas2 | 31 | 2 | 1 | 385 | Genetic Information Processing (09120) | K09951 |
| *ecfA*2 | Energy-coupling factor transporter ATP-binding protein EcfA2 | 13 | 20 | 8 | 378 | Environmental Information Processing (09130) | K16787 |
| ecf*T* | Energy-coupling factor transporter transmembrane protein EcfT | 13 | 20 | 8 | 378 | Environmental Information Processing (09130) | K16785 |
| *gal*K | Galactokinase | 33 | 0 | 3 | 383 | Metabolism (09130) | K00849 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *gal*T | Galactose-1-phosphate uridylyltransferase | 33 | 0 | 3 | 383 | Metabolism (09130) | K00965 |
| group_2123 | HTH-type transcriptional activator RhaS | 33 | 0 | 3 | 383 | | |
| *bag* | IgA FC receptor | 33 | 0 | 1 | 385 | | |
| group_416 | IS30 family transposase ISStin2 | 33 | 0 | 2 | 384 | Genetic Information Processing (09120) | K07498 |
| *ara*Q | L-arabinose transport system permease protein AraQ/ ABC.MS.P1; multiple sugar transport system permease protein - | 33 | 0 | 3 | 383 | Signalling and cellular processes (09183) | K02026 |
| *mel*D | Melibiose/raffinose/stachyose import permease protein MelD | 33 | 0 | 3 | 383 | | |
| *wbn*J | O-antigen biosynthesis glycosyltransferase WbnJ | 33 | 0 | 1 | 385 | | |
| *pho*B | Phosphate regulon transcriptional regulatory protein PhoB | 33 | 0 | 0 | 386 | | |
| *gat*A | PTS system galactitol-specific EIIA component | 33 | 0 | 3 | 383 | Metabolism (09130) | K02773 |
| *gat*C | PTS system galactitol-specific EIIC component | 33 | 0 | 3 | 383 | Metabolism (09130) | K02775 |
| *gat*B | PTS system galactitol-specific EIIB component | 33 | 0 | 3 | 383 | Metabolism (09130) | K02774 |
| *yes*O | ABC.MS.S; multiple sugar transport system substrate-binding protein | 33 | 0 | 3 | 383 | Signalling and cellular processes (09183) | K02027 |
| group_2468 | Putative multidrug export ATP-binding/permease protein | 13 | 20 | 8 | 378 | | |
| group_2128 | Putative multidrug export ATP-binding/permease protein | 13 | 20 | 8 | 378 | | |
| *yde*A | putative protease YdeA | 29 | 4 | 1 | 385 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *rha*D | Rhamnulose-1-phosphate aldolase | 33 | 0 | 3 | 383 | Metabolism (09130) | K01629 |
| *gal*E | UDP-glucose 4-epimerase | 33 | 0 | 3 | 383 | Metabolism (09130) | K01784 |

**Table 6. 8. Genes specific to CC17 as identified by Scoary v1.6.16** [513]**.** Each column shows (from left to right) the name and annotation of the gene as identified by Prokka v1.12 [497], number of CC17 strains positive for a specific gene, number of CC17 strains negative for a specific gene, number of other CC isolates with presence or absence of a specific gene, respectively, followed by the KEGG database id entry name and functional class (where available). 19 hypothetical protein encoding genes specific to CC17 were excluded from the table.

| **Gene** | **Annotation** | **No. of CC17 strain carrying specific gene** | **No. of CC17 strain devoid of specific gene** | **No. of other CC strains carrying specific gene** | **No. of other CC strains devoid of specific gene** | **Functional class with class reference in brackets** | **KEGG Pathway reference** |
|---|---|---|---|---|---|---|---|
| *aadK*1 | Aminoglycoside 6-adenylyltransferase | 96 | 2 | 8 | 314 | | |
| *aadK*2 | Aminoglycoside 6-adenylyltransferase | 96 | 0 | 0 | 324 | | |
| *ars*C | Arsenate reductase | 92 | 4 | 7 | 317 | Metabolism (09130) | K00537 |
| *asp*1 | Accessory Sec system protein Asp1 | 96 | 0 | 0 | 324 | Environmental Information Processing (09130) | K12268 |
| *asp*2 | Accessory Sec system protein Asp2 | 96 | 0 | 0 | 324 | Environmental Information Processing (09130) | K12269 |
| *bcr*A | Bacitracin transport ATP-binding protein BcrA | 93 | 3 | 1 | 323 | Environmental Information Processing (09130) | K19309 |
| *bin*3 | Putative transposon Tn552 DNA-invertase | 92 | 4 | 5 | 319 | | |
| *cre*C | Sensor protein CreC | 92 | 4 | 5 | 319 | | |
| *feu*C | Iron complex transport system permease protein | 94 | 2 | 3 | 321 | Environmental Information Processing (09130) | K02015 |
| group_2076 | Antitoxin epsilon | 91 | 5 | 10 | 314 | | |
| group_2374 | Aminoglycoside 3'-phosphotransferase | 94 | 2 | 5 | 319 | | |
| group_2599 | IS3 family transposase ISSag12 | 91 | 5 | 5 | 319 | Genetic Information Processing (09120) | K07498 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| group_2993 | Replication protein RepB | 93 | 3 | 6 | 318 | | |
| group_495 | FMN-binding protein | 94 | 2 | 5 | 319 | | |
| *inl*A | Internalin-A | 96 | 0 | 2 | 322 | | K13730 |
| *fim*A | type-2 fimbrial major subunit | 96 | 0 | 2 | 322 | | |
| *gtf*1 | Glucosyltransferase | 96 | 0 | 0 | 324 | | |
| *gtf*2 | N-acetylglucosaminyltransferase stabilizing protein | 96 | 2 | 7 | 315 | | |
| *hae*IIIM | Modification methylase HaeIII | 95 | 1 | 1 | 323 | | |
| *lyt*S | Two-component system, LytTR family, sensor histidine kinase | 93 | 3 | 8 | 316 | Environmental Information Processing (09130) | K07704 |
| *moa*A | GTP 3'8-cyclase | 91 | 5 | 10 | 314 | | |
| *nis*C | Nisin biosynthesis protein NisC | 93 | 3 | 2 | 322 | Environmental Information Processing (09130) | K20484 |
| *nis*I | Nisin immunity protein | 93 | 3 | 7 | 317 | Environmental Information Processing (09130) | K20489 |
| *pre* | Plasmid recombination enzyme | 94 | 2 | 7 | 317 | | |
| *reg*X3 | Sensory transduction protein regX3 | 96 | 2 | 0 | 322 | Environmental Information Processing (09130) | K07776 |
| *lep*B | Signal peptidase I | 94 | 2 | 2 | 322 | Metabolism (09130) | K03100 |
| *sec*A | Protein translocase subunit SecA | 96 | 2 | 0 | 322 | Metabolism (09130) | K03070 |
| *sec*Y | Protein translocase subunit SecY | 96 | 2 | 0 | 322 | Metabolism (09130) | K03076 |
| *sps*B | Signal peptidase IB | 96 | 2 | 0 | 322 | | |
| *str*A | Sortase A | 92 | 4 | 1 | 323 | Metabolism (09130) | K07284 |
| *top*B | DNA topoisomerase 3 | 95 | 1 | 8 | 316 | Genetic Information Processing (09120) | K03169 |
| *uvr*B | UvrABC system protein B | 91 | 5 | 4 | 320 | Genetic Information Processing (09120) | K03702 |
| *scp*A | C5a peptidase | | | | | | K08652 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *yyd*H | Putative peptide zinc metalloprotease protein YydH | 95 | 1 | 6 | 318 | Metabolism (09130) | K16922 |

**Table 6. 9. Genes specific to CC23 as identified by Scoary v1.6.16** [513]**.** Each column shows (from left to right) the name and annotation of the gene as identified by Prokka v1.12 [497], number of CC23 strains positive for a specific gene, number of CC23 strains negative for a specific gene, number of other CC isolates with presence or absence of a specific gene, respectively, followed by the KEGG database id entry name and functional class (where available). 46 hypothetical protein encoding genes specific to CC23 were excluded from the table.

| Gene | Annotation | No. of CC23 strain carrying specific gene | No. of CC23 strain devoid of specific gene | No. of other CC strains carrying specific gene | No. of other CC strains devoid of specific gene | Functional class | KEGG reference |
|---|---|---|---|---|---|---|---|
| group_1342 | 23-dihydroxybenzoate-AMP ligase | 50 | 2 | 1 | 367 | | |
| *ncs*B2 | 2-hydroxy-7-methoxy-5-methyl-1-naphthoate--CoA ligase | 51 | 1 | 1 | 367 | Metabolism (09130) | K20423 |
| *cyl*E | CylE uncharacterised protein | 50 | 2 | 0 | 368 | Signalling and cellular processes (09183) | K11052 |
| *cyl*K | CylK uncharacterised protein | 51 | 1 | 3 | 365 | Signalling and cellular processes (09183) | K11056 |
| *hts*T | Energy-coupling factor transport system substrate-specific component | 51 | 1 | 3 | 365 | Environmental Information Processing (09130) | K16926 |
| *mnt*R | HTH-type transcriptional regulator MntR, manganese transport regulator | 51 | 1 | 1 | 367 | Genetic Information Processing (09120) | K11924 |
| group_1593 | IS256 family transposase ISEnfa4 | 45 | 7 | 2 | 366 | | |
| *msb*A | Lipid A export ATP-binding/permease protein MsbA | 50 | 2 | 1 | 367 | Environmental Information Processing (09130) | K11085 |
| *glp*T | MFS transporter, OPA family, glycerol-3-phosphate transporter | 50 | 2 | 0 | 368 | Environmental Information Processing (09130) | K02445 |
| *dpn*M | Modification methylase DpnIIA | 45 | 7 | 6 | 362 | | |
| group_187 | Secreted 45 kDa protein | 49 | 3 | 2 | 366 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *btu*D | Vitamin B12 import ATP-binding protein BtuD | 49 | 3 | 2 | 368 | Environmental Information Processing (09130) | K06074 |
| *ram*B | XRE family transcriptional regulator, fatty acid utilization regulator | 45 | 7 | 7 | 361 | Genetic Information Processing (09120) | K07110 |

## 6.4 Discussion

In this chapter, whole genome sequence collection of GBS retrieved from Brazil, Canada, UK and United States population was examined for serotype distribution, virulence determinants, antibiotic resistance, phylogenetic relation and pan-GWAS. Generally, among the 26 Brazilian GBS genomes available for the analysis, serotype Ia (12/26, 46.1%) and V (10/26, 38.4%) were the most frequent, which is consistent with a recent study from Brazil [48] where serotype Ia and V were found predominating in individuals with GBS infections or colonization. However, in the same study [48], two more GBS serotypes (II, 19.1% and Ib, 18.7%) were found as the most common, which is different from our results where only two (7.4%) Brazil GBS isolates were found serotype Ib and none was serotype II. A similar predominance of serotype III (26.9%, 52/193), Ia (26.4%, 51/193), and V (15%, 29/193) has been observed among UK GBS dataset, which is in accord with a previous study from England and Wales where III (40%), Ia (24%), and V (14%) were found dominant among 4,878 GBS serotypes analysed [115]. Whereas the Canadian GBS genomes (serotype III and V) and United States GBS genomes (serotype V only) were deliberately selected for the thorough analysis of CC1 and CC17 specific genes since of the various clonal complexes (CCs), CC17 is overrepresented in GBS-infected newborns and CC1 mostly in adults for reasons that are still largely unknown.

The 417/447 (93.2%) GBS strains that were clustered into the most represented five major clonal complexes; CC1, CC10, CC17, CC19 and CC23 are also the most observed clades elsewhere [1, 665]. The serotype and MLST analyses showed that isolates of the same CC usually expressed one dominant serotype, indicating a relationship between CCs and serotypes. Serotype Ia clustered in CC23 and serotype V in CC1, while Ib strains clustered mostly into CC8/CC10 (91.3%) with a few into CC1 (8.6%). Of the 115 serotype III strains, two major CCs were detected, where 95 (82.6%) strains belonged to CC17 and 17 (14.7%) belonged to CC19 which is congruent with previous studies [1, 666].

To elicit GBS infection, a single virulence determinant is not adequate, however synchronized action of several pathogenicity factors would be required [283]. Serine rich repeat protein (Srr-1) promotes

GBS adherence to human epithelial (HEp2) and keratin cells while srr-2 enhances GBS virulence [667, 668]. The *srr*-1 gene was found largely associated with four CCs; CC1 (99%), CC8/CC10 (96.9%), CC19 (87.5%) and CC23 (98%) isolates as seen in a previous study [389] indicating increased adherence properties of GBS strains belonging to these CCs. While *srr*-2 was exclusively expressed by CC17 strains (88%), similar to the findings of a recent study [389], along with known increased pathogenicity of hypervirulent CC17 GBS isolates. Previously, *srr*-2 expression seemed to be restricted to serotype III/ST-17 strains [313]; however, in our case a single serotype IV/CC17 isolate (PHEGBS0476) was also found positive for *srr*-2.

Seven members of the alpha-like proteins (Alp) family, such as Cα (encoded by *bca*), Cβ (encoded by *bac*), Alp1 (*alp1*/ epsilon), Alp2 (*alp2*), Alp3 (*alp3*), Alp4 (*alp4*) and R4/Rib (encoded by *rib*) have been identified to date [317], which mediates GBS binding to human cervical epithelial cells and interacts with host cell glycosaminoglycan (GAG) to promote bacterial internalization [318]. This study showed that certain alpha like protein were predominantly expressed on the surface of specific GBS serotypes belonging to a specific CC; for example, *alp*1 was found predominantly in serotype V/CC1, V/ST19 and Ia/CC23 strains. However, none of the serotype II, Ib/CC8/CC10 and III/CC17 strains expressed *alp*1 gene, which is consistent with the previous findings [203, 319]. The gene *alp*2 was rarely expressed in some strains of serotype V, and II belonging to CC1(4%) and all serotype III/CC23 strains (3.8%), which is in agreement to previous results where *alp*2 was found in less proportion overall [325], and found significantly related to invasive strains [669]. The *bca* gene was expressed by all serotype Ib and II/CC8/CC10 strains (96.9%) except a single serotype II/ST1220 strain (PHEGBS0575) whereas the *rib* gene was found highly associated with serotype III CC17 (97.9%) and CC19 (87.5%) strains supporting the earlier reports [532]. The 92/94 (97.8%) of serotype III CC17 strains harboured the hypervirulent *hvg*A gene, which is a hallmark virulence gene and previously known to be carried only by ST17 [325], while this study shows that the *hvg*A gene was mainly detected CC17 (97.8%) strains, while other CCs; CC1 (3.8%) and CC8/CC10 (3%) strains also carried the *hvg*A gene, which has been noted by other investigators [311]. Furthermore, a strain with serotype IV/ST1351/CC17 was

found to carry the *hvg*A gene, which was markedly similar to a serotype/ genotype combination harbouring the *hvg*A gene seen in previous studies from France [269] and Ireland [532].

In the current study, pilus islands were found conserved among GBS strains. All GBS strains carried at least one of the three pilus islands. Strong association was observed between pilus island genes and specific CCs such as the combination of PI-1 + PI-2a found to be prevalent in CC1, CC8/CC10 and CC19 strains, while PI-1 + PI-2b were found exclusively in CC17 strains and PI-2a alone was found predominantly in CC23 strains, which has been noted before [389, 670-673].

One of the major objectives of this study was to investigate the antibiotic resistance determinants in all 417 GBS grouped into five CCs. Like the previous findings, a key proportion of strains (383/417, 91.8%) were carried tetracycline resistance genes: most frequently observed being the *tet*M gene in 357/383 (93.2%) strains, followed by *tet*O gene observed in 32/383 (8.3%) strains. However, this contradicts findings in a recent report on GBS from infants in China where *tet*O was found as the most commonly observed tetracycline resistance gene, followed by *tet*M in tetracycline resistant GBS [311], but agrees with the global consensus [1, 312, 389]. Clindamycin had been used as the main alternative to penicillin prophylaxis in women with penicillin allergy; however, since 2000 there has been a marked increase in resistance to clindamycin in patients of all ages [115]. Resistance to clindamycin and erythromycin is usually associated with three macrolide resistance genes: *erm*B, *erm*T, and *mef*A/*msr*D [674]. In this cohort, the erythromycin and clindamycin resistance mechanisms were not only limited to *erm*B (96/174, 55.1%), *erm*T (20/174,11.4%) and *mef*A/*msr*D (49/174, 28.1%) genes, but also mediated by *erm*A (49/174, 28.1%), *lsa*C (5/174, 2.8%) and *lnu*C (1/174, 0.5%) genes. The erythromycin-resistant strains carrying resistance genes in this study were similar from those in previous reports [505, 674].

Macrolide resistance was observed in 174/420 (41.4%) strains with a majority of the macrolide resistant GBS strains (114/174, 65.6%) clustered in CC1 (which were mostly serotype V (108/114, 94.7%)), as has also been observed in other studies [232, 541]. Similar to a previous study [505], certain macrolide

resistance genes were associated with certain CC isolates such as *erm*B (74/96, 77%), *erm*A (38/49, 77.5%) and *mef*A/*msr*D (40/49, 81.6%) found predominant in CC1 strains, while *erm*T was found particularly linked to CC17 strains (19/20, 95%) as observed in serotype III CC17 strains earlier [205]. Likewise, a high prevalence of aminoglycoside and chloramphenicol resistant GBS strains was observed in CC19, comparable to the findings of previous reports [214, 242], which reported accumulation of aminoglycoside and chloramphenicol resistance genes harbouring strains in CC19. Further it is worth noting that GBS isolates from UK have a high proportion of MDR strains conferring resistance to tetracycline, macrolide, aminoglycoside, and chloramphenicol in comparison to other region strains from Brazil, Canada, and United States.

This analysis showed that different CCs are defined by set of genes involved in GBS colonisation, niche adaptation and increased pathogenicity through acquisition of different virulence factors. For example, CC17 was observed to be the most invasive and hypervirulent lineage in comparison to CC1, CC8/CC10, CC19 and CC23 clones. CC17 specific genes mainly involved in micronutrient (iron) transport, signal transduction through two component systems and enhanced virulence through accessory secondary proteins. Among the CC17 specific genes, most of the genes identified have previously been described to cause GBS and other bacterial infections in human, for example the two component systems comprising sensor histidine kinases (HK) and response regulator (RR) proteins encoding gene (*lyt*S) was found exclusively in CC17 isolates as seen in earlier pan-GWAS study [362], which has significant role in bacterial signal transduction and is important for cellular survival, virulence, and cellular development [675]. An analogue of *lyt*S has also been identified in *Streptococcus mutans* which reveals its contribution to oxidative stress tolerance [676]. The *asp*1/*asp*2 - genes were also found specific to CC17 as identified earlier [362], which are a part of accessory secretory protein system, involved in the export of serine-rich glycoproteins (SRR) and important for virulence in several gram-positive species, including *Streptococcus gordonii* and *Staphylococcus aureus* [677]. The CC17 isolates also carried two translocase protein (secA and secY) which helps to pass *secretory proteins* across the cytoplasmic membrane [678], similar genes were found associated with iron acquisition, protein translocation, and virulence in *Acinetobacter baumanni* [679] and in earlier GWAS study

conducted on CC17 GBS [362]. Besides carrying specific genes, other distinctive features of CC17 isolates are mutations in CovS and STk kinases, which play a role in colonization and infection process of GBS [402].

Microbes secrete siderophores to acquire and solubilize ferric iron from the host through major iron transport systems as sufficient iron is needed during infection and post infection the host, any decrease in iron significantly reduces a pathogen's ability to colonize a host [680]. Majority of CC17 GBS isolates (97.9%) expressed *feu*C as also seen earlier [362], an iron uptake protein vs. the Feu ABC transporter identified in *Bacillus subtilis* for survival within host cell [681]. Various virulence factors in streptococci belong a diverse family of extracellular proteases that are involved in chemokines inactivation and employ immune cells at the infection site and expedite invasion of damaged cells [682, 683]. The CC17 specific *scp*A gene encodes C5a peptidase, which is a surface-associated serine protease that plays an important role in virulence of *S. agalactiae* by mediating invasion of epithelial cells [308]. Similarly, other CC17 specific genes in this analysed collection such as PI-IIb locus, part of which is represented by the CC17-associated genes *lepB, inlA,* and *fimA*, were similar as observed in earlier pan-GWAS of GBS [362].

CC23 associated genes mainly includes transcription factors and regulators that may regulate the expression of certain genes by turning them on and off [684]. Bacteria employ a variety of metal uptake and export systems and subtly control metal homeostasis by various transcriptional regulators, permitting them to adjust to alter environmental conditions and acquire metal ions which is essential for their survival in the environment or in their infected host [685]. One such example is CC23 specific *mnt*R [362], a manganese transport protein encoding gene that is known to provide GBS an advantage to acquire sufficient manganese which is required for its successful colonization and survival within host [686], similar to *S. aureus* that responds to this host-induced starvation by expressing metal transporters, such as MntH [686]. Further, bacterial toxins like (cylE and cylK) were also found to be CC23 specific that is known to promote GBS invasion to host cells and triggers host-cell lysis [668]. These results are in line with the findings of a recent pan-GWAS study of GBS [362].

In conflict, the genes defining CC1, CC8/CC10, and CC19 were involved with prokaryotic defence system, regulation, environmental sensing (PTS systems), as well as secretion systems. CC19 isolates had VirD4, a type IV secretion systems (T4SSs) protein gene, also observed earlier specific to CC19 GBS [362] that is associated with virulence effector translocation and mediate the conjugative transfer of plasmid DNA/ transposons into a wide range of bacterial species [687]. Whereas CC1 specific genes such as recombinases (*xer*C/*xer*D), translocases (*fts*K) and transposases including (Tn3 and IS5 family transposase) are putatively involved in site specific recombination [688, 689] which is conflicting with results of a recent study [362] where most of the CC1 specific genes were of unknown function and doesn't include genes found in this study. CC8/CC10 was characterized by genes involved in carbohydrate metabolism and uptake, such as the ABC transport system for multiple sugar transport that facilitates nutrient competition and survival [690]. In addition to this, bacterial defense CRISPR (*cas*1/*cas*2) genes were also found CC8/CC10 specific. All CC specific genes except CC1 identified in this study agrees to the recent pan GWAS conducted on wide-ranging publicly available GBS genomes from Kenya, USA, Canada, Malawi, and Netherlands [362].

This study has numerous limitations including a small and biased (serotype V and III from Canada and serotype V only from USA) publicly available GBS datasets were used for the GWAS analysis. Currently, limited GBS GWAS studies are available, this data could help to understand the genetic make-up of invasive and more colonizing strains, however, further studies are required on a broader and unbiased dataset. Secondly prokka software was used for annotation which failed to generate annotation for various genes, this may be resolved by using a more efficient annotation tool for the same purpose in future study. Lastly, this analysis is limited to the elicit genomic differences between GBS belong to different clades, further laboratory and epidemiological analysis will be needed to understand the biological processes of CC-specific genes.

In conclusion, this study reflects that GBS belong to particular CC contains specific set of genes giving potential benefit to bacteria to invade, colonize, better survive in the host cell and also to become more pathogenic by expressing virulent determinants.

# Chapter 7:

# General Discussion

## 7.1. General Discussion

GBS primarily colonizes the gastrointestinal and vaginal epithelium of 20–30% of healthy women [102], yet it is significantly associated with invasive diseases in newborns [156]. However, GBS burden among adults and those significant underlying diseases such as diabetes mellitus and cancer is considerably greater than previously recognized [691], resulting in substantial morbidity and mortality among adults as well as those observed for neonates [692]. Guidelines of implementation of prenatal screening and effective intrapartum antibiotic prophylaxis (IAP) have significantly reduced the early onset neonatal infections worldwide, however late-onset neonatal infections have largely been unaffected [339]. At present, no guidelines are available for the prevention of GBS disease in adults, but vaccines in development may hold potential could prevent iGBS disease across all at-risk population groups, including mother, fetus, infant, and the elderly or immunocompromised [122, 339]. The driving force behind rise in frequency of GBS disease among adults has not been fully explained, and recent trends in disease incidence have not been characterized. However, it is also possible that the rates are not increasing, just the recognition of their role in adult disease. In any event, the significance of GBS as a cause of severe infections among adults in different regions of the world are largely under investigated. There have been a couple of studies on the prevalence of iGBS disease in the UK [115, 517],. However, all were retrospective, geographically localized, and did not include information on GBS comprehensive antibiotic susceptibility testing (as opposed to collation of disparate hospital data), determination of MGEs carrying ARGs, potential capsular switch leading to different genotype/phenotype combinations and pan-GWAS. In contrast, detailed genomic analysis of adult iGBS, isolated from USA [76], Canada [338], France [269], Ireland [532] and Portugal [78], populations have been conducted to identify and characterize putative capsular switching events.

These unanswered queries and lack of information underpinned the significance and key purpose of this thesis: the epidemiological characterization of the GBS isolates responsible for an increasing number of invasive infections in adults of UK and Brazil, and carriage strains identified in antenatal screening of GBS colonized Australian pregnant women prior to delivery. In addition, quantitate distributions of

potential vaccine targets (capsule polysaccharides, surface proteins) and their associations with major antibiotic resistance markers within clonal lineages were also studied. This thesis objectives were divided into four main chapters (3-6);

The first results chapter (chapter 3) presented a detailed description of the epidemiology, serotype distribution, virulence determinants and clonal structure of the GBS populations causing invasive infections in adults in UK. Between January 2014 and May 2015 [693], isolates were mainly submitted from laboratories in England with a small proportion of submissions from elsewhere in the UK. Isolates received are from patients of all ages, with the majority from neonates (estimated >95%) and women of child-bearing age, with an estimated total adult isolate submission proportion of less than 11% (n=179) of all invasive cases. Information about pregnant/non-pregnant and postpartum status was not available for the women GBS was collected from. Only invasive isolates 179 iGBS isolates (178 from blood, and 1 from CSF) were used in this chapter to draw phylogenetic relationship and provides explicit genomic evidence for the existence of capsular switching in GBS. Rest of the isolates (n=14) were non-invasive and were included in chapter 4 and chapter 6. Large-scale recombination events with the exchange of the entire capsular locus, rather than switching based on transfer of small elements containing missing capsule-specific genes was found.

The second chapter (chapter 4) described an intensive genomic interrogation of whole genome sequences for systematic phenotypic antimicrobial resistance profiling and identification of mobile genetic elements carrying drug resistance genes in the clinical GBS strains (both invasive and non-invasive; n=193). High concordance between phenotypic and genotypic antibiotic resistance profiling was found, further the MGEs identified play a minor overall role in antibiotic resistance dissemination, including five novel MGEs identified in this study. Tn*916*/Tn*1545* family elements were prevalent in different CC and had insertion in specific intergenic sites, suggesting that horizontal gene transfer is fundamentally of very low prevalence in GBS, as also shown by the Tn*916* insertions by other authors [1] and very limited movement of these elements since the 1950s when tetracycline use was extensive.

The third results chapter (chapter 5) characterized a small available number of isolates (n=26) recovered from invasive infections in adults in Brazil and 171 isolates recovered from colonized pregnant women at <=22 weeks' gestation from Australia to study antibiotic susceptibility profiles and serotype distribution of GBS circulating in two distinct geographic regions and study populations. The source of all Brazilian isolates were blood culture and the date of collection for these isolates ranged between June 2010 and May 2017, however the isolate numbers (n=26) doesn't represent the total number of iGBS cases reported in Brazil between June 2010 and May 2017. The 26 iGBS analysed in this chapter, were collected in a tertiary care hospital of Southeast region of Brazil, Rio de Janeiro, Brazil. Currently there is no GBS surveillance available in Brazil that could reflect the total cases of GBS cases reported in different regions of Brazil per year, however compared to a previous study from the Rio de Janeiro, Brazil [371], where 3,647 specimens were collected from pregnant women living in Rio de Janeiro State, Brazil, and assessed for GBS colonization prevalence, over a period of 8 years (2008-2015), more GBS cases are expected to report to RJ Hospital over a period of seven years (2010-2017), however were not obtainable for this analysis. As for Australian pregnant women GBS dataset, a subset of isolates (n=171/347), collected from pregnant women at ≤22 weeks' gestation (visit one) as a part of larger cohort study, Predict1000 [384] was sent by Australian collaborators and were analysed knowing that understanding antibiotic resistance trends and surveillance of circulating, disease-causing target proteins and capsular polysaccharides are useful for treatment strategies, and to inform current vaccines targeting surface proteins and capsular polysaccharides. Addressing data gaps especially by world region and some at-risk populations is fundamental to evidence-based decision-making during vaccine design.

In chapter 6, an in depth pan-GWAS was performed on a total dataset of 447 iGBS genomes from adults, including 228 publicly available genomes from Canada [313], United States [76] in addition to UK (n=193) and Brazil (n=26) GBS genomes that were sequenced earlier in this study. Deliberately serotype III and V GBS were chosen from Canada and United States knowing they mediate severe

infection in neonates and adults, respectively to determine if there were lineage-specific genes playing role in enhanced GBS pathogenicity, colonization, and better survival in host.

### 7.1.1. GBS Epidemiology:

Overall, GBS recovered from British adults (both invasive and non-invasive strains), Brazilian adults (only invasive strains) and Australian pregnant women (asymptomatic carriage stains) consisted of a wide range of serotype and genotype combinations. It is obvious that a clonal genetic structure exists among the GBS isolates recovered from adults of Britain and Brazil, by the fact that some lineages appeared more frequently and are more likely to cause invasive infections in the population. Group B streptococcal infection has been documented as a crucial and apparently increasingly frequent cause of invasive disease in non-pregnant adults aged >65 years, especially in those with chronic medical conditions [133]. Contrary to this concept, iGBS infections were more frequently observed among females belonging to the childbearing age group (18-44) years in our UK dataset [693]. In 2014, 1676 and in 2015, 1870 cases of GBS bacteraemia were reported by laboratories in England, Wales and Northern Ireland to PHE [480, 481]. Isolates received are from patients of all ages, with the majority from neonates (estimated >95%) and women of child-bearing age, with an estimated total adult isolate submission proportion of less than 11% (n=179) of all invasive cases. In both years, GBS rates were higher in males than females in all age groups except the 15 to 44 years age group (females 2.2 and males 0.5/100,000) [480, 481] which agrees to most of our findings where more GBS samples collected in 2014-2015 were from male patients belonging to age groups (> 65 yrs) respectively, and for the 18-44 years age group more GBS were collected from female patients. A discrepancy observed for the number of iGBS isolates observed for the 45-64 years age group in this study where more isolates were recovered from female (n=12) than male (n=11), contrary to the PHE GBS surveillance report of 2014 and 2015, this difference could possible arise due to the total number of GBS bacteraemia cases reported in PHE reports but not the actual GBS samples recovered in the laboratories. Increased iGBS incidence in women of childbearing age of UK during 2015 and 2016 [517], and in our 2014 and 2015 UK GBS dataset suggests potential for post-partum GBS infection in female. There has been evidence of

increased prevalence of iGBS infections in younger female immunocompromised adults, for example in South Africa, younger female adults with trauma and HIV infection were found more commonly associated with iGBS [694], and in metropolitan Atlanta more than 40% of younger female adults aged between (18–64 years) identified with iGBS infection had diabetes [695]. Other health conditions, such as cirrhosis, history of stroke, breast cancer, decubitus ulcer, and neurogenic bladder, have also seen associated with increased risk of iGBS disease in younger females in multivariate analysis [695]. However, due to lack of data on pregnancy status and comorbidities of UK females in our dataset, the reason of increased GBS prevalence in young females compared to elderly adults remain unknown.

Contrary to our findings on UK iGBS dataset, more Brazilian male patients aged greater than 55 years had iGBS disease than females, supporting results published from different geographical locations where increased iGBS infections were reported in elderly male patients [20, 696-698]. Due to the lack of information of clinical diagnosis and medical underlying condition of patients infected with iGBS in UK in the metadata provided, reasons behind the sex and age group imbalance remain to be clarified.

### 7.1.2. Serotype distribution:

GBS serotypes correlate with pathogenic virulence [389] and are known to vary geography, temporally, and according to the patient age, at the onset of GBS disease [265, 345]. The most prevalent serotypes (Ia, II, Ib, III and V) account for more than 95% GBS infections reported worldwide including 93% of serotypes in Europe, 96% in the United States, and 89% in Western Pacific [265, 345].

In this thesis, the assessment of the incidence of serotypes distributed in British, Brazilian and Australian GBS populations disclosed some exciting findings. While for EOD, maternal GBS carriage is a necessity [699], the cause for GBS infections in adults is poorly understood, possibly arising initially from asymptotic colonization [281]. Among the British and Brazilian adult populations, the distribution of the GBS serotypes were different. Among British adults, the most common serotypes observed were

serotype III (26.9%, 52/193), Ia (26.4%, 51/193), and V (15%, 29/193), whereas among Brazilian GBS, serotype Ia (46.1%), V (38.4%), and Ib (7.6%) were most prevalent.

Despite serotype disparity observed between the two geographically distinct populations, the British GBS serotype results are consistent with a report on the same population published for isolates collected from 1991-2010 [691]. GBS serotype distribution for the Brazilian GBS found in this study is incongruent with some studies from Brazil, Europe and the United States where predominance of serotypes III and V prevails among invasive infections in younger and elderly adults, respectively [389, 609, 700, 701]; however, agrees with studies from Brazil [48] and Portugal [602] where serotype Ia was found dominant among GBS causing invasive infections in non-pregnant adults and may be subject to geographical and temporal biases between these sets of studies. The higher prevalence of serotype Ia (44.4%) and under presentation of serotype III (9.6%) in the Brazilian adult GBS collection, is in agreement with similar observations elsewhere [515, 602, 702], reinforces the differences between neonatal and adult invasive infections and the importance of serotype Ia and V, in adult invasive infection. It is well-known that GBS colonizes pregnant women and is transmitted to child either in utero, during the delivery or in the postnatal period [699]. Not much is known regarding the GBS transmission beyond pregnancy. Some reports revealed that GBS not only colonize non-pregnant women including both childbearing age and elderly women, but can also cause iGBS infections in these populations [281]. High predominance of serotype III among UK adults remains unclear, however suggests that non-pregnant populations (including men, children and non-pregnant women) may be colonized at birth or colonized (stably or sporadically) throughout their life, acting as GBS reservoirs and contributing thereby to its sustained dissemination to humans.

GBS isolates recovered from Australian pregnant women at ≤22 weeks' gestation had a wide range of serotypes with serotypes Ia (22.8%) and V (22.2%) being more frequent, but closely followed by serotypes III (20.5%) and II (15.8%), which is consistent with a recent report on the GBS serotype distribution in pregnant women in Australia, a subset of isolate was used in this analysis [384]. Three serotypes: Ia, III and V were the most identified GBS serotypes at both vaginal and rectal sites of women

pregnant ≤22 weeks, however differences in order of serotype prevalence were identified at both sites. The most prevalent serotype Ia (26.1%), III and V (21.4%) obtained from maternal vaginal sites, are also the leading GBS serotypes (Ia and III) causing neonatal disease worldwide [263], supporting the vertical transmission of GBS from mother to baby and its role in causing disease in neonates. Also worth mentioning is the importance of serotype Ia among neonatal infections: in Portugal serotype Ia was found to be more associated with EOD (31%), and serotype III was more common in LOD (63.6%) [364]. Serotype III isolates were confined to ST17 and ST19 (or SLVs) in our UK GBS dataset as found elsewhere [703] and more prevalent in female patients, an interesting correlate as it is also the most common serotype in infants [339, 691, 704]. Further worth noticing are the three Australian mothers colonized with different serotypes at vaginal and rectal sites, as also reported from Africa, Australia and USA [595, 605, 705, 706], highlighting the importance of identifying multiple serotype colonisation, which is often not investigated in epidemiological studies, but it is an important consideration for serotype-based vaccine development and implementation to ensure less abundant serotypes are not under-represented [707].

### 7.1.3. GBS virulence and genetic lineages.

GBS expresses a variety of virulence factors that are critically important for its pathogenicity [668]. Several studies [85, 281, 659] regarding the clonal structure of GBS population have identified a strong association between GBS virulence genes, specific serotypes and genetic lineages that causes invasive disease, yet differences have been reported in different regions reflecting the dynamic nature of GBS and further supporting continuous surveillance.

The main lineages associated with invasive infections in adults of the UK and Brazil described in chapter 3 and 5 mostly agree with those circulating in other countries [1, 86, 515, 532]. However, additional data analysis conducted in chapter 3 and chapter 6 underline the significance of local dynamics, indicating that genetic evolution of GBS presents with a geographic structure and may

depend on local factors. Serotype Ia ST23 isolates were the most prevalent lineage among UK GBS population (19.6%, 38/193) consistent with reports from Iceland [515], Portugal [602, 672] and Spain [670], but contrary to the findings of Brazilian GBS population which mostly comprised of serotype V ST1 isolates (64.6%, 9/26). For both UK and Brazil GBS datasets, we found expected correlates between virulence gene and pilus islands (PI) segregation to particular serotypes or CCs as reviewed previously [281], representing that the presence of a particular surface protein gene is both a feature of serotype and clonal property, indicating that horizontal gene transfer is exceptionally rare in GBS as also shown by the Tn*916* insertions in previous study [1].

In UK and Brazilian adult GBS datasets, certain alpha like proteins (which mediate GBS binding to human cervical epithelial cells and interacts with host cell glycosaminoglycan to promote bacterial internalization [708]) were predominantly expressed on the surface of specific GBS serotypes belonging to a specific CC. For example, serotype V/ST1/*alp*1, serotype Ib and II/ST8+ST12/*bca*+*bac*, and serotype Ia/ST23/*bib*A were predominated in CC1, CC8/CC10 and CC23, respectively. While serotype III/ST17/*hvg*A+PI-2b were detected in CC17 majorly and serotype III/ST17+ST19/*rib* were distributed between CC17 and CC19, which agrees with the findings of previous study where increased adherence, protective immunity and virulence were identified in iGBS isolates, harbouring *hvg*A and *rib* genes, and belonging to CC17 and CC19, respectively [402]. For GBS-EOD isolates, five capsular serotypes are relatively more common namely Ia, Ib, II, III and V, in contrast, the GBS-LOD isolates are more associated with serotype III and CNS infection [120, 154, 709, 710]. Genetic variation in the GBS genomes, within and outside the capsule biosynthesis region, influence the onset time for acute invasive disease and invasion of the meningeal tissue [711]. No statistically significant genetic variation with in *hvg*A influence the ability of GBS to invade the CNS, although such genes are generally essential for GBS meningeal tropism [711]. Altogether, abundance of serotype III (potentially *hvg*A positive) isolates in GBS-LOD and influence of genetic variation in *cps* biosynthesis region on the disease onset and meningeal tissue invasion defines the potential reason behind hyper-virulent strains causing infant disease who have comparatively a better immune system than neonates, rather expecting the most

virulent strains to be those causing disease in an immunocompetent host such as mid-life adults and pregnant women.

### 7.1.4. Antibiotic resistance and mobile genetic elements

Antimicrobial resistance in bacteria is increasing globally and the extensive use of antimicrobials is inevitably driving expanding prevalence and evolution of diverse MDR strains [712]. Several epidemiological studies reported an increase in the incidence of macrolide resistance in GBS, in particular in strains belonging to CC1 [78, 713]; which is consistent with chapter 4 and 6 findings where most of the UK and Brazil macrolide resistance gene positive GBS strains were characterized as serotype V/ST1 or a single locus variant of ST1 and belonged to CC1. This include four isolates that carried *erm*B in combination with *tet*M gene, both genes were harboured on a Tn*3872* element which has been previously reported to be restricted to CC1 only [1, 76, 567]. The limitation of macrolide resistance to serotype V and CC1 suggests selection and expansion of clones is the main factor driving clone expansion rather than horizontal gene transfer. One, conflicting result was observed in a Japanese study conducted by Morozumi *et al*. [714] in which (7/32, 21.9%) serotype Ia and (24/88, 27.3%) serotype III GBS isolates showed macrolide resistance and the frequency of serotype V isolates in 150 GBS isolates obtained from invasive infections in neonates was zero. These findings raise the questions whether the lack of serotype V GBS isolates is an epidemiological variation in Japan, or alternatively, whether macrolide use as an alternative to beta lactam can be attributed to the ability of GBS to switch capsular serotypes as seen in other studies [77, 269, 532, 715]. To elucidate if macrolide resistant serotype III and Ia underwent capsular switching from serotype V background, further genomic analysis are needed on Morozumi *et al* [714] dataset.

In bacterial populations, clonal expansion ensures vertical genetic inheritance throughout lineages, while horizontal gene transfer allows the rapid acquisition of new traits from external sources [434, 716]. The latter event occurs at low frequency, but its impact on microbial evolution and adaptation is

tremendous when it provides the recipient cell with a new selective advantage [717]. Horizontal gene transfer is driven by several mechanisms, like transformation, transduction and conjugation, the latter is a contact dependent process during which the unidirectional transfer of DNA takes place from a donor to a recipient cell, and this often relies on particular conjugative elements, like plasmids, ICEs and IMEs [717]. ICEs are the most common methods for introducing heterologous DNA into bacteria and propagate passively by host replication and cell division [717]. Wild-type ICEs vary greatly in size (20 kb to >500 kb) [718]and are mostly inserted specifically based on the serine or tyrosine recombinase and the targets are repeatable and well defined regions [444]. In all the analysed GBS datasets from the UK, Brazil and Australia, a high rate of tetracycline resistance was noted, as was noted by others [1, 505]. In agreement with data from the literature [505, 719] the *tet*M gene was the most frequent determinant among tetracycline non-susceptible strains, due to high resistance rates among GBS strains of all serotypes worldwide; therefore, tetracycline is no longer recommended for the treatment of GBS infections [153]. In UK cohort, of (154/175, 88%) *tet*M positive isolates, most of the isolates carried *tet*M by Tn*916*/Tn*916*-like elements (n=87, 56.4%) followed by Tn*5801* (n=67, 43.5%). These elements showed specific association with CC, and were inserted in the exact genomic sites as also seen in previous study [1], suggesting very limited movement of Tn*916* elements since the 1950's (when the extensive use of tetracycline has led to the selection of specific resistant and infectious GBS clone) because of remarkably rare horizontal gene transfer movement events in GBS and the one seen in the present day strains are the result of "one off" events that probably happened generations ago because of tetracycline selection.

Interestingly, out of 41 macrolides resistant UK GBS isolates, a single isolate carried macrolide resistance gene (*erm*T), harboured on a plasmid indicating this is a very uncommon mechanism of resistance gene dissemination in GBS. While four UK GBS isolates showed fluoroquinolone resistance due to mutation in gyrases and topoisomerases; for all other isolates, ARGs were acquired by MGEs including five novel MGEs: ICE*Sag*84 and ICE*Sag*100414 carrying *erm*A alone, ICE*Sag*662 containing *erm*B, *tet*S, *ant*(6-Ia) and *aph*(3'-III), ICE*Sag*71 carrying *erm*B and *tet*O, and ICE*Sag*139 containing *erm*A and the high gentamicin level resistance gene *aac*(6')-*aph*(2").

Of 193 UK GBS isolates, 40 (20.7%) carried multidrug resistance determinants and 41 MGEs were identified in total harbouring single (other than *tet*M) or multiple ARGs. Three most common target gene (TGs) favouring the integration of these 41 MGEs in UK GBS isolates were 3' end of *rum*A (29/41, 70.7%), followed by integration at 15 nucleotides prior to the 3' end of *rpl*L (21.9%, 9/41) and 3' end of *rps*I (7.3%, 3/41), which is in agreement with the finding of an extensive study carried out on 124 strains belonging to 27 streptococcal species, that described 15 TGs, as a common hotspot for ICEs integration in *Streptococcus* including sites observed in this study - *rum*A, *rpl*L and *rps*I. The *rp*lL gene, identified as the second most common target gene for MGEs in our UK human GBS dataset, was previously detected as the most frequent integration site for the ICEs in bovine *S. agalactiae* isolates *(*56/76, 73.6%) [365], which is contrary to this study finding where *rum*A observed as the most frequent integration site, and may reflect a difference between agricultural and human pathogen datasets.

Increased macrolide resistance rates were shown by our UK GBS isolates, which is consistent to the earlier reports in UK and other regions [389, 505, 691], in contrast, low resistance rates were observed for macrolides among the GBS isolates recovered from Brazilian adults and Australian colonised mothers. Low resistance rates have previously been reported in a study on maternal GBS carriage from Australia [621] and however opposite high resistance rates of macrolides were observed in earlier study from Brazil [380] as well as in Asia, Europe, United States and Canada [376, 379, 380, 397, 609, 619, 620]. Continuous surveillance of macrolide resistance will help clarify the mechanisms by which the dissemination of resistance occurs in GBS, and to which extent it depends on local genetic lineages.

Penicillin remains susceptible in both UK and Brazilian GBS cohort, and none of the mutations identified in PBP genes resulted in reduced penicillin susceptibility on AST. Penicillin is the first line of antibiotic used to treat GBS infections in patients of all ages, and remain effective against GBS since the last 80 years after its discovery in late 1930s [720], this is because PBP genes are the part of core genome, and not carried by MGEs which is more worrisome because resistance genes can spread at a faster rate to different bacteria through horizontal gene transfer compared to mutations causing an

antibiotic resistance. Penicillin inhibits cell wall synthesis and is found to be very effective against Gram-positive but not against Gram-negative bacteria (due to the presence of the outer membrane) or the tubercle bacillus (because of the extra thick cell wall) [721]. Interestingly, the few GBS UK isolates carrying the chloramphenicol resistance genes *cat*Q and *cat*(p*c*194) in combination with aminoglycoside, macrolide and tetracycline resistance genes were confined to CC19 and the majority (75%) of them were represented by serotype V/ST19 lineage consistent with previous reports [214, 242] proposing a clonal spread for this genetic determinant. Chloramphenicol is a rarely used drug in the United States because of its known severe adverse effects, such as bone marrow toxicity and grey baby syndrome [722], however in UK, chloramphenicol is safe to treat eye and ear infections in children aged younger than 2 years, where antibiotic treatment is required [723].

### 7.1.5. Capsular switching and pan-genome analysis:

Genetic recombination following a genetic transfer event can produce heterogeneous phylogenetic histories within sets of genes that share a common ancestral origin [724]. Delineating recombination events will enhance our understanding in genome evolution [724]. To determine the genetic relationships between strains and to discriminate within and between the large collections of GBS strains based on their genotypic and phenotypic characteristics, different typing method are used widely [725]. However, challenges are linked to obtain similar results through different typing methods, when dealing with diversified phenotype/genotype collections of GBS [132]. The seminal paper describing the concept of a pan-genome was based on an analysis of eight GBS genome sequences that separates pan-genome analysis into comparisons of core genome and a variable genome [726]. The expectation being that most strain associated genes are stable and can be separated from foreign and mobile elements, and in this way should be able to discern GBS evolution related to specific characteristics identifying intra- and inter-species horizontal gene transfer [727]. Sporadic reports are emerging of serotype-genotype mismatches clustered together with in a clade [269, 532, 715], providing vital evidence for emerging capsular switch events. Therefore, within a GBS population strains could share the same serotype, regardless of their diverse genetic backgrounds. This could impact future GBS

vaccination programs [247, 728, 729] through providing selective pressure for specific genotypes to escape vaccine coverage [730].

The extensive genomic investigations including MLST and phylogenetic approaches of characterising iGBS isolates from adults in the UK found a high number of capsular serotype-CC mismatches (14/179 isolates), from CC1-serotype V to serotype Ib (n=2), II (n=2) and VI (n=2); from CC19-serotype III to serotype V (n=5); from CC23-serotype Ia to serotype V (n=2); and CC-17-serotype III to serotype IV capsule (n=1). The capsule switching among ST1 and CC17 GBS strains was previously reported in USA [76], Canada [338], France [269], Ireland [532] and Portugal [78], confirming the geographical spread of these clones. However, compared to CC1 (47%) and other CCs (CC8/CC10 – 44%, CC19 – 24% and CC23 – 42%), the recombination rates in CC17 were found exceptionally low (3%) in Da cunha *et al* study [1], which is supported by other studies as well [205, 402], supporting the hypothesis that since after the expansion of major clones in the 1950's, CC17 isolates corresponds to a genetically isolated lineage, occupying a specific niche in the digestive tract with little interaction with other GBS clones [402]. In this analysis, no IS or IME motifs were found at the end of recombined *cps* locus therefore the only mechanism left is recombination of two large strands of DNA by homologous restriction at the ends. How the large piece of DNA was absorbed remained unclear. Clonal spread relative to common recombination can be proved by investigating consistency and retention of unique SNPS of the recombination/insertion ends by large scale WGS comparison.

The results of chapter 3 provided clear genomic proof for the presence of wide-spread capsular switching in GBS. These events have been previously reported from different geographical locations [76, 78, 269, 338, 532], with the suggested mechanism for capsule switch events by horizontal gene transfer of the whole capsular locus instead of genetic transfer of only the serotype-specific genes, which is in agreement by our findings. Capsular switch alone would not change the disease severity however it may be of critical importance in planning future vaccination strategies against this pathogen.

In Brazil, Australia and UK, a CPS based vaccine against serotypes Ia, Ib, II, III, IV and V could possibly offer protection to 80-95% of the population including neonates, pregnant women, and non-pregnant adults; yet the hurdle remains for GBS serotypes that may arise due to capsular switching and are currently not included in the hexavalent vaccine under development. Furthermore, non-typeable maternal carriage and nonpregnant isolated GBS strains [341, 731, 732], observed in different regions could escape CPS based vaccine coverage. Another aspect of potential vaccine failure may arise after the introduction of vaccine, as this could provide a selective pressure for virulence genotypes to escape vaccine coverage and this has been observed for *S. pneumoniae* [267], resulting in rare serotypes expansion not covered by vaccine and decrease in infection caused by serotypes included in the vaccine. In addition to this, capsular replacement can also affect bacteria phenotype, including virulence and host cell interactions [733]. Surface proteins and surface protein genes expressed by GBS strains can be useful in epidemiological studies and in formulation of non-capsule-based vaccines as these immunisations with these proteins has been shown to confer protective immunity in animal models [273, 692]. Unfortunately, no single gene or surface protein is sufficiently common for it to be considered as the basis for a successful pan-GBS vaccine therefore, a universal vaccine that could prevent all GBS disease is still beyond current knowledge.

Pan-genome wide association study of UK, Brazilian, United States and Canadian GBS strains in chapter 6 showed that different CCs are defined by a set of genes involved in GBS colonisation, niche adaptation and increased pathogenicity through acquisition of different virulence factors. For example, CC17 was predicted to be the most invasive and hypervirulent lineage compared to CC1, CC8/CC10, CC19 and CC23 clones due to the high prevalence of genes such as *lyt*S, *asp*1/*asp*2, *sec*A and *sec*Y, *feu*C, *scp*A, and PI-2b, associated with known roles in bacterial signal transduction [362], cellular survival [675], virulence [308, 362, 677], and cellular development [678], likely the underlying reason for the association of CC17 GBS lineages to cause neonatal diseases. While CC23 specific genes such as *mnt*R and *cyl*E and *cyl*K were more associated with known capacity for successful colonization [362, 668] and survival within host, which are absent in less invasive lineages. The findings of chapter 6 further highlighted that the lineage-specific basis of GBS niche adaptation and virulence, were not

limited to specific geographic areas, which is consistent with the observations of recent pan-GWAS study of GBS [362].

## 7.2. Limitations

There are several limitations in this thesis as the sample set is a cross sectional observational study limited to only referred isolates within adults of England and Wales, UK and Brazil over differing time periods of collection. Further, no information about the pregnancy, underlying medical conditions, outcomes of UK adult's patient after the sample collection was available and further restricted investigations of mortality rate associated with GBS infection in UK adults. The GBS specimens from pregnant women of Australia were limited and no metadata regarding patient age, medical conditions, subsequent neonatal disease arising from delivery by GBS colonized mothers was available. Further, when the Covid-19 pandemic hit, Cardiff University wet lab research was suspended from March until late September and therefore many further planned investigations for the Australian GBS cohort including further sequencing were not possible to complete.

## 7.3. Future studies

While many of the antimicrobial resistance genes were identified in mobile elements, no mating studies were performed to confirm or investigate their mobility. This would have been an important confirmation that these genetic elements had the capacity to disseminate the resistance phenotype. Furthermore, the underlying mechanism of capsular switching was not identifiable as there were no characteristic transposases or other IS elements at the boundary of insertion. Collection, re-sequencing and in-depth genomic analysis should be performed on all published strains with similar CC-capsular mismatches to determine if these events have sporadically occurred independently, or whether (despite their wide geographic reporting) they represent rare events that have arisen from a clonal event followed

by wide global dispersal. The combination of genes may represent a high fitness cost or other immunological/colonising disadvantage that is suppressing their emergence to greater prevalence.

# Chapter 8:

# References

1. 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; doi: 10.1038/ncomms5544.
2. Prevention of Group B Streptococcal Early-Onset Disease in Newborns: ACOG Committee Opinion, Number 782. Obstet Gynecol. 2019;134(1):1; doi: 10.1097/aog.0000000000003334.
3. Gynaecoloists RCoO: Women in preterm labour to be offered antibiotics to combat Group B Strep. In.; 13 September 2017. Press release.
4. Parks T, Barrett L, Jones N. Invasive streptococcal disease: a review for clinicians. British Medical Bulletin. 2015;115(1):77-89; doi: 10.1093/bmb/ldv027.
5. Lancefield RC. A serological differentiation of human and other groups of hemolytic streptococci. Journal of experimental medicine. 1933;57(4):571-95.
6. Facklam R, Elliott J. Identification, classification, and clinical relevance of catalase-negative, gram-positive cocci, excluding the streptococci and enterococci. Clinical microbiology reviews. 1995;8(4):479-95.
7. Facklam R. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clinical microbiology reviews. 2002;15(4):613-30.
8. Voluntary surveillance of streptococcal bacteraemia in England. https://www.gov.uk/government/publications/pyogenic-and-non-pyogenic-streptococcal-bacteraemia-annual-data-from-voluntary-surveillance. (1999-2019). Accessed January 2022.
9. Gray BM, Stevens DL. Streptococcal Infections. Bacterial Infections of Humans. 2009:743-82; doi: 10.1007/978-0-387-09843-2_35.
10. Steer AC, Lamagni T, Curtis N, Carapetis JR. Invasive group A streptococcal disease. Drugs. 2012;72(9):1213-27.
11. Group A streptococcal infections: guidance and data. Available at: https://www.gov.uk/government/collections/group-a-streptococcal-infections-guidance-and-data. Accessed October 2021.
12. Giuliano A, Lewis Jr F, Hadley K, Blaisdell FW. Bacteriology of necrotizing fasciitis. The American Journal of Surgery. 1977;134(1):52-7.
13. Sriskandan S. Severe peripartum sepsis. The journal of the Royal College of Physicians of Edinburgh. 2011;41(4):339-46.
14. Cantwell R, Clutton-Brock T, Cooper G, Dawson A, Drife J, Garrod D, et al. Saving mothers' lives: reviewing maternal deaths to make motherhood safer: 2006-2008. The eighth report of the confidential enquiries into maternal deaths in the United Kingdom. BJOG: an international journal of obstetrics and gynaecology. 2011;118:1-203.

15. Stevens DL. Streptococcal toxic-shock syndrome: spectrum of disease, pathogenesis, and new concepts in treatment. Emerging infectious diseases. 1995;1(3):69.
16. Efstratiou A. Group A streptococci in the 1990s. Journal of Antimicrobial Chemotherapy. 2000;45(suppl_1):3-12.
17. Lancefield RC, Hare R. The serological differentiation of pathogenic and non-pathogenic strains of hemolytic streptococci from parturient women. The Journal of experimental medicine. 1935;61(3):335.
18. Fry R. Fatal Infections by Haemolytie Streptococcus Group B. Lancet. 1938:199-201.
19. Broughton DD, Mitchell WG, Grossman M, Hadley WK, Cohen MS. Recurrence of group B streptococcal infection. The Journal of pediatrics. 1976;89(2):182-5.
20. Farley MM, Harvey RC, Stull T, Smith JD, Schuchat A, Wenger JD, et al. A population-based assessment of invasive disease due to group B Streptococcus in nonpregnant adults. N Engl J Med. 1993;328(25):1807-11; doi: 10.1056/nejm199306243282503.
21. Schuchat A. Epidemiology of group B streptococcal disease in the United States: shifting paradigms. Clinical microbiology reviews. 1998;11(3):497-513.
22. Watanabe S, Kirikae T, Miyoshi-Akiyama T. Complete genome sequence of Streptococcus dysgalactiae subsp. equisimilis 167 carrying Lancefield group C antigen and comparative genomics of S. dysgalactiae subsp. equisimilis strains. Genome Biol Evol. 2013;5(9):1644-51; doi: 10.1093/gbe/evt117.
23. Van Belkum A, Tassios P, Dijkshoorn L, Haeggman S, Cookson B, Fry N, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clinical Microbiology and Infection. 2007;13:1-46.
24. Muthukumar A, Zitterkopf NL, Payne D. Molecular tools for the detection and characterization of bacterial infections: a review. Laboratory Medicine. 2008;39(7):430-6.
25. Pradhan P, Tamang JP. Phenotypic and Genotypic Identification of Bacteria Isolated From Traditionally Prepared Dry Starters of the Eastern Himalayas. Frontiers in Microbiology. 2019;10:2526.
26. Donelli G, Vuotto C, Mastromarino P. Phenotyping and genotyping are both essential to identify and classify a probiotic microorganism. Microbial ecology in health and disease. 2013;24:10.3402/mehd.v24i0.20105; doi: 10.3402/mehd.v24i0.20105.
27. Krzyściak W, Pluskwa KK, Piątkowski J, Krzyściak P, Jurczak A, Kościelniak D, et al. The usefulness of biotyping in the determination of selected pathogenicity determinants in Streptococcus mutans. BMC Microbiology. 2014;14(1):194; doi: 10.1186/1471-2180-14-194.
28. Toi CS, Cleaton-Jones PE, Daya NP. Mutans streptococci and other caries-associated acidogenic bacteria in five-year-old children in South Africa. Oral Microbiol Immunol. 1999;14(4):238-43; doi: 10.1034/j.1399-302x.1999.140407.x.

29. Dabrowska-Szponar M, Kubielas H. [Serologic and genetic analysis of group B streptococci isolated from adult patients]. Med Dosw Mikrobiol. 2006;58(4):283-9.
30. Spellerberg B. Pathogenesis of neonatal Streptococcusagalactiae infections. Microbes and infection. 2000;2(14):1733-42.
31. Identification of Streptococcus species, Enterococcus species and morphologically similar organisms. In: UK Standards for Microbiology Investigations (UK SMI ID 4). 2021.
32. Rosa-Fraile M, Spellerberg B. Reliable detection of group B Streptococcus in the clinical laboratory. Journal of clinical microbiology. 2017;55(9):2590-8.
33. Binghuai L, Yanli S, Shuchen Z, Fengxia Z, Dong L, Yanchao C. Use of MALDI-TOF mass spectrometry for rapid identification of group B Streptococcus on chromID Strepto B agar. International Journal of Infectious Diseases. 2014;27:44-8.
34. Wilkinson HW. CAMP-disk test for presumptive identification of group B streptococci. J Clin Microbiol. 1977;6(1):42-5.
35. Bae BH, Bottone EJ. Modified Christie–Atkins–Munch-Petersen (CAMP) test for direct identification of hemolytic and nonhemolytic group B streptococci on primary plating. Canadian journal of microbiology. 1980;26(4):539-42.
36. Winn WC. Koneman's color atlas and textbook of diagnostic microbiology. Lippincott williams & wilkins; 2006.
37. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P-E, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clinical Infectious Diseases. 2009;49(4):543-51.
38. Nabet C, Kone AK, Dia AK, Sylla M, Gautier M, Yattara M, et al. New assessment of Anopheles vector species identification using MALDI-TOF MS. Malaria Journal. 2021;20(1):33; doi: 10.1186/s12936-020-03557-2.
39. Lartigue M-F, Héry-Arnaud G, Haguenoer E, Domelier A-S, Schmit P-O, Van Der Mee-Marquet N, et al. Identification of Streptococcus agalactiae isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Journal of clinical microbiology. 2009;47(7):2284-7.
40. Lartigue M-F, Kostrzewa M, Salloum M, Haguenoer E, Héry-Arnaud G, Domelier A-S, et al. Rapid detection of “highly virulent” Group B Streptococcus ST-17 and emerging ST-1 clones by MALDI-TOF mass spectrometry. Journal of microbiological methods. 2011;86(2):262-5.
41. Rothen J, Pothier JF, Foucault F, Blom J, Nanayakkara D, Li C, et al. Subspecies typing of Streptococcus agalactiae based on ribosomal subunit protein mass variation by MALDI-TOF MS. Frontiers in microbiology. 2019;10:471.
42. Suwantarat N, Grundy M, Rubin M, Harris R, Miller JA, Romagnoli M, et al. Recognition of Streptococcus pseudoporcinus Colonization in Women as a Consequence of Using Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Group B

Streptococcus Identification. J Clin Microbiol. 2015;53(12):3926-30; doi: 10.1128/jcm.02363-15.

43. Buchan BW, Ledeboer NA. Emerging technologies for the clinical microbiology laboratory. Clin Microbiol Rev. 2014;27(4):783-822; doi: 10.1128/cmr.00003-14.
44. Slotved H-C, Hoffmann S. Evaluation of procedures for typing of group B Streptococcus: a retrospective study. PeerJ. 2017;5:e3105-e; doi: 10.7717/peerj.3105.
45. Kapatai G, Patel D, Efstratiou A, Chalker VJ. Comparison of molecular serotyping approaches of Streptococcus agalactiae from genomic sequences. BMC genomics. 2017;18(1):429; doi: 10.1186/s12864-017-3820-5.
46. 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.
47. Benson JA, Flores AE, Baker CJ, Hillier SL, Ferrieri P. Improved methods for typing nontypeable isolates of group B streptococci. International journal of medical microbiology. 2002;292(1):37-42.
48. Dutra VG, Alves VM, Olendzki AN, Dias CA, de Bastos AF, Santos GO, et al. Streptococcus agalactiae in Brazil: serotype distribution, virulence determinants and antimicrobial susceptibility. BMC infectious diseases. 2014;14(1):1-9.
49. Brigtsen AK, Dedi L, Melby KK, Holberg-Petersen M, Radtke A, Lyng RV, et al. Comparison of PCR and serotyping of Group B Streptococcus in pregnant women: the Oslo GBS-study. Journal of microbiological methods. 2015;108:31-5.
50. Poyart C, Tazi A, Réglier-Poupet H, Billoët A, Tavares N, Raymond J, et al. Multiplex PCR assay for rapid and accurate capsular typing of group B streptococci. Journal of clinical microbiology. 2007;45(6):1985-8.
51. Andrade PD, Russo JdS, Gouveia JB, Costa CRC, Oliveira KG, Gianetti M, et al. Molecular characterization of group B Streptococcus serotypes by multiplex polymerase chain reaction. MedicalExpress. 2017;4.
52. Reller LB, Weinstein M, Jorgensen JH, Ferraro MJ. Antimicrobial Susceptibility Testing: A Review of General Principles and Contemporary Practices. Clinical Infectious Diseases. 2009;49(11):1749-55; doi: 10.1086/647952.
53. Joseph NM, Sistla S, Dutta TK, Badhe AS, Rasitha D, Parija SC. Reliability of Kirby-Bauer disk diffusion method for detecting meropenem resistance among non-fermenting gram-negative bacilli. Indian J Pathol Microbiol. 2011;54(3):556-60; doi: 10.4103/0377-4929.85092.
54. Wayne P. Clinical and laboratory standards institute. Performance standards for antimicrobial susceptibility testing. 2011.
55. Nachnani S, Scuteri A, Newman MG, Avanessian AB, Lomeli SL. E-test: a new technique for antimicrobial susceptibility testing for periodontal microorganisms. J Periodontol. 1992;63(7):576-83; doi: 10.1902/jop.1992.63.7.576.

56. Somily AM, Babay HAH. Superiority of D-zone Testing Method over Standard Method to detect Rnducible Resistance in Gram Positive Bacteria: a Prospective Surveillance from a Teaching Hospital in Saudi Arabia. International journal of health sciences. 2008;2(2):8-16.
57. Kassim A, Omuse G, Premji Z, Revathi G. Comparison of Clinical Laboratory Standards Institute and European Committee on Antimicrobial Susceptibility Testing guidelines for the interpretation of antibiotic susceptibility at a University teaching hospital in Nairobi, Kenya: a cross-sectional study. Annals of clinical microbiology and antimicrobials. 2016;15:21-; doi: 10.1186/s12941-016-0135-3.
58. Van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clinical microbiology reviews. 2001;14(3):547-60.
59. Maslow JN, Mulligan ME, Arbeit RD. Molecular epidemiology: application of contemporary techniques to the typing of microorganisms. Clinical Infectious Diseases. 1993:153-62.
60. Schiffman MH, Bauer HM, Lorincz AT, Manos MM, Byrne JC, Glass AG, et al. Comparison of Southern blot hybridization and polymerase chain reaction methods for the detection of human papillomavirus DNA. Journal of Clinical Microbiology. 1991;29(3):573-7.
61. Takahashi S, Adderson EE, Nagano Y, Nagano N, Briesacher MR, Bohnsack JF. Identification of a highly encapsulated, genetically related group of invasive type III group B streptococci. Journal of Infectious Diseases. 1998;177(4):1116-9.
62. Tamura GS, Herndon M, Przekwas J, Rubens CE, Ferrieri P, Hillier SL. Analysis of Restriction Fragment Length Polymorphisms of the Insertion Sequence IS1381 in Group B Streptococci. The Journal of Infectious Diseases. 2000;181(1):364-8.
63. Lee PY, Costumbrado J, Hsu C-Y, Kim YH. Agarose gel electrophoresis for the separation of DNA fragments. Journal of visualized experiments : JoVE. 2012;(62):3923; doi: 10.3791/3923.
64. Fasola E, Livdahl C, Ferrieri P. Molecular analysis of multiple isolates of the major serotypes of group B streptococci. Journal of clinical microbiology. 1993;31(10):2616-20.
65. Gordillo M, Singh K, Baker C, Murray B. Typing of group B streptococci: comparison of pulsed-field gel electrophoresis and conventional electrophoresis. Journal of Clinical Microbiology. 1993;31(6):1430-4.
66. Zhao W, Zou W, Chen JJ. Topic modeling for cluster analysis of large biological and medical datasets. BMC bioinformatics. 2014;15 Suppl 11(Suppl 11):S11-S; doi: 10.1186/1471-2105-15-S11-S11.
67. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. Journal of clinical microbiology. 1995;33(9):2233.

68. Ferrieri P, Lynfield R, Creti R, Flores AE. Serotype IV and invasive group B Streptococcus disease in neonates, Minnesota, USA, 2000–2010. Emerging infectious diseases. 2013;19(4):551.
69. Roychowdhury T, Vishnoi A, Bhattacharya A. Next-Generation Anchor Based Phylogeny (NexABP): constructing phylogeny from next-generation sequencing data. Scientific reports. 2013;3:2634-; doi: 10.1038/srep02634.
70. Hansen SM, Uldbjerg N, Kilian M, Sorensen UB. Dynamics of Streptococcus agalactiae colonization in women during and after pregnancy and in their infants. J Clin Microbiol. 2004;42(1):83-9; doi: 10.1128/jcm.42.1.83-89.2004.
71. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences. 1998;95(6):3140-5.
72. Jolley KA, Maiden MC. Using MLST to study bacterial variation: prospects in the genomic era. Future microbiology. 2014;9(5):623-30.
73. Chin C-S, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, et al. The origin of the Haitian cholera outbreak strain. New England Journal of Medicine. 2011;364(1):33-42.
74. Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Henderson DK, Palmore TN, et al. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Science translational medicine. 2012;4(148):148ra16-ra16.
75. Posada D, Crandall KA, Holmes EC. Recombination in evolutionary genomics. Annual Review of Genetics. 2002;36(1):75-97.
76. Flores AR, Galloway-Peña J, Sahasrabhojane P, Saldaña M, Yao H, Su X, et al. Sequence type 1 group B Streptococcus, an emerging cause of invasive disease in adults, evolves by small genetic changes. Proceedings of the National Academy of Sciences. 2015;112(20):6431-6.
77. 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. Emerging infectious diseases. 2016;22(11):1941.
78. Lopes E, Fernandes T, Machado MP, Carriço JA, Melo-Cristino J, Ramirez M, et al. Increasing macrolide resistance among Streptococcus agalactiae causing invasive disease in non-pregnant adults was driven by a single capsular-transformed lineage, Portugal, 2009 to 2015. Euro Surveill. 2018;23(21); doi: 10.2807/1560-7917.es.2018.23.21.1700473.
79. Hauser E, Mellmann A, Semmler T, Stoeber H, Wieler LH, Karch H, et al. Phylogenetic and molecular analysis of food-borne shiga toxin-producing Escherichia coli. Applied and environmental microbiology. 2013;79(8):2731-40.
80. Raphael BH, Baker DJ, Nazarian E, Lapierre P, Bopp D, Kozak-Muiznieks NA, et al. Genomic resolution of outbreak-associated Legionella pneumophila serogroup 1 isolates from New York State. Applied and environmental microbiology. 2016;82(12):3582-90.

81. Kingry LC, Rowe LA, Respicio-Kingry LB, Beard CB, Schriefer ME, Petersen JM. Whole genome multilocus sequence typing as an epidemiologic tool for Yersinia pestis. Diagnostic microbiology and infectious disease. 2016;84(4):275-80.
82. Basic-Hammer N, Vogel V, Basset P, Blanc DS. Impact of recombination on genetic variability within Staphylococcus aureus clonal complexes. Infection, Genetics and Evolution. 2010;10(7):1117-23.
83. Lan R, Reeves PR, Octavia S. Population structure, origins and evolution of major Salmonella enterica clones. Infection, Genetics and Evolution. 2009;9(5):996-1005.
84. Turrientes M-C, González-Alba J-M, del Campo R, Baquero M-R, Cantón R, Baquero F, et al. Recombination blurs phylogenetic groups routine assignment in Escherichia coli: setting the record straight. PloS one. 2014;9(8):e105395.
85. Lin E, Zou S, Wang Y, Lee C-C, Chiu C-H, Feng Y. Phylogeny, recombination, and invasiveness of group B Streptococcus revealed by genomic comparisons of its global strains. European Journal of Clinical Microbiology & Infectious Diseases. 2021;40(3):581-90; doi: 10.1007/s10096-020-04067-4.
86. 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; doi: 10.1128/jcm.41.6.2530-2536.2003.
87. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. Journal of bacteriology. 2004;186(5):1518-30.
88. Francisco AP, Bugalho M, Ramirez M, Carriço JA. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. BMC bioinformatics. 2009;10(1):152.
89. Kong F, Gowan S, Martin D, James G, Gilbert GL. Serotype identification of group B streptococci by PCR and sequencing. Journal of clinical microbiology. 2002;40(1):216-26.
90. Martins E, Melo-Cristino J, Ramirez M. Reevaluating the serotype II capsular locus of Streptococcus agalactiae. Journal of clinical microbiology. 2007;45(10):3384-6.
91. Manning SD, Lacher DW, Davies HD, Foxman B, Whittam TS. DNA polymorphism and molecular subtyping of the capsular gene cluster of group B streptococcus. Journal of clinical microbiology. 2005;43(12):6113-6.
92. Kong F, Martin D, James G, Gilbert GL. Towards a genotyping system for Streptococcus agalactiae (group B streptococcus): use of mobile genetic elements in Australasian invasive isolates. Journal of medical microbiology. 2003;52(4):337-44.
93. Sutcliffe J, Grebe T, Tait-Kamradt A, Wondrack L. Detection of erythromycin-resistant determinants by PCR. Antimicrobial agents and chemotherapy. 1996;40(11):2562-6.

94. Land M, Hauser L, Jun S-R, Nookaew I, Leuze MR, Ahn T-H, et al. Insights from 20 years of bacterial genome sequencing. Functional & integrative genomics. 2015;15(2):141-61; doi: 10.1007/s10142-015-0433-4.

95. Fraser CM, Eisen JA, Nelson KE, Paulsen IT, Salzberg SL. The value of complete microbial genome sequencing (you get what you pay for). Journal of bacteriology. 2002;184(23):6403-5; doi: 10.1128/jb.184.23.6403-6405.2002.

96. Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, Msadek T, et al. Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease. Molecular microbiology. 2002;45(6):1499-513.

97. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proceedings of the National Academy of Sciences. 2005;102(39):13950-5.

98. Tettelin H, Masignani V, Cieslewicz MJ, Eisen JA, Peterson S, Wessels MR, et al. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae. Proceedings of the National Academy of Sciences. 2002;99(19):12391-6.

99. Rouli L, Merhej V, Fournier P-E, Raoult D. The bacterial pangenome as a new tool for analysing pathogenic bacteria. New microbes and new infections. 2015;7:72-85.

100. Libster R, Edwards KM, Levent F, Edwards MS, Rench MA, Castagnini LA, et al. Long-term outcomes of group B streptococcal meningitis. Pediatrics. 2012;130(1):e8-e15.

101. Hartley J, Li Y, Kunkel L, Crowcroft NS. The burden of infant group B streptococcal infections in Ontario: Analysis of administrative data to estimate the potential benefits of new vaccines. Human vaccines & immunotherapeutics. 2019;15(1):193-202.

102. Doran KS, Nizet V. Molecular pathogenesis of neonatal group B streptococcal infection: no longer in its infancy. Molecular microbiology. 2004;54(1):23-31.

103. Polin RA. Management of neonates with suspected or proven early-onset bacterial sepsis. Pediatrics. 2012;129(5):1006-15.

104. Martinez E, Mintegi S, Vilar B, Martinez MJ, Lopez A, Catediano E, et al. Prevalence and predictors of bacterial meningitis in young infants with fever without a source. The Pediatric infectious disease journal. 2015;34(5):494-8.

105. Lim DV, Kanarek KS, Peterson ME. Magnitude of colonization and sepsis by group B streptococci in newborn infants. Current Microbiology. 1982;7(2):99-101.

106. Boyer K, Gotoff S. Strategies for Chemoprophylaxis of GBS Early-Onset Infections1. In: Neonatal group B streptococcal infections. Karger Publishers; 1985. p. 267-80.

107. 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 Suppl 4:D20-6; doi: 10.1016/j.vaccine.2012.11.056.

108. Melin P. Neonatal group B streptococcal disease: from pathogenesis to preventive strategies. Clinical Microbiology and Infection. 2011;17(9):1294-303; doi: https://doi.org/10.1111/j.1469-0691.2011.03576.x.
109. Streptococcal early-onset disease in newborns: ACOG committee opinion, number 797. Obstet Gynecol, 135 2020;pp. e51-e72.
110. Grimwood K, Darlow B, Gosling I, Green R, Lennon D, Martin D, et al. Early-onset neonatal group B streptococcal infections in New Zealand 1998− 1999. Journal of paediatrics and child health. 2002;38(3):272-7.
111. Darlow BA, Voss L, Lennon DR, Grimwood K. Early-onset neonatal group B streptococcus sepsis following national risk-based prevention guidelines. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2016;56(1):69-74.
112. López Sastre JB, Fernández Colomer B, Coto Cotallo GD, Ramos Aparicio A. Trends in the epidemiology of neonatal sepsis of vertical transmission in the era of group B streptococcal prevention. Acta Paediatrica. 2005;94(4):451-7.
113. Hafner E, Sterniste W, Rosen A, Schuchter K, Plattner M, Asboth F, et al. Group B streptococci during pregnancy: a comparison of two screening and treatment protocols. American journal of obstetrics and gynecology. 1998;179(3):677-81.
114. Alhhazmi A, Hurteau D, Tyrrell GJ. Epidemiology of invasive group B streptococcal disease in Alberta, Canada, from 2003 to 2013. Journal of clinical microbiology. 2016;54(7):1774-81.
115. Lamagni TL, Keshishian C, Efstratiou A, Guy R, Henderson KL, Broughton K, et al. Emerging trends in the epidemiology of invasive group B streptococcal disease in England and Wales, 1991-2010. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;57(5):682-8; doi: 10.1093/cid/cit337.
116. Bekker V, Bijlsma MW, van de Beek D, Kuijpers TW, van der Ende A. Incidence of invasive group B streptococcal disease and pathogen genotype distribution in newborn babies in the Netherlands over 25 years: a nationwide surveillance study. The Lancet Infectious Diseases. 2014;14(11):1083-9.
117. Slotved HC, Hoffmann S. The Epidemiology of Invasive Group B Streptococcus in Denmark From 2005 to 2018. Front Public Health. 2020;8:40; doi: 10.3389/fpubh.2020.00040.
118. Horváth-Puhó E, Snoek L, van Kassel MN, Gonçalves BP, Chandna J, Procter SR, et al. Prematurity Modifies the Risk of Long-term Neurodevelopmental Impairments After Invasive Group B Streptococcus Infections During Infancy in Denmark and the Netherlands. Clinical Infectious Diseases. 2021;74(Supplement_1):S44-S53; doi: 10.1093/cid/ciab774.
119. Dauby N, Chamekh M, Melin P, Slogrove AL, Goetghebuer T. Increased risk of group B streptococcus invasive infection in HIV-exposed but uninfected infants: a review of the evidence and possible mechanisms. Frontiers in immunology. 2016;7:505.

120. O'Sullivan CP, Lamagni T, Patel D, Efstratiou A, Cunney R, Meehan M, et al. Group B streptococcal disease in UK and Irish infants younger than 90 days, 2014-15: a prospective surveillance study. Lancet Infect Dis. 2019;19(1):83-90; doi: 10.1016/s1473-3099(18)30555-3.

121. Cailes B, Kortsalioudaki C, Buttery J, Pattnayak S, Greenough A. Epidemiology of UK neonatal infections: the neonIN infection surveillance network. 2018;103(6):F547-f53; doi: 10.1136/archdischild-2017-313203.

122. 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; doi: 10.1093/cid/cix664.

123. Bingen E, Denamur E, Lambert-Zechovsky N, Aujard Y, Brahimi N, Geslin P, et al. Analysis of DNA restriction fragment length polymorphism extends the evidence for breast milk transmission in Streptococcus agalactiae late-onset neonatal infection. Journal of Infectious Diseases. 1992;165(3):569-73.

124. 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; doi: https://doi.org/10.1016/j.pedneo.2021.05.007.

125. Bartlett AW, Smith B, George CRR, McMullan B, Kesson A, Lahra MM, et al. Epidemiology of Late and Very Late Onset Group B Streptococcal Disease: Fifteen-Year Experience From Two Australian Tertiary Pediatric Facilities. The Pediatric Infectious Disease Journal. 2017;36(1):20-4; doi: 10.1097/inf.0000000000001345.

126. Edwards MS, Baker C. Group B streptococcal infections. GROUP. 2005;12:c0060.

127. Kim EJ, Oh KY, Kim MY, Seo YS, Shin J-H, Song YR, et al. Risk factors for group B streptococcus colonization among pregnant women in Korea. Epidemiology and health. 2011;33.

128. Terry R, Kelly F, Gauzer C, Jeitler M. Risk factors for maternal colonization with group B beta-hemolytic streptococci. The Journal of the American Osteopathic Association. 1999;99(11):571-3.

129. Regan JA, Klebanoff MA, Nugent RP. The epidemiology of group B streptococcal colonization in pregnancy. Vaginal Infections and Prematurity Study Group. Obstetrics and gynecology. 1991;77(4):604-10.

130. Krohn MA, Hillier SL, Baker CJ. Maternal peripartum complications associated with vaginal group B streptococci colonization. The Journal of infectious diseases. 1999;179(6):1410-5.

131. Stoll BJ, Schuchat A. Maternal carriage of group B streptococci in developing countries. The Pediatric infectious disease journal. 1998;17(6):499-503.

132. Skoff TH, Farley MM, Petit S, Craig AS, Schaffner W, Gershman K, et al. Increasing burden of invasive group B streptococcal disease in nonpregnant adults, 1990–2007. Clinical Infectious Diseases. 2009;49(1):85-92.
133. Farley MM, Strasbaugh LJ. Group B streptococcal disease in nonpregnant adults. Clinical Infectious Diseases. 2001;33(4):556-61.
134. Domingo P, Barquet N, Alvarez M, Coll P, Nava J, Garau J. Group B streptococcal meningitis in adults: report of twelve cases and review. Clinical infectious diseases. 1997;25(5):1180-7.
135. Sambola A, Miro J, Tornos M, Almirante B, Moreno-Torrico A, Gurgui M, et al. Streptococcus agalactiae infective endocarditis: analysis of 30 cases and review of the literature, 1962–1998. Clinical infectious diseases. 2002;34(12):1576-84.
136. Eickel V, Kahl B, Reinisch B, Dübbers A, Küster P, Brandt C, et al. Emergence of respiratory Streptococcus agalactiae isolates in cystic fibrosis patients. PLoS One. 2009;4(2):e4650.
137. Harrison LH, Ali A, Dwyer DM, Libonati JP, Reeves MW, Elliott JA, et al. Relapsing invasive group B streptococcal infection in adults. Annals of internal medicine. 1995;123(6):421-7.
138. Rodriguez-Granger J, Alvargonzalez J, Berardi A, Berner R, Kunze M, Hufnagel M, et al. Prevention of group B streptococcal neonatal disease revisited. The DEVANI European project. European journal of clinical microbiology & infectious diseases. 2012;31(9):2097-104.
139. Bliss SJ, Manning SD, Tallman P, Baker CJ, Pearlman MD, Marrs CF, et al. Group B Streptococcus colonization in male and nonpregnant female university students: a cross-sectional prevalence study. Clinical Infectious Diseases. 2002;34(2):184-90.
140. Edwards MS, Rench MA, Palazzi DL, Baker CJ. Group B streptococcal colonization and serotype-specific immunity in healthy elderly persons. Clinical Infectious Diseases. 2005;40(3):352-7.
141. Manning SD, Neighbors K, Tallman PA, Gillespie B, Marrs CF, Borchardt SM, et al. Prevalence of group B streptococcus colonization and potential for transmission by casual contact in healthy young men and women. Clinical Infectious Diseases. 2004;39(3):380-8.
142. Manning SD, Springman AC, Million AD, Milton NR, McNamara SE, Somsel PA, et al. Association of group B Streptococcus colonization and bovine exposure: a prospective cross-sectional cohort study. PLoS One. 2010;5(1):e8795.
143. Oliveira I, De Mattos M, Pinto T, Ferreira-Carvalho B, Benchetrit L, Whiting A, et al. Genetic relatedness between group B streptococci originating from bovine mastitis and a human group B Streptococcus type V cluster displaying an identical pulsed-field gel electrophoresis pattern. Clinical microbiology and infection. 2006;12(9):887-93.
144. Sukhnanand S, Dogan B, Ayodele MO, Zadoks RN, Craver MPJ, Dumas NB, et al. Molecular subtyping and characterization of bovine and human Streptococcus agalactiae isolates. Journal of clinical microbiology. 2005;43(3):1177-86.

145. Foxman B, Gillespie B, Manning S, Marrs CF. Risk factors for group B streptococcal colonization: potential for different transmission systems by capsular type. Annals of epidemiology. 2007;17(11):854-62.
146. Tan S, Lin Y, Foo K, Koh HF, Tow C, Zhang Y, et al. Group B Streptococcus Serotype III Sequence Type 283 Bacteremia Associated with Consumption of Raw Fish, Singapore. Emerging infectious diseases. 2016;22(11):1970-3; doi: 10.3201/eid2211.160210.
147. Le Doare K, O'Driscoll M, Turner K, Seedat F, Russell NJ, Seale AC, et al. Intrapartum antibiotic chemoprophylaxis policies for the prevention of group B streptococcal disease worldwide: systematic review. Clinical Infectious Diseases. 2017;65(suppl_2):S143-S51.
148. Harper I. The importance of group B streptococci as human pathogens in the British Isles. Journal of clinical pathology. 1971;24(5):438-41.
149. Schuchat A, Whitney CG, Zangwill K. Prevention of perinatal group B streptococcal disease: a public health perspective. 1996.
150. Schrag SJ, Zywicki S, Farley MM, Reingold AL, Harrison LH, Lefkowitz LB, et al. Group B streptococcal disease in the era of intrapartum antibiotic prophylaxis. New England Journal of Medicine. 2000;342(1):15-20.
151. Prevention CfDCa. Perinatal group B streptococcal disease after universal screening recommendations--United States, 2003-2005. MMWR Morbidity and mortality weekly report. 2007;56(28):701.
152. Phares CR, Lynfield R, Farley MM, Mohle-Boetani J, Harrison LH, Petit S, et al. Epidemiology of invasive group B streptococcal disease in the United States, 1999-2005. Jama. 2008;299(17):2056-65.
153. Verani JR, McGee L, Schrag SJ. Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010. MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports. 2010;59(Rr-10):1-36.
154. 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; doi: 10.1001/jamapediatrics.2018.4826.
155. Filkins L, Hauser JR, Robinson-Dunn B, Tibbetts R, Boyanton BL, Revell P. American Society for Microbiology Provides 2020 Guidelines for Detection and Identification of Group B Streptococcus. Journal of clinical microbiology. 2020.
156. Puopolo KM, Lynfield R, Cummings JJ, Diseases CoI. Management of infants at risk for group B streptococcal disease. Pediatrics. 2019;144(2).
157. Cromwell D, Joffey T, Van der Meulen J, Dhillon C, Hughes R, Murphy D. The prevention of early-onset neonatal group B streptococcal disease in UK obstetric units: an audit of reported

practice in England, Scotland, Wales and Northern Ireland. London: Royal College of Obstetricians and Gynaecologists. 2007.

158. Prevention of early-onset neonatal group B streptococcal disease: green-top guideline No. 36. BJOG. 2017;volume 124, pp. e280-e305.

159. Freitas FTdM, Romero GAS. Early-onset neonatal sepsis and the implementation of group B streptococcus prophylaxis in a Brazilian maternity hospital: a descriptive study. The Brazilian Journal of Infectious Diseases. 2017;21(1):92-7; doi: https://doi.org/10.1016/j.bjid.2016.09.013.

160. Pediatria SBd. http://www.sbp.com.br/publicacoes/. 2018.

161. Costa HdPF, BRITO A. Prevenção da doença perinatal pelo estreptococo do grupo B. Manual da Sociedade Brasileira de Pediatria. 2011:1-18.

162. Garland SM, Fliegner JR. Group B streptococcus (GBS) and neonatal infections: the case for intrapartum chemoprophylaxis. Australian and New Zealand Journal of Obstetrics and Gynaecology. 1991;31(2):119-22.

163. Braye K, Foureur M, de Waal K, Jones M, Putt E, Ferguson J. Group B streptococcal screening, intrapartum antibiotic prophylaxis, and neonatal early-onset infection rates in an Australian local health district: 2006-2016. Plos one. 2019;14(4):e0214295.

164. Maternal Group B Streptococcus (GBS) and minimisation of neonatal early-onset GBS sepsis. Minstry of Health. NSW government. 2016.

165. The Royal Australian and New Zealand College of Obstetricians and Gynaecologists (RANZCOG). Maternal Group B Streptococcus in pregnancy: screening and management .C-Obs 19. 2016.

166. Kurz E, Davis D. Routine culture-based screening versus risk-based management for the prevention of early-onset group B streptococcus disease in the neonate: a systematic review. JBI Evidence Synthesis. 2015;13(3):206-46.

167. Van Dyke MK, Phares CR, Lynfield R, Thomas AR, Arnold KE, Craig AS, et al. Evaluation of universal antenatal screening for group B streptococcus. New England Journal of Medicine. 2009;360(25):2626-36.

168. McGee L, Schrag S, Verani JR. Prevention of perinatal Group B streptococcal disease; revised guidelines from CDC, 2010. 2010.

169. Berardi A, Rossi C, Lugli L, Creti R, Bacchi Reggiani ML, Lanari M, et al. Group B streptococcus late-onset disease: 2003-2010. Pediatrics. 2013;131(2):e361-8; doi: 10.1542/peds.2012-1231.

170. Romain AS, Cohen R, Plainvert C, Joubrel C, Béchet S, Perret A, et al. Clinical and Laboratory Features of Group B Streptococcus Meningitis in Infants and Newborns: Study of 848 Cases in France, 2001-2014. Clin Infect Dis. 2018;66(6):857-64; doi: 10.1093/cid/cix896.

171. Moroi H, Kimura K, Kotani T, Tsuda H, Banno H, Jin W, et al. Isolation of group B Streptococcus with reduced β-lactam susceptibility from pregnant women. Emerging microbes & infections. 2019;8(1):2-7; doi: 10.1080/22221751.2018.1557987.
172. McCall SJ, Bunch KJ, Brocklehurst P, D'Arcy R, Hinshaw K, Kurinczuk JJ, et al. The incidence, characteristics, management and outcomes of anaphylaxis in pregnancy: a population-based descriptive study. BJOG: An International Journal of Obstetrics & Gynaecology. 2018;125(8):965-71.
173. Azad MB, Konya T, Persaud RR, Guttman DS, Chari RS, Field CJ, et al. Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study. Bjog. 2016;123(6):983-93; doi: 10.1111/1471-0528.13601.
174. Seedat F, Stinton C, Patterson J, Geppert J, Tan B, Robinson ER, et al. Adverse events in women and children who have received intrapartum antibiotic prophylaxis treatment: a systematic review. BMC pregnancy and childbirth. 2017;17(1):247-; doi: 10.1186/s12884-017-1432-3.
175. Bonofiglio L, Gagetti P, Gabarrot GG, Kaufman S, Mollerach M, Toresani I, et al. Susceptibility to β-lactams in β-hemolytic streptococci. Revista Argentina de microbiologia. 2018;50(4):431-5.
176. Cooper K, Abbott F, Gould I. Reduced penicillin susceptibility of group B Streptococcus: an assessment of emergence in Grampian, Scotland. British Journal of Biomedical Science. 2016;73(1):25-7.
177. Dahesh S, Hensler ME, Van Sorge NM, Gertz Jr RE, Schrag S, Nizet V, et al. Point mutation in the group B streptococcal pbp2x gene conferring decreased susceptibility to β-lactam antibiotics. Antimicrobial agents and chemotherapy. 2008;52(8):2915-8.
178. Kimura K, Suzuki S, Wachino J-i, Kurokawa H, Yamane K, Shibata N, et al. First molecular characterization of group B streptococci with reduced penicillin susceptibility. Antimicrobial agents and chemotherapy. 2008;52(8):2890-7.
179. Longtin J, Vermeiren C, Shahinas D, Tamber GS, McGeer A, Low DE, et al. Novel mutations in a patient isolate of Streptococcus agalactiae with reduced penicillin susceptibility emerging after long-term oral suppressive therapy. Antimicrob Agents Chemother. 2011;55(6):2983-5; doi: 10.1128/aac.01243-10.
180. 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. Journal of Antimicrobial Chemotherapy. 2015;70(10):2725-8; doi: 10.1093/jac/dkv203.
181. Murayama SY, Seki C, Sakata H, Sunaoshi K, Nakayama E, Iwata S, et al. Capsular type and antibiotic resistance in Streptococcus agalactiae isolates from patients, ranging from newborns

to the elderly, with invasive infections. Antimicrob Agents Chemother. 2009;53(6):2650-3; doi: 10.1128/aac.01716-08.

182. Kimura K, Nagano N, Nagano Y, Wachino J-i, Suzuki S, Shibayama K, et al. Predominance of sequence type 1 group with serotype VI among group B streptococci with reduced penicillin susceptibility identified in Japan. Journal of antimicrobial chemotherapy. 2011;66(11):2460-4.

183. Compain F, Hays C, Touak G, Dmytruk N, Trieu-Cuot P, Joubrel C, et al. Molecular characterization of Streptococcus agalactiae isolates harboring small erm(T)-carrying plasmids. Antimicrobial agents and chemotherapy. 2014;58(11):6928-30; doi: 10.1128/AAC.03855-14.

184. Goffin C, Ghuysen J-M. Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. Microbiology and molecular biology reviews. 1998;62(4):1079-93.

185. Nagano N, Nagano Y, Kimura K, Tamai K, Yanagisawa H, Arakawa Y. Genetic heterogeneity in pbp genes among clinically isolated group B Streptococci with reduced penicillin susceptibility. Antimicrob Agents Chemother. 2008;52(12):4258-67; doi: 10.3201/eid1904.121572.

186. Kimura K, Wachino J, Kurokawa H, Matsui M, Suzuki S, Yamane K, et al. High cephalosporin resistance due to amino acid substitutions in PBP1A and PBP2X in a clinical isolate of group B Streptococcus. J Antimicrob Chemother. 2013;68(7):1533-6; doi: 10.1093/jac/dkt060.

187. Nagano N, Nagano Y, Toyama M, Kimura K, Shibayama K, Arakawa Y. Penicillin-susceptible group B streptococcal clinical isolates with reduced cephalosporin susceptibility. J Clin Microbiol. 2014;52(9):3406-10; doi: 10.1128/jcm.01291-14.

188. Di Renzo GC, Melin P, Berardi A, Blennow M, Carbonell-Estrany X, Donzelli G, et al. Intrapartum GBS screening and antibiotic prophylaxis: a European consensus conference. The Journal of Maternal-Fetal & Neonatal Medicine. 2015;28(7):766-82.

189. Back EE, O'grady EJ, Back JD. High rates of perinatal group B Streptococcus clindamycin and erythromycin resistance in an upstate New York hospital. Antimicrobial agents and chemotherapy. 2012;56(2):739-42.

190. Gygax SE, Schuyler JA, Kimmel LE, Trama JP, Mordechai E, Adelson ME. Erythromycin and clindamycin resistance in group B streptococcal clinical isolates. Antimicrobial agents and chemotherapy. 2006;50(5):1875-7.

191. Metcalf BJ, Gertz RE, Jr., Gladstone RA, Walker H, Sherwood LK, Jackson D, et al. Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. Clin Microbiol Infect. 2016;22(1):60.e9-.e29; doi: 10.1016/j.cmi.2015.08.027.

192. Garza-Ramos G, Xiong L, Zhong P, Mankin A. Binding site of macrolide antibiotics on the ribosome: new resistance mutation identifies a specific interaction of ketolides with rRNA. Journal of bacteriology. 2001;183(23):6898-907; doi: 10.1128/JB.183.23.6898-6907.2001.

193. Leclercq R. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. Clin Infect Dis. 2002;34(4):482-92; doi: 10.1086/324626.

194. Siibak T, Peil L, Xiong L, Mankin A, Remme J, Tenson T. Erythromycin- and Chloramphenicol-Induced Ribosomal Assembly Defects Are Secondary Effects of Protein Synthesis Inhibition. Antimicrobial Agents and Chemotherapy. 2009;53(2):563-71; doi: 10.1128/aac.00870-08.

195. Leclercq R, Courvalin P. Resistance to macrolides and related antibiotics in Streptococcus pneumoniae. Antimicrobial agents and chemotherapy. 2002;46(9):2727-34; doi: 10.1128/aac.46.9.2727-2734.2002.

196. Hollenbeck BL, Rice LB. Intrinsic and acquired resistance mechanisms in enterococcus. Virulence. 2012;3(5):421-33; doi: 10.4161/viru.21282.

197. Seo YS, Srinivasan U, Oh K-Y, Shin J-H, Chae JD, Kim MY, et al. Changing molecular epidemiology of group B streptococcus in Korea. Journal of Korean medical science. 2010;25(6):817-23.

198. Malbruny B, Werno AM, Anderson TP, Murdoch DR, Leclercq R. A new phenotype of resistance to lincosamide and streptogramin A-type antibiotics in Streptococcus agalactiae in New Zealand. J Antimicrob Chemother. 2004;54(6):1040-4; doi: 10.1093/jac/dkh493.

199. Brzychczy-Wloch M, Gosiewski T, Bodaszewska M, Pabian W, Bulanda M, Kochan P, et al. Genetic characterization and diversity of Streptococcus agalactiae isolates with macrolide resistance. J Med Microbiol. 2010;59(Pt 7):780-6; doi: 10.1099/jmm.0.018176-0.

200. Figueira-Coelho J, Ramirez M, Salgado MJ, Melo-Cristino J. Streptococcus agalactiae in a large Portuguese teaching hospital: antimicrobial susceptibility, serotype distribution, and clonal analysis of macrolide-resistant isolates. Microb Drug Resist. 2004;10(1):31-6; doi: 10.1089/107662904323047772.

201. de Azavedo JC, McGavin M, Duncan C, Low DE, McGeer A. Prevalence and mechanisms of macrolide resistance in invasive and noninvasive group B streptococcus isolates from Ontario, Canada. Antimicrob Agents Chemother. 2001;45(12):3504-8; doi: 10.1128/aac.45.12.3504-3508.2001.

202. Castor ML, Whitney CG, Como-Sabetti K, Facklam RR, Ferrieri P, Bartkus JM, et al. Antibiotic resistance patterns in invasive group B streptococcal isolates. Infectious diseases in obstetrics and gynecology. 2008;2008.

203. Sadowy E, Matynia B, Hryniewicz W. Population structure, virulence factors and resistance determinants of invasive, non-invasive and colonizing Streptococcus agalactiae in Poland. Journal of Antimicrobial Chemotherapy. 2010;65(9):1907-14; doi: 10.1093/jac/dkq230.

204. Dipersio LP, Dipersio JR. Identification of an erm(T) gene in strains of inducibly clindamycin-resistant group B Streptococcus. Diagn Microbiol Infect Dis. 2007;57(2):189-93; doi: 10.1016/j.diagmicrobio.2006.07.013.

205. 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;6:20047.

206. Guo Y, Deng X, Liang Y, Zhang L, Zhao G-P, Zhou Y. The draft genomes and investigation of serotype distribution, antimicrobial resistance of group B Streptococcus strains isolated from urine in Suzhou, China. Annals of Clinical Microbiology and Antimicrobials. 2018;17(1):28; doi: 10.1186/s12941-018-0280-y.

207. Marimón JM, Valiente A, Ercibengoa M, García-Arenzana JM, Pérez-Trallero E. Erythromycin resistance and genetic elements carrying macrolide efflux genes in Streptococcus agalactiae. Antimicrobial agents and chemotherapy. 2005;49(12):5069-74; doi: 10.1128/AAC.49.12.5069-5074.2005.

208. von Both U, Ruess M, Mueller U, Fluegge K, Sander A, Berner R. A serotype V clone is predominant among erythromycin-resistant Streptococcus agalactiae isolates in a southwestern region of Germany. J Clin Microbiol. 2003;41(5):2166-9; doi: 10.1128/jcm.41.5.2166-2169.2003.

209. Yao J, Moellering R. Antibacterial Agents en Manual of Clinical Microbiology. Patrick Murray y col. 1999.

210. Ruppen C, Hemphill A, Sendi P. In vitro activity of gentamicin as an adjunct to penicillin against biofilm group B Streptococcus. Journal of Antimicrobial Chemotherapy. 2016;72(2):444-7; doi: 10.1093/jac/dkw447.

211. Menzies D, Benedetti A, Paydar A, Martin I, Royce S, Pai M, et al. Effect of duration and intermittency of rifampin on tuberculosis treatment outcomes: a systematic review and meta-analysis. PLoS Med. 2009;6(9):e1000146; doi: 10.1371/journal.pmed.1000146.

212. Sendi P, Furitsch M, Mauerer S, Florindo C, Kahl BC, Shabayek S, et al. Chromosomally and extrachromosomally mediated high-level gentamicin resistance in Streptococcus agalactiae. Antimicrobial agents and chemotherapy. 2016;60(3):1702-7.

213. Liddy H, Holliman R. Group B Streptococcus highly resistant to gentamicin. Journal of Antimicrobial Chemotherapy. 2002;50(1):142-3.

214. 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.

215. Bolukaoto JY, Monyama CM, Chukwu MO, Lekala SM, Nchabeleng M, Maloba MRB, et al. Antibiotic resistance of Streptococcus agalactiae isolated from pregnant women in Garankuwa, South Africa. BMC Res Notes. 2015;8:364-; doi: 10.1186/s13104-015-1328-0.

216. Versporten A, Bielicki J, Drapier N, Sharland M, Goossens H, Group AP, et al. The Worldwide Antibiotic Resistance and Prescribing in European Children (ARPEC) point prevalence survey: developing hospital-quality indicators of antibiotic prescribing for children. Journal of Antimicrobial Chemotherapy. 2016;71(4):1106-17.
217. Puopolo KM, Benitz WE, Zaoutis TE, Diseases CoI. Management of neonates born at≥ 35 0/7 weeks' gestation with suspected or proven early-onset bacterial sepsis. Pediatrics. 2018;142(6).
218. Ruppen C, Lupo A, Decosterd L, Sendi P. Is penicillin plus gentamicin synergistic against clinical group B Streptococcus isolates?: an in vitro study. Frontiers in microbiology. 2016;7:1680.
219. Galimand M, Sabtcheva S, Courvalin P, Lambert T. Worldwide disseminated armA aminoglycoside resistance methylase gene is borne by composite transposon Tn1548. Antimicrob Agents Chemother. 2005;49(7):2949-53; doi: 10.1128/aac.49.7.2949-2953.2005.
220. Doi Y, Arakawa Y. 16S ribosomal RNA methylation: emerging resistance mechanism against aminoglycosides. Clin Infect Dis. 2007;45(1):88-94; doi: 10.1086/518605.
221. Over U, Gür D, Unal S, Miller GH. The changing nature of aminoglycoside resistance mechanisms and prevalence of newly recognized resistance mechanisms in Turkey. Clin Microbiol Infect. 2001;7(9):470-8; doi: 10.1046/j.1198-743x.2001.00284.x.
222. Magnet S, Courvalin P, Lambert T. Resistance-nodulation-cell division-type efflux pump involved in aminoglycoside resistance in Acinetobacter baumannii strain BM4454. Antimicrob Agents Chemother. 2001;45(12):3375-80; doi: 10.1128/aac.45.12.3375-3380.2001.
223. Magnet S, Smith TA, Zheng R, Nordmann P, Blanchard JS. Aminoglycoside resistance resulting from tight drug binding to an altered aminoglycoside acetyltransferase. Antimicrob Agents Chemother. 2003;47(5):1577-83; doi: 10.1128/aac.47.5.1577-1583.2003.
224. Garneau-Tsodikova S, Labby KJ. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. MedChemComm. 2016;7(1):11-27; doi: 10.1039/C5MD00344J.
225. Ramirez MS, Tolmasky ME. Aminoglycoside modifying enzymes. Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy. 2010;13(6):151-71; doi: 10.1016/j.drup.2010.08.003.
226. Vetting MW, LP SdC, Yu M, Hegde SS, Magnet S, Roderick SL, et al. Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys. 2005;433(1):212-26; doi: 10.1016/j.abb.2004.09.003.
227. Li C, Sapugahawatte DN, Yang Y, Wong KT, Lo NWS, Ip M. Multidrug-Resistant Streptococcus agalactiae Strains Found in Human and Fish with High Penicillin and Cefotaxime Non-Susceptibilities. Microorganisms. 2020;8(7):1055; doi: 10.3390/microorganisms8071055.

228. Chopra I, Roberts M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev. 2001;65(2):232-60 ; second page, table of contents; doi: 10.1128/mmbr.65.2.232-260.2001.

229. De Mouy D, Cavallo J-D, Leclercq R, Fabre R. Antibiotic Susceptibility and Mechanisms of Erythromycin Resistance in Clinical Isolates of Streptococcus agalactiae: French Multicenter Study. Antimicrobial Agents and Chemotherapy. 2001;45(8):2400-2; doi: 10.1128/aac.45.8.2400-2402.2001.

230. Culebras E, Rodriguez-Avial I, Betriu C, Redondo M, Picazo JJ. Macrolide and tetracycline resistance and molecular relationships of clinical strains of Streptococcus agalactiae. Antimicrob Agents Chemother. 2002;46(5):1574-6; doi: 10.1128/aac.46.5.1574-1576.2002.

231. Poyart C, Jardy L, Quesne G, Berche P, Trieu-Cuot P. Genetic basis of antibiotic resistance in Streptococcus agalactiae strains isolated in a French hospital. Antimicrob Agents Chemother. 2003;47(2):794-7; doi: 10.1128/aac.47.2.794-797.2003.

232. Gherardi G, Imperi M, Baldassarri L, Pataracchia M, Alfarone G, Recchia S, et al. Molecular epidemiology and distribution of serotypes, surface proteins, and antibiotic resistance among group B streptococci in Italy. J Clin Microbiol. 2007;45(9):2909-16; doi: 10.1128/jcm.00999-07.

233. Schoeffler AJ, Berger JM. DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. Q Rev Biophys. 2008;41(1):41-101; doi: 10.1017/s003358350800468x.

234. Wu H-M, Janapatla RP, Ho Y-R, Hung K-H, Wu C-W, Yan J-J, et al. Emergence of fluoroquinolone resistance in group B streptococcal isolates in Taiwan. Antimicrobial agents and chemotherapy. 2008;52(5):1888-90.

235. Kawamura Y, Fujiwara H, Mishima N, Tanaka Y, Tanimoto A, Ikawa S, et al. First Streptococcus agalactiae isolates highly resistant to quinolones, with point mutations in gyrA and parC. Antimicrobial agents and chemotherapy. 2003;47(11):3605-9.

236. Woodford N. Biological counterstrike: antibiotic resistance mechanisms of Gram-positive cocci. Clin Microbiol Infect. 2005;11 Suppl 3:2-21; doi: 10.1111/j.1469-0691.2005.01140.x.

237. Wehbeh W, Rojas-Diaz R, Li X, Mariano N, Grenner L, Segal-Maurer S, et al. Fluoroquinolone-resistant Streptococcus agalactiae: epidemiology and mechanism of resistance. Antimicrobial agents and chemotherapy. 2005;49(6):2495-7; doi: 10.1128/AAC.49.6.2495-2497.2005.

238. Neemuchwala A, Teatero S, Patel SN, Fittipaldi N. Fluoroquinolone resistance among clonal complex 1 group B Streptococcus strains. Canadian Journal of Infectious Diseases and Medical Microbiology. 2016;2016.

239. Arias B, Kovacec V, Vigliarolo L, Suárez M, Tersigni C, Müller L, et al. Fluoroquinolone-resistant Streptococcus agalactiae invasive isolates recovered in Argentina. Microbial Drug Resistance. 2019;25(5):739-43.

240. Gao K, Gao C, Huang L, Guan X, Ji W, Chang C-Y, et al. Predominance of III/ST19 and Ib/ST10 Lineages With High Multidrug Resistance in Fluoroquinolone-Resistant Group B Streptococci Isolates in Which a New Integrative and Conjugative Element Was Identified. Frontiers in Microbiology. 2021;11:3436.

241. Trieu-Cuot P, de Cespedes G, Bentorcha F, Delbos F, Gaspar E, Horaud T. Study of heterogeneity of chloramphenicol acetyltransferase (CAT) genes in streptococci and enterococci by polymerase chain reaction: characterization of a new CAT determinant. Antimicrob Agents Chemother. 1993;37(12):2593-8; doi: 10.1128/aac.37.12.2593.

242. Simoni S, Vincenzi C, Brenciani A, Morroni G, Bagnarelli P, Giovanetti E, et al. Molecular characterization of Italian isolates of fluoroquinolone-resistant streptococcus agalactiae and relationships with chloramphenicol resistance. Microbial Drug Resistance. 2018;24(3):225-31.

243. Johri AK, Paoletti LC, Glaser P, Dua M, Sharma PK, Grandi G, et al. Group B Streptococcus: global incidence and vaccine development. Nature Reviews Microbiology. 2006;4(12):932.

244. Edwards MS. Group B streptococcal conjugate vaccine: a timely concept for which the time has come. Human vaccines. 2008;4(6):444-8.

245. Wessels MR, Rubens CE, Benedi VJ, Kasper DL. Definition of a bacterial virulence factor: sialylation of the group B streptococcal capsule. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(22):8983-7; doi: 10.1073/pnas.86.22.8983.

246. Kong F, Gowan S, Martin D, James G, Gilbert GL. Molecular profiles of group B streptococcal surface protein antigen genes: relationship to molecular serotypes. J Clin Microbiol. 2002;40(2):620-6; doi: 10.1128/jcm.40.2.620-626.2002.

247. Sinha A, Lieu TA, Paoletti LC, Weinstein MC, Platt R. The projected health benefits of maternal group B streptococcal vaccination in the era of chemoprophylaxis. Vaccine. 2005;23(24):3187-95.

248. Baker CJ, Carey VJ, Rench MA, Edwards MS, Hillier SL, Kasper DL, et al. Maternal antibody at delivery protects neonates from early onset group B streptococcal disease. The Journal of infectious diseases. 2014;209(5):781-8; doi: 10.1093/infdis/jit549.

249. Dangor Z, Kwatra G, Izu A, Adrian P, Cutland CL, Velaphi S, et al. Correlates of protection of serotype-specific capsular antibody and invasive Group B Streptococcus disease in South African infants. Vaccine. 2015;33(48):6793-9.

250. Fabbrini M, Rigat F, Rinaudo CD, Passalaqua 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.

251. Lin F-YC, Philips III JB, Azimi PH, Weisman LE, Clark P, Rhoads GG, et al. Level of maternal antibody required to protect neonates against early-onset disease caused by group B Streptococcus type Ia: a multicenter, seroepidemiology study. The Journal of infectious diseases. 2001;184(8):1022-8.

252. Lancefield RC, Freimer EH. Type-specific polysaccharide antigens of group B streptococci. Epidemiology & Infection. 1966;64(2):191-203.

253. Baker CJ. Immunization to prevent group B streptococcal disease: victories and vexations. The Journal of infectious diseases. 1990;161(5):917-21.

254. Kobayashi M, Vekemans J, Baker CJ, Ratner AJ, Le Doare K, Schrag SJ. Group B Streptococcus vaccine development: present status and future considerations, with emphasis on perspectives for low and middle income countries. F1000Research. 2016;5.

255. Bröker M. Potential protective immunogenicity of tetanus toxoid, diphtheria toxoid and Cross Reacting Material 197 (CRM197) when used as carrier proteins in glycoconjugates. Human vaccines & immunotherapeutics. 2016;12(3):664-7.

256. Baker CJ, Rench MA, McInnes P. Immunization of pregnant women with group B streptococcal type III capsular polysaccharide-tetanus toxoid conjugate vaccine. Vaccine. 2003;21(24):3468-72.

257. Baker CJ, Paoletti LC, Rench MA, Guttormsen H-K, Edwards MS, Kasper DL. Immune response of healthy women to 2 different group B streptococcal type V capsular polysaccharide-protein conjugate vaccines. The Journal of infectious diseases. 2004;189(6):1103-12.

258. Paoletti LC, Kasper DL. Glycoconjugate vaccines to prevent group B streptococcal infections. Expert opinion on biological therapy. 2003;3(6):975-84.

259. Baker CJ, Paoletti LC, Wessels MR, Guttormsen H-K, Rench MA, Hickman ME, et al. Safety and immunogenicity of capsular polysaccharide—tetanus toxoid conjugate vaccines for group B streptococcal types Ia and Ib. The Journal of infectious diseases. 1999;179(1):142-50.

260. Baker CJ, Paoletti LC, Rench MA, Guttormsen H-K, Carey VJ, Hickman ME, et al. Use of capsular polysaccharide—tetanus toxoid conjugate vaccine for type II group B streptococcus in healthy women. The Journal of infectious diseases. 2000;182(4):1129-38.

261. Kasper DL, Paoletti LC, Wessels MR, Guttormsen H-K, Carey VJ, Jennings HJ, et al. Immune response to type III group B streptococcal polysaccharide-tetanus toxoid conjugate vaccine. The Journal of clinical investigation. 1996;98(10):2308-14.

262. Madhi SA, Koen A, Cutland CL, Jose L, Govender N, Wittke F, et al. Antibody kinetics and response to routine vaccinations in infants born to women who received an investigational trivalent group B Streptococcus polysaccharide CRM197-conjugate vaccine during pregnancy. Clinical Infectious Diseases. 2017;65(11):1897-904.

263. 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. Clinical Infectious Diseases. 2017;65(suppl_2):S160-S72.

264. 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.

265. 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.

266. 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.

267. Aguiar S, Serrano I, Pinto F, Melo-Cristino J, Ramirez M, Pathogens PSGftSoR. Changes in Streptococcus pneumoniae serotypes causing invasive disease with non-universal vaccination coverage of the seven-valent conjugate vaccine. Clinical Microbiology and Infection. 2008;14(9):835-43.

268. Dangor Z, Cutland CL, Izu A, Kwatra G, Trenor S, Lala SG, et al. Temporal Changes in Invasive Group B Streptococcus Serotypes: Implications for Vaccine Development. PloS one. 2016;11(12):e0169101-e; doi: 10.1371/journal.pone.0169101.

269. 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.

270. Gravekamp C, Kasper DL, Paoletti LC, Madoff LC. Alpha C protein as a carrier for type III capsular polysaccharide and as a protective protein in group B streptococcal vaccines. Infection and immunity. 1999;67(5):2491-6.

271. Cheng Q, Debol S, Lam H, Eby R, Edwards L, Matsuka Y, et al. Immunization with C5a peptidase or peptidase-type III polysaccharide conjugate vaccines enhances clearance of group B streptococci from lungs of infected mice. Infection and immunity. 2002;70(11):6409-15.

272. Larsson C, Holmgren J, Lindahl G, Bergquist C. Intranasal immunization of mice with group B streptococcal protein rib and cholera toxin B subunit confers protection against lethal infection. Infection and immunity. 2004;72(2):1184-7.

273. Areschoug T, Stålhammar-Carlemalm M, Larsson C, Lindahl G. Group B streptococcal surface proteins as targets for protective antibodies: identification of two novel proteins in strains of serotype V. Infect Immun. 1999;67(12):6350-7.

274. Adli M. The CRISPR tool kit for genome editing and beyond. Nature communications. 2018;9(1):1911-; doi: 10.1038/s41467-018-04252-2.

275. Wang NY, Patras KA, Seo HS, Cavaco CK, Rösler B, Neely MN, et al. Group B streptococcal serine-rich repeat proteins promote interaction with fibrinogen and vaginal colonization. J Infect Dis. 2014;210(6):982-91; doi: 10.1093/infdis/jiu151.

276. Santillan DA, Rai KK, Santillan MK, Krishnamachari Y, Salem AK, Hunter SK. Efficacy of polymeric encapsulated C5a peptidase–based group B streptococcus vaccines in a murine model. American journal of obstetrics and gynecology. 2011;205(3):249. e1-. e8.

277. Bohnsack JF, Widjaja K, Ghazizadeh S, Rubens CE, Hillyard DR, Parker CJ, et al. A role for C5 and C5a-ase in the acute neutrophil response to group B streptococcal infections. J Infect Dis. 1997;175(4):847-55; doi: 10.1086/513981.

278. Santillan DA, Andracki ME, Hunter SK. Protective immunization in mice against group B streptococci using encapsulated C5a peptidase. American journal of obstetrics and gynecology. 2008;198(1):114. e1-. e6.

279. Lauer P, Rinaudo CD, Soriani M, Margarit I, Maione D, Rosini R, et al. Genome analysis reveals pili in Group B Streptococcus. Science. 2005;309(5731):105; doi: 10.1126/science.1111563.

280. Margarit I, Rinaudo CD, Galeotti CL, Maione D, Ghezzo C, Buttazzoni E, et al. Preventing bacterial infections with pilus-based vaccines: the group B streptococcus paradigm. J Infect Dis. 2009;199(1):108-15; doi: 10.1086/595564.

281. Shabayek S, Spellerberg B. Group B Streptococcal Colonization, Molecular Characteristics, and Epidemiology. Frontiers in microbiology. 2018;9:437-; doi: 10.3389/fmicb.2018.00437.

282. Cieslewicz MJ, Chaffin D, Glusman G, Kasper D, Madan A, Rodrigues S, et al. Structural and genetic diversity of group B streptococcus capsular polysaccharides. Infection and immunity. 2005;73(5):3096-103.

283. 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-; doi: 10.3389/fimmu.2018.00602.

284. Vornhagen J, Waldorf KMA, Rajagopal L. Perinatal group B streptococcal infections: virulence factors, immunity, and prevention strategies. Trends in microbiology. 2017;25(11):919-31.

285. Maisey HC, Doran KS, Nizet V. Recent advances in understanding the molecular basis of group B Streptococcus virulence. Expert reviews in molecular medicine. 2008;10:e27-e; doi: 10.1017/S1462399408000811.

286. Doran KS, Chang JC, Benoit VM, Eckmann L, Nizet V. Group B streptococcal beta-hemolysin/cytolysin promotes invasion of human lung epithelial cells and the release of interleukin-8. J Infect Dis. 2002;185(2):196-203; doi: 10.1086/338475.

287. Spellerberg B, Rozdzinski E, Martin S, Weber-Heynemann J, Schnitzler N, Lütticken R, et al. Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of Streptococcus agalactiae to human laminin. Infection and immunity. 1999;67(2):871-8; doi: 10.1128/IAI.67.2.871-878.1999.

288. Beckmann C, Waggoner JD, Harris TO, Tamura GS, Rubens CE. Identification of novel adhesins from Group B streptococci by use of phage display reveals that C5a peptidase mediates fibronectin binding. Infection and immunity. 2002;70(6):2869-76; doi: 10.1128/iai.70.6.2869-2876.2002.

289. Schubert A, Zakikhany K, Schreiner M, Frank R, Spellerberg B, Eikmanns BJ, et al. A fibrinogen receptor from group B Streptococcus interacts with fibrinogen by repetitive units with novel ligand binding sites. Mol Microbiol. 2002;46(2):557-69; doi: 10.1046/j.1365-2958.2002.03177.x.

290. Seo HS, Xiong YQ, Sullam PM. Role of the serine-rich surface glycoprotein Srr1 of Streptococcus agalactiae in the pathogenesis of infective endocarditis. PLoS One. 2013;8(5):e64204; doi: 10.1371/journal.pone.0064204.

291. Maisey HC, Hensler M, Nizet V, Doran KS. Group B streptococcal pilus proteins contribute to adherence to and invasion of brain microvascular endothelial cells. J Bacteriol. 2007;189(4):1464-7; doi: 10.1128/jb.01153-06.

292. 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.

293. Al Safadi R, Amor S, Hery-Arnaud G, Spellerberg B, Lanotte P, Mereghetti L, et al. Enhanced expression of lmb gene encoding laminin-binding protein in Streptococcus agalactiae strains harboring IS1548 in scpB-lmb intergenic region. PloS one. 2010;5(5):e10794-e; doi: 10.1371/journal.pone.0010794.

294. Franken C, Haase G, Brandt C, Weber-Heynemann J, Martin S, Lämmler C, et al. Horizontal gene transfer and host specificity of beta-haemolytic streptococci: the role of a putative composite transposon containing scpB and lmb. Mol Microbiol. 2001;41(4):925-35; doi: 10.1046/j.1365-2958.2001.02563.x.

295. Udo EE, Boswihi SS, Al-Sweih N. Genotypes and Virulence Genes in Group B Streptococcus Isolated in the Maternity Hospital, Kuwait. Medical Principles and Practice. 2013;22(5):453-7; doi: 10.1159/000349932.

296. Springman AC, Lacher DW, Wu G, Milton N, Whittam TS, Davies HD, et al. Selection, recombination, and virulence gene diversity among group B streptococcal genotypes. Journal of bacteriology. 2009;191(17):5419-27.

297. Bensing BA, Khedri Z, Deng L, Yu H, Prakobphol A, Fisher SJ, et al. Novel aspects of sialoglycan recognition by the Siglec-like domains of streptococcal SRR glycoproteins. Glycobiology. 2016;26(11):1222-34; doi: 10.1093/glycob/cww042.

298. Seo HS, Minasov G, Seepersaud R, Doran KS, Dubrovska I, Shuvalova L, et al. Characterization of fibrinogen binding by glycoproteins Srr1 and Srr2 of Streptococcus agalactiae. J Biol Chem. 2013;288(50):35982-96; doi: 10.1074/jbc.M113.513358.

299. Konto-Ghiorghi Y, Mairey E, Mallet A, Duménil G, Caliot E, Trieu-Cuot P, et al. Dual role for pilus in adherence to epithelial cells and biofilm formation in Streptococcus agalactiae. PLoS pathogens. 2009;5(5):e1000422-e; doi: 10.1371/journal.ppat.1000422.

300. 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. Mol Microbiol. 2006;61(1):126-41; doi: 10.1111/j.1365-2958.2006.05225.x.

301. Vengadesan K, Ma X, Dwivedi P, Ton-That H, Narayana SV. A model for group B Streptococcus pilus type 1: the structure of a 35-kDa C-terminal fragment of the major pilin GBS80. J Mol Biol. 2011;407(5):731-43; doi: 10.1016/j.jmb.2011.02.024.

302. Springman AC, Lacher DW, Waymire EA, Wengert SL, Singh P, Zadoks RN, et al. Pilus distribution among lineages of group B Streptococcus: an evolutionary and clinical perspective. BMC microbiology. 2014;14(1):159.

303. Maione D, Margarit I, Rinaudo CD, Masignani 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; doi: 10.1126/science.1109869.

304. Pezzicoli A, Santi I, Lauer P, Rosini R, Rinaudo D, Grandi G, et al. Pilus Backbone Contributes to Group B Streptococcus Paracellular Translocation through Epithelial Cells. The Journal of Infectious Diseases. 2008;198(6):890-8; doi: 10.1086/591182.

305. Chattopadhyay D, Carey AJ, Caliot E, Webb RI, Layton JR, Wang Y, et al. Phylogenetic lineage and pilus protein Spb1/SAN1518 affect opsonin-independent phagocytosis and intracellular survival of Group B Streptococcus. Microbes and infection. 2011;13(4):369-82; doi: 10.1016/j.micinf.2010.12.009.

306. Telford JL, Barocchi MA, Margarit I, Rappuoli R, Grandi G. Pili in gram-positive pathogens. Nat Rev Microbiol. 2006;4(7):509-19; doi: 10.1038/nrmicro1443.

307. Fieber C, Kovarik P. Responses of innate immune cells to group A Streptococcus. Frontiers in cellular and infection microbiology. 2014;4:140-; doi: 10.3389/fcimb.2014.00140.

308. Cheng Q, Stafslien D, Purushothaman SS, Cleary P. The group B streptococcal C5a peptidase is both a specific protease and an invasin. Infection and immunity. 2002;70(5):2408-13; doi: 10.1128/iai.70.5.2408-2413.2002.

309. Nanra JS, Buitrago SM, Crawford S, Ng J, Fink PS, Hawkins J, et al. Capsular polysaccharides are an important immune evasion mechanism for Staphylococcus aureus. Human vaccines & immunotherapeutics. 2013;9(3):480-7; doi: 10.4161/hv.23223.

310. Santi I, Scarselli M, Mariani M, Pezzicoli A, Masignani V, Taddei A, et al. BibA: a novel immunogenic bacterial adhesin contributing to group B Streptococcus survival in human blood. Mol Microbiol. 2007;63(3):754-67; doi: 10.1111/j.1365-2958.2006.05555.x.

311. Li J, Ji W, Gao K, Zhou H, Zhang L, Mu X, et al. Molecular characteristics of group B Streptococcus isolates from infants in southern mainland China. BMC Infectious Diseases. 2019;19(1):812; doi: 10.1186/s12879-019-4434-0.

312. Kekic D, Gajic I, Opavski N, Kojic M, Vukotic G, Smitran A, et al. Trends in molecular characteristics and antimicrobial resistance of group B streptococci: a multicenter study in Serbia, 2015–2020. Scientific Reports. 2021;11(1):540; doi: 10.1038/s41598-020-79354-3.

313. Teatero S, McGeer A, Low DE, Li A, Demczuk W, Martin I, et al. Characterization of invasive group B streptococcus strains from the greater Toronto area, Canada. J Clin Microbiol. 2014;52(5):1441-7; doi: 10.1128/jcm.03554-13.

314. Kardos S, Tóthpál A, Laub K, Kristóf K, Ostorházi E, Rozgonyi F, et al. High prevalence of group B streptococcus ST17 hypervirulent clone among non-pregnant patients from a Hungarian venereology clinic. BMC Infectious Diseases. 2019;19(1):1009; doi: 10.1186/s12879-019-4626-7.

315. Soriano M, Santi I, Taddei A, Rappuoli R, Grandi G, Telford JL. Group B Streptococcus Crosses Human Epithelial Cells by a Paracellular Route. The Journal of Infectious Diseases. 2006;193(2):241-50; doi: 10.1086/498982.

316. Wilkinson HW, Eagon RG. Type-specific antigens of group B type Ic streptococci. Infection and immunity. 1971;4(5):596-604; doi: 10.1128/IAI.4.5.596-604.1971.

317. Maeland JA, Afset JE, Lyng RV, Radtke A. Survey of immunological features of the alpha-like proteins of Streptococcus agalactiae. Clin Vaccine Immunol. 2015;22(2):153-9; doi: 10.1128/CVI.00643-14.

318. Bolduc G, Baron M, Gravekamp C, Lachenauer C, Madoff L. The alpha C protein mediates internalization of group B Streptococcus within human cervical epithelial cells. Cellular microbiology. 2002;4(11):751-8.

319. Gabrielsen C, Mæland JA, Lyng RV, Radtke A, Afset JE. Molecular characteristics of Streptococcus agalactiae strains deficient in alpha-like protein encoding genes. J Med Microbiol. 2017;66(1):26-33; doi: 10.1099/jmm.0.000412.

320. Michel JL, Madoff LC, Kling DE, Kasper DL, Ausubel FM. Cloned alpha and beta C-protein antigens of group B streptococci elicit protective immunity. Infection and immunity. 1991;59(6):2023-8.

321. Smith BL, Flores A, Dechaine J, Krepela J, Bergdall A, Ferrieri P. Gene encoding the group B streptococcal protein R4, its presence in clinical reference laboratory isolates & R4 protein pepsin sensitivity. Indian J Med Res. 2004;119 Suppl:213-20.

322. Lachenauer CS, Creti R, Michel JL, Madoff LC. Mosaicism in the alpha-like protein genes of group B streptococci. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(17):9630-5; doi: 10.1073/pnas.97.17.9630.

323. Maeland JA, Bevanger L, Lyng RV. Antigenic determinants of alpha-like proteins of Streptococcus agalactiae. Clinical and diagnostic laboratory immunology. 2004;11(6):1035-9; doi: 10.1128/CDLI.11.6.1035-1039.2004.

324. Creti R, Fabretti F, Orefici G, von Hunolstein C. Multiplex PCR assay for direct identification of group B streptococcal alpha-protein-like protein genes. J Clin Microbiol. 2004;42(3):1326-9; doi: 10.1128/jcm.42.3.1326-1329.2004.

325. Lindahl G, Stålhammar-Carlemalm M, Areschoug T. Surface proteins of Streptococcus agalactiae and related proteins in other bacterial pathogens. Clin Microbiol Rev. 2005;18(1):102-27; doi: 10.1128/cmr.18.1.102-127.2005.

326. Baron MJ, Bolduc GR, Goldberg MB, Aupérin TC, Madoff LC. Alpha C protein of group B Streptococcus binds host cell surface glycosaminoglycan and enters cells by an actin-dependent mechanism. The Journal of biological chemistry. 2004;279(23):24714-23; doi: 10.1074/jbc.M402164200.

327. Jerlström PG, Talay SR, Valentin-Weigand P, Timmis KN, Chhatwal GS. Identification of an immunoglobulin A binding motif located in the beta-antigen of the c protein complex of group B streptococci. Infection and immunity. 1996;64(7):2787-93; doi: 10.1128/IAI.64.7.2787-2793.1996.

328. 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. J Exp Med. 1993;177(6):1593-603; doi: 10.1084/jem.177.6.1593.

329. Song JY, Lim JH, Lim S, Yong Z, Seo HS. Progress toward a group B streptococcal vaccine. Human vaccines & immunotherapeutics. 2018;14(11):2669-81; doi: 10.1080/21645515.2018.1493326.

330. Persson E, Berg S, Bevanger L, Bergh K, Valsö-Lyng R, Trollfors B. Characterisation of invasive group B streptococci based on investigation of surface proteins and genes encoding surface proteins. Clin Microbiol Infect. 2008;14(1):66-73; doi: 10.1111/j.1469-0691.2007.01877.x.

331. Deng L, Kasper DL, Krick TP, Wessels MR. Characterization of the linkage between the type III capsular polysaccharide and the bacterial cell wall of group B Streptococcus. J Biol Chem. 2000;275(11):7497-504; doi: 10.1074/jbc.275.11.7497.

332. Paoletti LC, Kasper DL. Surface Structures of Group B Streptococcus Important in Human Immunity. Microbiol Spectr. 2019;7(2); doi: 10.1128/microbiolspec.GPP3-0001-2017.

333. Cieslewicz MJ, Kasper DL, Wang Y, Wessels MR. Functional analysis in type Ia group B Streptococcus of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. J Biol Chem. 2001;276(1):139-46; doi: 10.1074/jbc.M005702200.

334. Rubens CE, Wessels MR, Heggen LM, Kasper DL. Transposon mutagenesis of type III group B Streptococcus: correlation of capsule expression with virulence. Proceedings of the National Academy of Sciences of the United States of America. 1987;84(20):7208-12; doi: 10.1073/pnas.84.20.7208.

335. Brochet M, Couve E, Zouine M, Vallaeys T, Rusniok C, Lamy MC, et al. Genomic diversity and evolution within the species Streptococcus agalactiae. Microbes Infect. 2006;8(5):1227-43; doi: 10.1016/j.micinf.2005.11.010.

336. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27(7):1009-10; doi: 10.1093/bioinformatics/btr039.

337. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000;16(10):944-5.

338. 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; doi: 10.1128/jcm.01615-16.

339. 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.

340. Barcaite E, Bartusevicius A, Tameliene R, Kliucinskas M, Maleckiene L, Nadisauskiene R. Prevalence of maternal group B streptococcal colonisation in European countries. Acta obstetricia et gynecologica Scandinavica. 2008;87(3):260-71.

341. 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):336.

342. Hiller JE, McDonald HM, Darbyshire P, Crowther CA. Antenatal screening for Group B Streptococcus: a diagnostic cohort study. BMC pregnancy and childbirth. 2005;5(1):1-8.

343. Angstetra D, Ferguson J, Giles WB. Institution of universal screening for Group B streptococcus (GBS) from a risk management protocol results in reduction of early-onset GBS

disease in a tertiary obstetric unit. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2007;47(5):378-82.

344. Diedrick MJ, Flores AE, Hillier SL, Creti R, Ferrieri P. Clonal analysis of colonizing group B Streptococcus, serotype IV, an emerging pathogen in the United States. Journal of clinical microbiology. 2010;48(9):3100-4.

345. Edmond KM, Kortsalioudaki C, Scott S, Schrag SJ, Zaidi AK, Cousens S, et al. Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. Lancet. 2012;379(9815):547-56; doi: 10.1016/s0140-6736(11)61651-6.

346. Lachenauer CS, Kasper D, Shimada J, Ichiman Y, Ohtsuka H, Kaku M, et al. Serotypes VI and VIII predominate among group B streptococci isolated from pregnant Japanese women. The Journal of infectious diseases. 1999;179(4):1030-3.

347. Ueno H, Yamamoto Y, Yamamichi A, Kikuchi K, Kobori S, Miyazaki M. Characterization of group B streptococcus isolated from women in Saitama city, Japan. Japanese journal of infectious diseases. 2012;65(6):516-21.

348. Kimura K, Matsubara K, Yamamoto G, Shibayama K, Arakawa Y. Active screening of group B streptococci with reduced penicillin susceptibility and altered serotype distribution isolated from pregnant women in Kobe, Japan. Japanese journal of infectious diseases. 2013;66(2):158-60.

349. Dhanoa A, Karunakaran R, Puthucheary S. Serotype distribution and antibiotic susceptibility of group B streptococci in pregnant women. Epidemiology & Infection. 2010;138(7):979-81.

350. Eskandarian N, Neela V, Ismail Z, Puzi S, Hamat R, Desa M, et al. Group B streptococcal bacteremia in a major teaching hospital in Malaysia: a case series of eighteen patients. International Journal of Infectious Diseases. 2013;17(9):e777-e80.

351. Karunakaran R, Raja NS, Hafeez A, Puthucheary SD. Group B Streptococcus infection: epidemiology, serotypes, and antimicrobial susceptibility of selected isolates in the population beyond infancy (excluding females with genital tract-and pregnancy-related isolates) at the University Malaya Medical Centre, Kuala Lumpur. Jpn J Infect Dis. 2009;62(3):192-4.

352. Lee C-C, Hsu J-F, Prasad Janapatla R, Chen C-L, Zhou Y-L, Lien R, et al. 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):13525; doi: 10.1038/s41598-019-49977-2.

353. Eskandarian N, Ismail Z, Neela V, van Belkum A, Desa MNM, Amin Nordin S. Antimicrobial susceptibility profiles, serotype distribution and virulence determinants among invasive, non-invasive and colonizing Streptococcus agalactiae (group B streptococcus) from Malaysian patients. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2015;34(3):579-84; doi: 10.1007/s10096-014-2265-x.

354. Ho Y-R, Li C-M, Su H-P, Wu J-H, Tseng Y-C, Lin Y-J, et al. Variation in the number of tandem repeats and profile of surface protein genes among invasive group B streptococci correlates with patient age. Journal of clinical microbiology. 2007;45(5):1634-6.

355. Tien N, Ho C-M, Lin H-J, Shih M-C, Ho M-W, Lin H-C, et al. Multilocus sequence typing of invasive group B Streptococcus in central area of Taiwan. Journal of Microbiology, Immunology and Infection. 2011;44(6):430-4.

356. Wang YH, Su LH, Hou JN, Yang TH, Lin TY, Chu C, et al. Group B streptococcal disease in nonpregnant patients: emergence of highly resistant strains of serotype Ib in Taiwan in 2006 to 2008. J Clin Microbiol. 2010;48(7):2571-4; doi: 10.1128/jb.01130-09.

357. Wong SS, Tsui K, Liu Q-D, Lin L-C, Tsai CR, Chen L-C, et al. Serotypes, surface proteins, and clinical syndromes of invasive Group B streptococcal infections in northern Taiwan, 1998–2009. Journal of Microbiology, Immunology and Infection. 2011;44(1):8-14.

358. Lin H-C, Chen C-J, Chiang K-H, Yen T-Y, Ho C-M, Hwang K-P, et al. Clonal dissemination of invasive and colonizing clonal complex 1 of serotype VI group B Streptococcus in central Taiwan. Journal of Microbiology, Immunology and Infection. 2016;49(6):902-9.

359. Hoshina K, Suzuki Y, Nishida H, Kaneko K, Matsuda S, Kobayashi M, et al. Trend of neonatal group B streptococcal infection during the last 15 years. Pediatrics international. 2002;44(6):641-6.

360. Ekelund K, Slotved H-C, Nielsen HU, Kaltoft MS, Konradsen HB. Emergence of invasive serotype VIII group B streptococcal infections in Denmark. Journal of clinical microbiology. 2003;41(9):4442-4.

361. Martins ER, Florindo C, Martins F, Aldir I, Borrego MJ, Brum L, et al. Streptococcus agalactiae serotype Ib as an agent of meningitis in two adult nonpregnant women. J Clin Microbiol. 2007;45(11):3850-2; doi: 10.1128/jcm.01358-07.

362. 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):e00728-20.

363. 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.

364. Martins ER, Pessanha MA, Ramirez M, Melo-Cristino J. Analysis of group B streptococcal isolates from infants and pregnant women in Portugal revealing two lineages with enhanced invasiveness. J Clin Microbiol. 2007;45(10):3224-9; doi: 10.1128/jcm.01182-07.

365. Haenni M, Saras E, Bertin S, Leblond P, Madec J-Y, Payot S. Diversity and mobility of integrative and conjugative elements in bovine isolates of Streptococcus agalactiae, S. dysgalactiae subsp. dysgalactiae, and S. uberis. Applied and environmental microbiology. 2010;76(24):7957-65.

366. Rato MG, Bexiga R, Florindo C, Cavaco LM, Vilela CL, Santos-Sanches I. Antimicrobial resistance and molecular epidemiology of streptococci from bovine mastitis. Veterinary microbiology. 2013;161(3-4):286-94.

367. Evans JJ, Bohnsack JF, Klesius PH, Whiting AA, Garcia JC, Shoemaker CA, et al. Phylogenetic relationships among Streptococcus agalactiae isolated from piscine, dolphin, bovine and human sources: a dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. Journal of medical microbiology. 2008;57(11):1369-76.

368. Delannoy CM, Crumlish M, Fontaine MC, Pollock J, Foster G, Dagleish MP, et al. Human Streptococcus agalactiae strains in aquatic mammals and fish. BMC microbiology. 2013;13(1):41.

369. Li L, Wang R, Huang Y, Huang T, Luo F, Huang W, et al. High Incidence of Pathogenic Streptococcus agalactiae ST485 Strain in Pregnant/Puerperal Women and Isolation of Hyper-Virulent Human CC67 Strain. Frontiers in Microbiology. 2018;9(50); doi: 10.3389/fmicb.2018.00050.

370. Navarro-Torné A, Curcio D, Moïsi JC, Jodar L. Burden of invasive group B Streptococcus disease in non-pregnant adults: A systematic review and meta-analysis. PloS one. 2021;16(9):e0258030-e; doi: 10.1371/journal.pone.0258030.

371. 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.

372. do Nascimento CS, Dos Santos NF, Ferreira RC, Taddei CR. Streptococcus agalactiae in pregnant women in Brazil: prevalence, serotypes, and antibiotic resistance. Brazilian Journal of Microbiology. 2019;50(4):943-52.

373. Richtmann R, Silva CdA, Rodrigues TT, Costa HP, Baltieri SR, Kusano E. Evaluation of a strategy to prevent early neonatal group B streptococcus infection: A prospective cohort study. Journal of Pediatric Infectious Diseases. 2012;7(04):145-9.

374. Simoes JA, Alves VMN, Fracalanzza SEL, Camargo RPSd, Mathias L, Milanez HMBP, et al. Phenotypical characteristics of group B streptococcus in parturients. Brazilian Journal of Infectious Diseases. 2007;11:261-6.

375. Soares GCT, Alviano DS, Santos GdS, Alviano CS, Mattos-Guaraldi AL, Nagao PE. Prevalence of Group B Streptococcus serotypes III and V in pregnant women of Rio de Janeiro, Brazil. Brazilian Journal of Microbiology. 2013;44:869-72.

376. Palmeiro JK, Dalla-Costa LM, Fracalanzza SE, Botelho AC, da Silva Nogueira K, Scheffer MC, et al. Phenotypic and genotypic characterization of group B streptococcal isolates in southern Brazil. Journal of clinical microbiology. 2010;48(12):4397-403.

377. Benchetrit L, Fracalanzza S, Peregrino H, Camelo A, Sanches L. Carriage of Streptococcus agalactiae in women and neonates and distribution of serological types: a study in Brazil. Journal of clinical microbiology. 1982;15(5):787-90.

378. Corrêa ABdA, Silva LGd, Pinto TdCA, Oliveira ICMd, Fernandes FG, Costa NSd, et al. The genetic diversity and phenotypic characterisation of Streptococcus agalactiae isolates from Rio de Janeiro, Brazil. Memórias do Instituto Oswaldo Cruz. 2011;106(8):1002-6.

379. Pinto TCA, Costa NS, Souza ARV, Silva LGd, Corrêa ABdA, Fernandes FG, et al. Distribution of serotypes and evaluation of antimicrobial susceptibility among human and bovine Streptococcus agalactiae strains isolated in Brazil between 1980 and 2006. Brazilian Journal of Infectious Diseases. 2013;17(2):131-6.

380. Duarte RS, Bellei BC, Miranda OP, Brito MAVP, Teixeira LM. Distribution of Antimicrobial Resistance and Virulence-Related Genes among Brazilian Group B Streptococci Recovered from Bovine and Human Sources. Antimicrobial Agents and Chemotherapy. 2005;49(1):97-103; doi: 10.1128/aac.49.1.97-103.2005.

381. Singh T, Barnes EH, Isaacs D. Early-onset neonatal infections in Australia and New Zealand, 2002–2012. Archives of Disease in Childhood-Fetal and Neonatal Edition. 2019;104(3):F248-F52.

382. Ko DW, Zurynski Y, Gilbert GL, Group GS. Group B streptococcal disease and genotypes in A ustralian infants. Journal of paediatrics and child health. 2015;51(8):808-14.

383. Zhao Z, Kong F, Zeng X, Gidding H, Morgan J, Gilbert G. Distribution of genotypes and antibiotic resistance genes among invasive Streptococcus agalactiae (group B streptococcus) isolates from Australasian patients belonging to different age groups. Clinical Microbiology and Infection. 2008;14(3):260-7.

384. Furfaro LL, Nathan EA, Chang BJ, Payne MS. Group B streptococcus prevalence, serotype distribution and colonization dynamics in Western Australian pregnant women. J Med Microbiol. 2019;68(5):728-40; doi: 10.1099/jmm.0.000980.

385. Langley G, Schaffner W, Farley MM, Lynfield R, Bennett NM, Reingold A, et al. Twenty Years of Active Bacterial Core Surveillance. Emerging infectious diseases. 2015;21(9):1520-8; doi: 10.3201/eid2109.141333.

386. Laboratory- and population-based surveillance system for invasive bacterial pathogens of public health importance. https://www.cdc.gov/abcs/reports-findings/surv-reports.html. (1997-2019). Accessed January 2022.

387. Prevention of perinatal group B streptococcal disease: a public health perspective. Centers for Disease Control and Prevention. 1996;MMWR Recomm Rep, 45, pp. 1-24.

388. Control CfD, Prevention. Prevention of perinatal group B streptococcal disease. MMWR Morbid Mortal Wkly Rep. 2002;51(RR 11).

389. 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. Clin Infect Dis. 2021;72(6):1004-13; doi: 10.1093/cid/ciaa151.

390. Bryant J, Chewapreecha C, Bentley SD. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. Future Microbiol. 2012;7(11):1283-96; doi: 10.2217/fmb.12.108.

391. Schürch AC, Arredondo-Alonso S, Willems RJL, Goering RV. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. Clinical Microbiology and Infection. 2018;24(4):350-4; doi: https://doi.org/10.1016/j.cmi.2017.12.016.

392. Mira A, Ochman H, Moran NA. Deletional bias and the evolution of bacterial genomes. Trends Genet. 2001;17(10):589-96; doi: 10.1016/s0168-9525(01)02447-7.

393. Moran NA. Microbial minimalism: genome reduction in bacterial pathogens. Cell. 2002;108(5):583-6; doi: 10.1016/s0092-8674(02)00665-7.

394. Toft C, Andersson SG. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet. 2010;11(7):465-75; doi: 10.1038/nrg2798.

395. 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.

396. 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.

397. Dogan B, Schukken Y, Santisteban C, Boor KJ. Distribution of serotypes and antimicrobial resistance genes among Streptococcus agalactiae isolates from bovine and human hosts. Journal of Clinical Microbiology. 2005;43(12):5899-906.

398. Ramaswamy SV, Ferrieri P, Madoff LC, Flores AE, Kumar N, Tettelin H, et al. Identification of novel cps locus polymorphisms in nontypable group B Streptococcus. Journal of medical microbiology. 2006;55(6):775-83.

399. Creti R, Imperi M, Pataracchia M, Alfarone G, Recchia S, Baldassarri L. Identification and molecular characterization of a S. agalactiae strain lacking the capsular locus. European journal of clinical microbiology & infectious diseases. 2012;31(3):233-5.

400. 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.

401. Chaffin DO, Beres SB, Yim HH, Rubens CE. The serotype of type Ia and III group B streptococci is determined by the polymerase gene within the polycistronic capsule operon. Journal of bacteriology. 2000;182(16):4466-77.

402. Almeida A, Rosinski-Chupin I, Plainvert C, Douarre P-E, Borrego MJ, Poyart C, et al. Parallel evolution of group B Streptococcus hypervirulent clonal complex 17 unveils new pathoadaptive mutations. MSystems. 2017;2(5):e00074-17.

403. Kapatai G, Sheppard CL, Al-Shahib A, Litt DJ, Underwood AP, Harrison TG, et al. Whole genome sequencing of Streptococcus pneumoniae: development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. PeerJ. 2016;4:e2477-e; doi: 10.7717/peerj.2477.

404. Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nature reviews Genetics. 2011;13(1):36-46; doi: 10.1038/nrg3117.

405. Alkan C, Sajjadian S, Eichler EE. Limitations of next-generation genome sequence assembly. Nat Methods. 2011;8(1):61-5; doi: 10.1038/nmeth.1527.

406. Phillippy AM, Schatz MC, Pop M. Genome assembly forensics: finding the elusive mis-assembly. Genome Biol. 2008;9(3):R55; doi: 10.1186/gb-2008-9-3-r55.

407. Wetzel J, Kingsford C, Pop M. Assessing the benefits of using mate-pairs to resolve repeats in de novo short-read prokaryotic assemblies. BMC Bioinformatics. 2011;12:95; doi: 10.1186/1471-2105-12-95.

408. Thompson LH, Schild D. Homologous recombinational repair of DNA ensures mammalian chromosome stability. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2001;477(1):131-53; doi: https://doi.org/10.1016/S0027-5107(01)00115-4.

409. Johnson CM, Grossman AD. Integrative and Conjugative Elements (ICEs): What They Do and How They Work. Annual review of genetics. 2015;49:577-601; doi: 10.1146/annurev-genet-112414-055018.

410. Soltani J, Poorabbas B, Miri N, Mardaneh J. Health care associated infections, antibiotic resistance and clinical outcome: A surveillance study from Sanandaj, Iran. World journal of clinical cases. 2016;4(3):63-70; doi: 10.12998/wjcc.v4.i3.63.

411. Partridge SR, Kwong SM, Firth N, Jensen SO. Mobile genetic elements associated with antimicrobial resistance. Clinical microbiology reviews. 2018;31(4).

412. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiology Reviews. 2014;38(5):865-91; doi: 10.1111/1574-6976.12067.

413. Domingues S, da Silva GJ, Nielsen KM. Integrons: Vehicles and pathways for horizontal dissemination in bacteria. Mobile genetic elements. 2012;2(5):211-23; doi: 10.4161/mge.22967.

414. Chandler M, Mahillon J. Insertion sequences revisited. In: Mobile DNA II. American Society of Microbiology; 2002. p. 305-66.

415. Hickman AB, Dyda F. Mechanisms of DNA transposition. Mobile DNA III. 2015:529-53.
416. Siguier P, Gourbeyre E, Varani A, Ton-Hoang B, Chandler M. Everyman's guide to bacterial insertion sequences. Mobile DNA III. 2015:555-90.
417. Siguier P, Gagnevin L, Chandler M. The new IS1595 family, its relation to IS1 and the frontier between insertion sequences and transposons. Research in Microbiology. 2009;160(3):232-41; doi: https://doi.org/10.1016/j.resmic.2009.02.003.
418. Siguier P, Gourbeyre E, Chandler M. Known knowns, known unknowns and unknown unknowns in prokaryotic transposition. Current Opinion in Microbiology. 2017;38:171-80; doi: https://doi.org/10.1016/j.mib.2017.06.005.
419. Nicolas E, Lambin M, Dandoy D, Galloy C, Nguyen N, Oger CA, et al. The Tn3-family of replicative transposons. Mobile DNA III. 2015:693-726.
420. Partridge SR, Hall RM. Evolution of transposons containing blaTEM genes. Antimicrobial agents and chemotherapy. 2005;49(3):1267-8.
421. Gillings MR. Integrons: past, present, and future. Microbiology and molecular biology reviews. 2014;78(2):257-77.
422. Akrami F, Rajabnia M, Pournajaf A. Resistance integrons; A mini review. Caspian Journal of Internal Medicine. 2019;10(4):370.
423. Boucher Y, Labbate M, Koenig JE, Stokes H. Integrons: mobilizable platforms that promote genetic diversity in bacteria. Trends in microbiology. 2007;15(7):301-9.
424. Cambray G, Guerout A-M, Mazel D. Integrons. Annual review of genetics. 2010;44:141-66.
425. Messier N, Roy PH. Integron integrases possess a unique additional domain necessary for activity. Journal of bacteriology. 2001;183(22):6699-706.
426. Partridge SR, Recchia GD, Scaramuzzi C, Collis CM, Stokes H, Hall RM. Definition of the attI1 site of class 1 integrons. Microbiology. 2000;146(11):2855-64.
427. Collis CM, Hall RM. Expression of antibiotic resistance genes in the integrated cassettes of integrons. Antimicrobial agents and chemotherapy. 1995;39(1):155-62.
428. Lévesque C, Brassard S, Lapointe J, Roy PH. Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integron. Gene. 1994;142(1):49-54.
429. Hall RM, Brookes DE, Stokes H. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Molecular microbiology. 1991;5(8):1941-59.
430. Rowe-Magnus DA, Guérout A-M, Mazel D. Super-integrons. Research in microbiology. 1999;150(9-10):641-51.
431. Stokes H, O'gorman D, Recchia GD, Parsekhian M, Hall RM. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. Molecular microbiology. 1997;26(4):731-45.

432. Hocquet D, Llanes C, Thouverez M, Kulasekara HD, Bertrand X, Plésiat P, et al. Evidence for induction of integron-based antibiotic resistance by the SOS response in a clinical setting. PLoS pathogens. 2012;8(6):e1002778.

433. Guglielmini J, Quintais L, Garcillán-Barcia MP, de la Cruz F, Rocha EP. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS Genet. 2011;7(8):e1002222; doi: 10.1371/journal.pgen.1002222.

434. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005;3(9):722-32; doi: 10.1038/nrmicro1235.

435. Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EP, de la Cruz F. Mobility of plasmids. Microbiol Mol Biol Rev. 2010;74(3):434-52; doi: 10.1128/mmbr.00020-10.

436. Horodniceanu T, Bouanchaud D, Bieth G, Chabbert Y. R plasmids in Streptococcus agalactiae (group B). Antimicrobial agents and chemotherapy. 1976;10(5):795-801.

437. Woodbury RL, Klammer KA, Xiong Y, Bailiff T, Glennen A, Bartkus JM, et al. Plasmid-borne erm (T) from invasive, macrolide-resistant Streptococcus pyogenes strains. Antimicrobial agents and chemotherapy. 2008;52(3):1140-3.

438. DiPersio LP, DiPersio JR, Beach JA, Loudon AM, Fuchs AM. Identification and characterization of plasmid-borne erm (T) macrolide resistance in group B and group A Streptococcus. Diagnostic microbiology and infectious disease. 2011;71(3):217-23.

439. Delavat F, Miyazaki R, Carraro N, Pradervand N, van der Meer JR. The hidden life of integrative and conjugative elements. FEMS microbiology reviews. 2017;41(4):512-37.

440. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. Nature Reviews Microbiology. 2004;2(5):414-24.

441. Novick RP, Christie GE, Penadés JR. The phage-related chromosomal islands of Gram-positive bacteria. Nature Reviews Microbiology. 2010;8(8):541-51.

442. Bellanger X, Payot S, Leblond-Bourget N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiology Reviews. 2014;38(4):720-60; doi: 10.1111/1574-6976.12058.

443. Ghinet MG, Bordeleau E, Beaudin J, Brzezinski R, Roy S, Burrus V. Uncovering the prevalence and diversity of integrating conjugative elements in actinobacteria. PLoS One. 2011;6(11):e27846; doi: 10.1371/journal.pone.0027846.

444. Ambroset C, Coluzzi C, Guédon G, Devignes M-D, Loux V, Lacroix T, et al. New Insights into the Classification and Integration Specificity of Streptococcus Integrative Conjugative Elements through Extensive Genome Exploration. Frontiers in Microbiology. 2016;6(1483); doi: 10.3389/fmicb.2015.01483.

445. Roberts AP, Chandler M, Courvalin P, Guédon G, Mullany P, Pembroke T, et al. Revised nomenclature for transposable genetic elements. Plasmid. 2008;60(3):167-73.

446. Tomich P, An F, Damle S, Clewell D. Plasmid-related transmissibility and multiple drug resistance in Streptococcus faecalis subsp. zymogenes strain DS16. Antimicrobial Agents and Chemotherapy. 1979;15(6):828-30.

447. Rice LB. Tn916 family conjugative transposons and dissemination of antimicrobial resistance determinants. Antimicrobial agents and chemotherapy. 1998;42(8):1871-7; doi: 10.1128/AAC.42.8.1871.

448. Roberts AP, Mullany P. Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. FEMS Microbiology Reviews. 2011;35(5):856-71; doi: 10.1111/j.1574-6976.2011.00283.x.

449. Varaldo PE, Montanari MP, Giovanetti E. Genetic elements responsible for erythromycin resistance in streptococci. Antimicrobial agents and chemotherapy. 2009;53(2):343-53.

450. Puopolo KM, Klinzing DC, Lin MP, Yesucevitz DL, Cieslewicz MJ. A composite transposon associated with erythromycin and clindamycin resistance in group B Streptococcus. Journal of medical microbiology. 2007;56(Pt 7):947-55; doi: 10.1099/jmm.0.47131-0.

451. Brenciani A, Bacciaglia A, Vecchi M, Vitali LA, Varaldo PE, Giovanetti E. Genetic elements carrying erm (B) in Streptococcus pyogenes and association with tet (M) tetracycline resistance gene. Antimicrobial agents and chemotherapy. 2007;51(4):1209-16.

452. Clewell D, An FY, White BA, Gawron-Burke C. Streptococcus faecalis sex pheromone (cAM373) also produced by Staphylococcus aureus and identification of a conjugative transposon (Tn918). Journal of bacteriology. 1985;162(3):1212-20.

453. Christie P, Korman R, Zahler S, Adsit J, Dunny G. Two conjugation systems associated with Streptococcus faecalis plasmid pCF10: identification of a conjugative transposon that transfers between S. faecalis and Bacillus subtilis. Journal of Bacteriology. 1987;169(6):2529-36.

454. Courvalin P, Carlier C. Tn 1545: a conjugative shuttle transposon. Molecular and General Genetics MGG. 1987;206(2):259-64.

455. Cochetti I, Tili E, Mingoia M, Varaldo PE, Montanari MP. erm (B)-carrying elements in tetracycline-resistant pneumococci and correspondence between Tn1545 and Tn6003. Antimicrobial agents and chemotherapy. 2008;52(4):1285-90.

456. Garnier F, Taourit S, Glaser P, Courvalin P, Galimand M. Characterization of transposon Tn1549, conferring VanB-type resistance in Enterococcus spp. Microbiology (Reading). 2000;146 ( Pt 6):1481-9; doi: 10.1099/00221287-146-6-1481.

457. Del Grosso M, d'Abusco AS, Iannelli F, Pozzi G, Pantosti A. Tn2009, a Tn916-like element containing mef (E) in Streptococcus pneumoniae. Antimicrobial agents and chemotherapy. 2004;48(6):2037-42.

458. Li Y, Tomita H, Lv Y, Liu J, Xue F, Zheng B, et al. Molecular characterization of erm (B)-and mef (E)-mediated erythromycin-resistant Streptococcus pneumoniae in China and complete DNA sequence of Tn2010. Journal of applied microbiology. 2011;110(1):254-65.

459. Del Grosso M, Camilli R, Barbabella G, Northwood JB, Farrell DJ, Pantosti A. Genetic resistance elements carrying mef subclasses other than mef (A) in Streptococcus pyogenes. Antimicrobial agents and chemotherapy. 2011;55(7):3226-30.

460. Horaud T, Delbos F, de Cespédès G. Tn 3702, a conjugative transposon in Enterococcus faecalis. FEMS microbiology letters. 1990;72(1-2):189-94.

461. McDougal LK, Tenover FC, Lee LN, Rasheed JK, Patterson JE, Jorgensen JH, et al. Detection of Tn917-like sequences within a Tn916-like conjugative transposon (Tn3872) in erythromycin-resistant isolates of Streptococcus pneumoniae. Antimicrobial agents and chemotherapy. 1998;42(9):2312-8.

462. Fletcher HM, Marri L, Daneo-Moore L. Transposon-916-Iike Elements in Clinical Isolates of Enterococcus faecium. Microbiology. 1989;135(11):3067-77.

463. Ayoubi P, Kilic A, Vijayakumar M. Tn5253, the pneumococcal omega (cat tet) BM6001 element, is a composite structure of two conjugative transposons, Tn5251 and Tn5252. Journal of Bacteriology. 1991;173(5):1617-22.

464. Rice LB, Marshall SH, Carias LL. Tn5381, a conjugative transposon identifiable as a circular form in Enterococcus faecalis. Journal of bacteriology. 1992;174(22):7308-15.

465. Roberts AP, Johanesen PA, Lyras D, Mullany P, Rood JI. Comparison of Tn5397 from Clostridium difficile, Tn916 from Enterococcus faecalis and the CW459tet (M) element from Clostridium perfringens shows that they have similar conjugation regions but different insertion and excision modulesThe GenBank accession numbers for the sequences in this paper are AF333235 (Tn5397) and AF329848 [part of CW459tet (M)]. Microbiology. 2001;147(5):1243-51.

466. Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, et al. Whole genome sequencing of meticillin-resistant Staphylococcus aureus. The Lancet. 2001;357(9264):1225-40.

467. Brouwer MS, Mullany P, Roberts AP. Characterization of the conjugative transposon Tn 6000 from Enterococcus casseliflavus 664.1 H1 (formerly Enterococcus faecium 664.1 H1). FEMS microbiology letters. 2010;309(1):71-6.

468. Palmieri C, Mingoia M, Massidda O, Giovanetti E, Varaldo PE. Streptococcus pneumoniae transposon Tn1545/Tn6003 changes to Tn6002 due to spontaneous excision in circular form of the erm (B)-and aphA3-containing macrolide-aminoglycoside-streptothricin (MAS) element. Antimicrobial agents and chemotherapy. 2012;56(11):5994-7.

469. Soge OO, Beck NK, White TM, No DB, Roberts MC. A novel transposon, Tn 6009, composed of a Tn 916 element linked with a Staphylococcus aureus mer operon. Journal of Antimicrobial Chemotherapy. 2008;62(4):674-80.

470. De Vries LE, Christensen H, Skov RL, Aarestrup FM, Agersø Y. Diversity of the tetracycline resistance gene tet (M) and identification of Tn 916-and Tn 5801-like (Tn 6014) transposons

in Staphylococcus aureus from humans and animals. Journal of antimicrobial chemotherapy. 2009;64(3):490-500.

471. Rice LB, Carias LL, Rudin S, Hutton RA, Marshall S. Multiple copies of functional, Tet (M)-encoding Tn916-like elements in a clinical Enterococcus faecium isolate. Plasmid. 2010;64(3):150-5.
472. Roberts AP, Mullany P. A modular master on the move: the Tn916 family of mobile genetic elements. Trends Microbiol. 2009;17(6):251-8; doi: 10.1016/j.tim.2009.03.002.
473. Brochet M, Rusniok C, Couvé 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; doi: 10.1073/pnas.0803654105.
474. Brochet M, Couvé E, Bercion R, Sire J-M, Glaser P. Population structure of human isolates of Streptococcus agalactiae from Dakar and Bangui. Journal of clinical microbiology. 2009;47(3):800-3.
475. Whittle G, Hamburger N, Shoemaker NB, Salyers AA. A bacteroides conjugative transposon, CTnERL, can transfer a portion of itself by conjugation without excising from the chromosome. J Bacteriol. 2006;188(3):1169-74; doi: 10.1128/jb.188.3.1169-1174.2006.
476. Guérillot R, Da Cunha V, Sauvage E, Bouchier C, Glaser P. Modular evolution of TnGBSs, a new family of integrative and conjugative elements associating insertion sequence transposition, plasmid replication, and conjugation for their spreading. Journal of bacteriology. 2013;195(9):1979-90; doi: 10.1128/JB.01745-12.
477. Zhou K, Xie L, Han L, Guo X, Wang Y, Sun J. ICESag37, a novel integrative and conjugative element carrying antimicrobial resistance genes and potential virulence factors in Streptococcus agalactiae. Frontiers in microbiology. 2017;8:1921.
478. Morici E, Simoni S, Brenciani A, Giovanetti E, Varaldo PE, Mingoia M. A new mosaic integrative and conjugative element from Streptococcus agalactiae carrying resistance genes for chloramphenicol (catQ) and macrolides [mef (I) and erm (TR)]. Journal of Antimicrobial Chemotherapy. 2016;72(1):64-7.
479. Campisi E, Rosini R, Ji W, Guidotti S, Rojas-López M, Geng G, et al. Genomic analysis reveals multi-drug resistance clusters in group B Streptococcus CC17 hypervirulent isolates causing neonatal invasive disease in southern mainland China. Frontiers in microbiology. 2016;7:1265.
480. England PH: Laboratory surveillance of pyogenic and non-pyogenic streptococcal bacteraemia in England, Wales and Northern Ireland. In.; 2014.
481. England PH: Laboratory surveillance of pyogenic and non-pyogenic streptococcal bacteraemia in England, Wales and Northern Ireland. In.; 2015.

482. 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. Journal of microbiological methods. 2010;80(2):212-4.
483. Jorgensen JH, Turnidge JD. Susceptibility test methods: dilution and disk diffusion methods. In: Manual of Clinical Microbiology, Eleventh Edition. American Society of Microbiology; 2015. p. 1253-73.
484. Hudzicki J. Kirby-Bauer Disk Diffusion Susceptibility Test Protocol. American Society for Microbiology. 2009.
485. Wayne P. CLSI: Performance standards for antimicrobial susceptibility testing; Twenty-seventh informational supplement. 2017;CLSI document M100-S22.
486. Wayne P. Performance standards for antimicrobial susceptibility testing: eighteenth informational Supplement. 2008;M100-S18 CLSI.
487. Singh T, Deshmukh AB, Chitnis V, Bajpai T. Inducible clindamycin resistance among the clinical isolates of Staphylococcus aureus in a tertiary care hospital. International Journal of Health & Allied Sciences. 2016;5(2):111.
488. Baker CN, Stocker SA, Culver DH, Thornsberry C. Comparison of the E Test to agar dilution, broth microdilution, and agar diffusion susceptibility testing techniques by using a special challenge set of bacteria. J Clin Microbiol. 1991;29(3):533-8; doi: 10.1128/jcm.29.3.533-538.1991.
489. Citron DM, Ostovari MI, Karlsson A, Goldstein EJ. Evaluation of the E test for susceptibility testing of anaerobic bacteria. J Clin Microbiol. 1991;29(10):2197-203; doi: 10.1128/jcm.29.10.2197-2203.1991.
490. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS computational biology. 2017;13(6):e1005595.
491. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20; doi: 10.1093/bioinformatics/btu170.
492. Andrews S: FastQC: a quality control tool for high throughput sequence data. In.: Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom; 2010.
493. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology. 2012;19(5):455-77.
494. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS one. 2014;9(11):e112963.
495. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957-63; doi: 10.1093/bioinformatics/btr507.

496. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5; doi: 10.1093/bioinformatics/btt086.

497. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9; doi: 10.1093/bioinformatics/btu153.

498. Ripma LA, Simpson MG, Hasenstab-Lehman K. Geneious! Simplified genome skimming methods for phylogenetic systematic studies: A case study in Oreocarya (Boraginaceae). Appl Plant Sci. 2014;2(12):apps.1400062; doi: 10.3732/apps.1400062.

499. Yamamoto S, Miyake K, Koike Y, Watanabe M, Machida Y, Ohta M, et al. Molecular characterization of type-specific capsular polysaccharide biosynthesis genes of Streptococcus agalactiae type Ia. J Bacteriol. 1999;181(17):5176-84.

500. Watanabe M, Miyake K, Yanae K, Kataoka Y, Koizumi S, Endo T, et al. Molecular characterization of a novel β1, 3-galactosyltransferase for capsular polysaccharide synthesis by Streptococcus agalactiae type Ib. The journal of biochemistry. 2002;131(2):183-91.

501. Inouye M, Dashnow H, Raven L-A, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. Genome Medicine. 2014;6(11):90; doi: 10.1186/s13073-014-0090-6.

502. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. Journal of clinical microbiology. 2012;50(4):1355-61.

503. Nascimento M, Sousa A, Ramirez M, Francisco AP, Carriço JA, Vaz C. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. Bioinformatics. 2017;33(1):128-9; doi: 10.1093/bioinformatics/btw582.

504. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. Journal of Antimicrobial Chemotherapy. 2012;67(11):2640-4; doi: 10.1093/jac/dks261.

505. Metcalf BJ, Chochua S, Gertz RE, Jr., 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. Clin Microbiol Infect. 2017;23(8):574.e7-.e14; doi: 10.1016/j.cmi.2017.02.021.

506. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, et al. Multiple sequence alignment with the Clustal series of programs. Nucleic acids research. 2003;31(13):3497-500; doi: 10.1093/nar/gkg500.

507. Fischer P, Pawlowski A, Cao D, Bell D, Kitson G, Darsley M, et al. Safety and immunogenicity of a prototype recombinant alpha-like protein subunit vaccine (GBS-NN) against Group B Streptococcus in a randomised placebo-controlled double-blind phase 1 trial in healthy adult women. Vaccine. 2021;39(32):4489-99; doi: https://doi.org/10.1016/j.vaccine.2021.06.046.

508. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312-3; doi: 10.1093/bioinformatics/btu033.
509. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic acids research. 2019;47(W1):W256-w9; doi: 10.1093/nar/gkz239.
510. 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. 2015;43(3):e15-e; doi: 10.1093/nar/gku1196.
511. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC genomics. 2011. doi: 10.1186/1471-2164-12-402.
512. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.
513. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome biology. 2016;17(1):238.
514. Le Doare K, Heath PT. An overview of global GBS epidemiology. Vaccine. 2013;31:D7-D12.
515. Björnsdóttir E, Martins E, Erlendsdóttir H, Haraldsson G, Melo-Cristino J, Kristinsson K, et al. Changing epidemiology of group B streptococcal infections among adults in Iceland: 1975–2014. Clinical Microbiology and Infection. 2016;22(4):379. e9-. e16.
516. 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; doi: 10.1001/jamainternmed.2018.7269.
517. Collin SM, Shetty N, Lamagni T. Invasive group B streptococcus infections in adults, England, 2015–2016. Emerging Infectious Diseases. 2020;26(6):1174.
518. Fraser C, Hanage WP, Spratt BG. Recombination and the nature of bacterial speciation. Science. 2007;315(5811):476-80; doi: 10.1126/science.1127573.
519. Perron GG, Lee AE, Wang Y, Huang WE, Barraclough TG. Bacterial recombination promotes the evolution of multi-drug-resistance in functionally diverse populations. Proceedings of the Royal Society B: Biological Sciences. 2012;279(1733):1477-84.
520. Thomas CM, Nielsen KM. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nature reviews microbiology. 2005;3(9):711-21.
521. Gelot C, Le-Guen T, Ragu S, Lopez B. Double-Strand Break Repair: Homologous Recombination in Mammalian Cells. In: Genome Stability. Elsevier; 2016. p. 337-51.
522. Tham K-C, Kanaar R, Lebbink JH. Mismatch repair and homeologous recombination. DNA repair. 2016;38:75-83.
523. Jenkins S, Mukherjee S, Heyer W-D. DNA repair by homologous recombination. 2016.

524. Bishop AJ, Schiestl RH. Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects. Human molecular genetics. 2000;9(16):2427-334.

525. Lasko D, Cavenee W, Nordenskjöld M. Loss of constitutional heterozygosity in human cancer. Annual review of genetics. 1991;25(1):281-314.

526. Calabretta B, Robberson DL, Barrera-Saldaña HA, Lambrou TP, Saunders GF. Genome instability in a region of human DNA enriched in Alu repeat sequences. Nature. 1982;296(5854):219-25.

527. Spratt BG, Hanage WP, Feil EJ. The relative contributions of recombination and point mutation to the diversification of bacterial clones. Curr Opin Microbiol. 2001;4(5):602-6; doi: 10.1016/s1369-5274(00)00257-5.

528. Madhi SA, Cutland CL, Jose L, Koen A, Govender N, Wittke F, et al. 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;16(8):923-34.

529. Absalon J, Segall N, Block SL, Center KJ, Scully IL, Giardina PC, et al. Safety and immunogenicity of a novel hexavalent group B streptococcus conjugate vaccine in healthy, non-pregnant adults: a phase 1/2, randomised, placebo-controlled, observer-blinded, dose-escalation trial. The Lancet Infectious Diseases. 2021;21(2):263-74.

530. Feikin DR, Kagucia EW, Loo JD, Link-Gelles R, Puhan MA, Cherian T, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. PLoS medicine. 2013;10(9):e1001517.

531. Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. Nature genetics. 2013;45(6):656-63.

532. 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. 2014;33(7):1155-62.

533. Collin SM, Groves N, O'Sullivan C, Jauneikaite E, Patel D, Cunney R, et al. Uncovering Infant Group B Streptococcal (GBS) Disease Clusters in the United Kingdom and Ireland Through Genomic Analysis: A Population-based Epidemiological Study. Clinical Infectious Diseases. 2021;72(9):e296-e302.

534. Gogarten JP, Townsend JP. Horizontal gene transfer, genome innovation and evolution. Nature Reviews Microbiology. 2005;3(9):679-87.

535. Bröker G, Spellerberg B. Surface proteins of Streptococcus agalactiae and horizontal gene transfer. International journal of medical microbiology. 2004;294(2-3):169-75.

536. Sørensen UBS, Poulsen K, Ghezzo C, Margarit I, Kilian M. Emergence and global dissemination of host-specific Streptococcus agalactiae clones. MBio. 2010;1(3):e00178-10.

537. Teatero S, Lemire P, Dewar K, Wasserscheid J, Calzas C, Mallo GV, et al. Genomic recombination leading to decreased virulence of group B Streptococcus in a mouse model of adult invasive disease. Pathogens. 2016;5(3):54.

538. Campisi E, Rinaudo CD, Donati C, Barucco M, Torricelli G, Edwards MS, et al. Serotype IV Streptococcus agalactiae ST-452 has arisen from large genomic recombination events between CC23 and the hypervirulent CC17 lineages. Scientific reports. 2016;6(1):1-11.

539. Kalimuddin S, Chen SL, Lim CT, Koh TH, Tan TY, Kam M, et al. 2015 epidemic of severe Streptococcus agalactiae sequence type 283 infections in Singapore associated with the consumption of raw freshwater fish: a detailed analysis of clinical, epidemiological, and bacterial sequencing data. Clinical Infectious Diseases. 2017;64(suppl_2):S145-S52.

540. Weinberger DM, Malley R, Lipsitch M. Serotype replacement in disease after pneumococcal vaccination. The Lancet. 2011;378(9807):1962-73.

541. Bergseng H, Rygg M, Bevanger L, Bergh K. Invasive group B streptococcus (GBS) disease in Norway 1996–2006. European journal of clinical microbiology & infectious diseases. 2008;27(12):1193-9.

542. Pitts SI, Maruthur NM, Langley GE, Pondo T, Shutt KA, Hollick R, et al.: Obesity, diabetes, and the risk of invasive group B streptococcal disease in nonpregnant adults in the United States. In: Open forum infectious diseases. vol. 5: Oxford University Press US; 2018: ofy030.

543. Shabayek S, Abdalla S. Macrolide-and tetracycline-resistance determinants of colonizing group B streptococcus in women in Egypt. Journal of medical microbiology. 2014;63(10):1324-7.

544. Mengist HM, Zewdie O, Belew A, Dabsu R. Prevalence and drug susceptibility pattern of group B Streptococci (GBS) among pregnant women attending antenatal care (ANC) in Nekemte Referral Hospital (NRH), Nekemte, Ethiopia. BMC research notes. 2017;10(1):388; doi: 10.1186/s13104-017-2725-3.

545. Pediatrics AAo. Prevention of Group B Streptococcal Early-Onset Disease in Newborns. Pediatrics. 2019;144(2).

546. Brown CE, Christmas JT, Bawdon RE. Placental transfer of cefazolin and piperacillin in pregnancies remote from term complicated by Rh isoimmunization. American journal of obstetrics and gynecology. 1990;163(3):938-43.

547. Groff SM, Fallatah W, Yang S, Murphy J, Crutchfield C, Marzinke M, et al. Effect of maternal obesity on maternal-fetal transfer of preoperative cefazolin at cesarean section. The Journal of Pediatric Pharmacology and Therapeutics. 2017;22(3):227-32.

548. Mitchell TF, Pearlman MD, Chapman RL, Bhatt-Mehta V, Faix RG. Maternal and transplacental pharmacokinetics of cefazolin. Obstetrics & Gynecology. 2001;98(6):1075-9.

549. Allegaert K, Van Mieghem T, Verbesselt R, de Hoon J, Rayyan M, Devlieger R, et al. Cefazolin pharmacokinetics in maternal plasma and amniotic fluid during pregnancy. American journal of obstetrics and gynecology. 2009;200(2):170. e1-. e7.

550. Plainvert C, Hays C, Touak G, Joubrel-Guyot C, Dmytruk N, Frigo A, et al. Multidrug-Resistant Hypervirulent Group B Streptococcus in Neonatal Invasive Infections, France, 2007–2019. Emerging infectious diseases. 2020;26(11):2721.

551. Genovese C, D'Angeli F, Di Salvatore V, Tempera G, Nicolosi D. Streptococcus agalactiae in pregnant women: serotype and antimicrobial susceptibility patterns over five years in Eastern Sicily (Italy). European Journal of Clinical Microbiology & Infectious Diseases. 2020;39(12):2387-96; doi: 10.1007/s10096-020-03992-8.

552. do Nascimento CS, Dos Santos NFB, Ferreira RCC, Taddei CR. Streptococcus agalactiae in pregnant women in Brazil: prevalence, serotypes, and antibiotic resistance. Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]. 2019;50(4):943-52; doi: 10.1007/s42770-019-00129-8.

553. Hughes R, Brocklehurst P, Steer P, Heath P, Stenson B. on behalf of the Royal College of Obstetricians and Gynaecologists. Prevention of early-onset neonatal group B streptococcal disease. Green-top Guideline No. 36. BJOG. 2017;124(12):e280-305.

554. Mohamed Sadaka S, Abdelsalam Aly H, Ahmed Meheissen M, Orief YI, Mohamed Arafa B. Group B streptococcal carriage, antimicrobial susceptibility, and virulence related genes among pregnant women in Alexandria, Egypt. Alexandria Journal of Medicine. 2018;54(1):69-76; doi: https://doi.org/10.1016/j.ajme.2017.01.003.

555. Park C, Nichols M, Schrag SJ. Two cases of invasive vancomycin-resistant group B streptococcus infection. New England Journal of Medicine. 2014;370(9):885-6.

556. McKessar SJ, Berry AM, Bell JM, Turnidge JD, Paton JC. Genetic characterization of vanG, a novel vancomycin resistance locus of Enterococcus faecalis. Antimicrobial agents and chemotherapy. 2000;44(11):3224-8.

557. Depardieu F, Bonora MG, Reynolds PE, Courvalin P. The vanG glycopeptide resistance operon from Enterococcus faecalis revisited. Molecular microbiology. 2003;50(3):931-48.

558. Goessweiner-Mohr N, Arends K, Keller W, Grohmann E. Conjugation in gram-positive bacteria. Microbiology spectrum. 2014;2(4):2.4. 19.

559. Tosi M, Roat E, De Biasi S, Munari E, Venturelli S, Coloretti I, et al. Multidrug resistant bacteria in critically ill patients: a step further antibiotic therapy. J Emerg Crit Care Med. 2018;2:103.

560. Palmieri C, Magi G, Creti R, Baldassarri L, Imperi M, Gherardi G, et al. Interspecies mobilization of an erm (T)-carrying plasmid of Streptococcus dysgalactiae subsp. equisimilis by a coresident ICE of the ICE Sa2603 family. Journal of Antimicrobial Chemotherapy. 2013;68(1):23-6.

561. Ciric L, Jasni A, de Vries E, Agerso Y, Mullany P, Roberts AP. The Tn916/Tn1545 Family of Conjugative Transposon. In: Landes Bioscience; 2013.
562. Salyers AA, Shoemaker NB, Stevens AM, Li LY. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol Rev. 1995;59(4):579-90; doi: 10.1128/mr.59.4.579-590.1995.
563. Muñoz-López M, García-Pérez JL. DNA transposons: nature and applications in genomics. Curr Genomics. 2010;11(2):115-28; doi: 10.2174/138920210790886871.
564. Franke AE, Clewell D. Evidence for a chromosome-borne resistance transposon (Tn916) in Streptococcus faecalis that is capable of" conjugal" transfer in the absence of a conjugative plasmid. Journal of Bacteriology. 1981;145(1):494-502.
565. Burrus V, Pavlovic G, Decaris B, Guédon G. Conjugative transposons: the tip of the iceberg. Molecular microbiology. 2002;46(3):601-10.
566. Vandecraen J, Chandler M, Aertsen A, Van Houdt R. The impact of insertion sequences on bacterial genome plasticity and adaptability. Critical reviews in microbiology. 2017;43(6):709-30.
567. Oppegaard O, Skrede S, Mylvaganam H, Kittang BR. Emerging threat of antimicrobial resistance in β-hemolytic streptococci. Frontiers in microbiology. 2020;11:797.
568. Mingoia M, Morici E, Tili E, Giovanetti E, Montanari MP, Varaldo PE. Characterization of Tn5801.Sag, a variant of Staphylococcus aureus Tn916 family transposon Tn5801 that is widespread in clinical isolates of Streptococcus agalactiae. Antimicrobial agents and chemotherapy. 2013;57(9):4570-4; doi: 10.1128/AAC.00521-13.
569. León-Sampedro R, Novais C, Peixe L, Baquero F, Coque TM. Diversity and Evolution of the Tn5801-tet(M)-Like Integrative and Conjugative Elements among Enterococcus, Streptococcus, and Staphylococcus. 2016;60(3):1736-46; doi: 10.1128/aac.01864-15.
570. Giovanetti E, Brenciani A, Tiberi E, Bacciaglia A, Varaldo PE. ICESp2905, the erm (TR)-tet (O) element of Streptococcus pyogenes, is formed by two independent integrative and conjugative elements. Antimicrobial agents and chemotherapy. 2012;56(1):591-4.
571. Del Grosso M, Camilli R, Rizzi E, Pietrelli A, De Bellis G, Pantosti A. ICESpy009, a conjugative genetic element carrying mef (E) in Streptococcus pyogenes. Antimicrobial agents and chemotherapy. 2016;60(7):3906-12.
572. Achard A, Villers C, Pichereau V, Leclercq R. New lnu (C) gene conferring resistance to lincomycin by nucleotidylation in Streptococcus agalactiae UCN36. Antimicrobial agents and chemotherapy. 2005;49(7):2716-9.
573. Buu-Hoï A, Le Bouguenec C, Horaud T. High-level chromosomal gentamicin resistance in Streptococcus agalactiae (group B). Antimicrob Agents Chemother. 1990;34(6):985-8; doi: 10.1128/aac.34.6.985.

574. Galimand M, Lambert T, Gerbaud G, Courvalin P. High-level aminoglycoside resistance in the beta-hemolytic group G Streptococcus isolate BM2721. Antimicrob Agents Chemother. 1999;43(12):3008-10; doi: 10.1128/aac.43.12.3008.

575. Kaufhold A, Potgieter E. Chromosomally mediated high-level gentamicin resistance in Streptococcus mitis. Antimicrob Agents Chemother. 1993;37(12):2740-2; doi: 10.1128/aac.37.12.2740.

576. Raabe VN, Shane AL. Group B Streptococcus (Streptococcus agalactiae). Microbiology spectrum. 2019;7(2):10.1128/microbiolspec.GPP3-0007-2018; doi: 10.1128/microbiolspec.GPP3-0007-2018.

577. Doumith M, Mushtaq S, Martin V, Chaudhry A, Adkin R, Coelho J, et al. Genomic sequences of Streptococcus agalactiae with high-level gentamicin resistance, collected in the BSAC bacteraemia surveillance. Journal of Antimicrobial Chemotherapy. 2017;72(10):2704-7; doi: 10.1093/jac/dkx207.

578. Tsai M-H, Hsu J-F, Lai M-Y, Lin L-C, Chu S-M, Huang H-R, et al. Molecular characteristics and antimicrobial resistance of group B Streptococcus strains causing invasive disease in neonates and adults. Frontiers in microbiology. 2019;10:264.

579. Motallebirad T, Fazeli H, Ghahiri A, Shokri D, Jalalifar S, Moghim S, et al. Prevalence, population structure, distribution of serotypes, pilus islands and resistance genes among erythromycin-resistant colonizing and invasive Streptococcus agalactiae isolates recovered from pregnant and non-pregnant women in Isfahan, Iran. BMC Microbiology. 2021;21(1):139; doi: 10.1186/s12866-021-02186-2.

580. Mudzana R, Mavenyengwa RT, Gudza-Mugabe M. Analysis of virulence factors and antibiotic resistance genes in group B streptococcus from clinical samples. BMC Infectious Diseases. 2021;21(1):125; doi: 10.1186/s12879-021-05820-6.

581. Heelan JS, Hasenbein ME, McAdam AJ. Resistance of group B streptococcus to selected antibiotics, including erythromycin and clindamycin. Journal of clinical microbiology. 2004;42(3):1263-4.

582. Dela Cruz WP, Richardson JY, Broestler JM, Thornton JA, Danaher PJ. Rapid determination of macrolide and lincosamide resistance in group B streptococcus isolated from vaginal-rectal swabs. Infect Dis Obstet Gynecol. 2007;2007:46581-; doi: 10.1155/2007/46581.

583. Grossman TH. Tetracycline Antibiotics and Resistance. Cold Spring Harb Perspect Med. 2016;6(4):a025387-a; doi: 10.1101/cshperspect.a025387.

584. Cookson AL, Noel S, Hussein H, Perry R, Sang C, Moon CD, et al. Transposition of Tn916 in the four replicons of the Butyrivibrio proteoclasticus B316(T) genome. FEMS Microbiol Lett. 2011;316(2):144-51; doi: 10.1111/j.1574-6968.2010.02204.x.

585. Mullany P, Williams R, Langridge GC, Turner DJ, Whalan R, Clayton C, et al. Behavior and target site selection of conjugative transposon Tn916 in two different strains of toxigenic

Clostridium difficile. Applied and environmental microbiology. 2012;78(7):2147-53; doi: 10.1128/AEM.06193-11.

586. Rosen GH, Randis TM, Desai PV, Sapra KJ, Ma B, Gajer P, et al. Group B Streptococcus and the Vaginal Microbiota. J Infect Dis. 2017;216(6):744-51; doi: 10.1093/infdis/jix395.

587. Berardi A, Rossi C, Creti R, China M, Gherardi G, Venturelli C, et al. Group B streptococcal colonization in 160 mother-baby pairs: a prospective cohort study. J Pediatr. 2013;163(4):1099-104.e1; doi: 10.1016/j.jpeds.2013.05.064.

588. Kvist LJ, Larsson BW, Hall-Lord ML, Steen A, Schalén C. The role of bacteria in lactational mastitis and some considerations of the use of antibiotic treatment. Int Breastfeed J. 2008;3:6; doi: 10.1186/1746-4358-3-6.

589. Baker CJ, Barrett FF. Transmission of group B streptococci among parturient women and their neonates. J Pediatr. 1973;83(6):919-25; doi: 10.1016/s0022-3476(73)80524-4.

590. Benitz WE, Gould JB, Druzin ML. Risk factors for early-onset group B streptococcal sepsis: estimation of odds ratios by critical literature review. Pediatrics. 1999;103(6):e77-e.

591. Boyer K, Gadzala C, Burd L, Fisher D, Paton J, Gotoff S. Selective intrapartum chemoprophylaxis of neonatal group B streptococcal early-onset disease. I. Epidemiologic rationale. Journal of Infectious Diseases. 1983;148(5):795-801.

592. Hosoda A, Gatayama R, Moriyama S, Ishii N, Yamada K, Matsuzaki Y, et al. The first case of recurrent ultra late onset group B streptococcal sepsis in a 3-year-old child. IDCases. 2016;7:16-8; doi: 10.1016/j.idcr.2016.11.007.

593. Hanna M, Noor A. Streptococcus Group B. In: StatPearls [Internet] 2020;Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK553143/?report=classic.

594. Kwatra G, Cunnington 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. The Lancet Infectious Diseases. 2016;16(9):1076-84; doi: https://doi.org/10.1016/S1473-3099(16)30055-X.

595. Taylor KL: A study of group B streptococcus in Brisbane: the epidemiology, detection by PCR assay and serovar prevalence. In.: Queensland University of Technology; 2006.

596. Zeng X, Kong F, Wang H, Darbar A, Gilbert GL. Simultaneous detection of nine antibiotic resistance-related genes in Streptococcus agalactiae using multiplex PCR and reverse line blot hybridization assay. Antimicrobial agents and chemotherapy. 2006;50(1):204-9.

597. Ellis S, Kotiw M, Garland S. Restriction endonuclease analysis of group B streptococcal isolates from two distinct geographical regions. Journal of Hospital Infection. 1996;33(4):279-87.

598. Slotved H, Dayie N, Banini J, Frimodt-Møller N. Carriage and serotype distribution of Streptococcus agalactiae in third trimester pregnancy Jul 21; 17 (n Ghana. BMC Pr 1186/s12884-birth. 2017.

599. Furfaro LL, Chang BJ, Payne MS. Perinatal Streptococcus agalactiae epidemiology and surveillance targets. Clinical microbiology reviews. 2018;31(4).

600. Nakamura PA, Schuab RBB, Neves FP, Pereira CF, de Paula GR, Barros RR. Antimicrobial resistance profiles and genetic characterisation of macrolide resistant isolates of Streptococcus agalactiae. Memórias do Instituto Oswaldo Cruz. 2011;106(2):119-22.

601. Melo SC, Santos NC, Oliveira M, Scodro RB, Cardoso RF, Pádua RA, et al. ANTIMICROBIAL SUSCEPTIBILITY OF Streptococcus agalactiae ISOLATED FROM PREGNANT WOMEN. Rev Inst Med Trop Sao Paulo. 2016;58:83; doi: 10.1590/s1678-9946201658083.

602. Martins ER, Melo-Cristino J, Ramirez M. Dominance of serotype Ia among group B Streptococci causing invasive infections in nonpregnant adults in Portugal. J Clin Microbiol. 2012;50(4):1219-27; doi: 10.1128/jcm.05488-11.

603. Lamy M-C, Dramsi S, Billoët A, Réglier-Poupet H, Tazi A, Raymond J, et al. Rapid detection of the "highly virulent" group B Streptococcus ST-17 clone. Microbes and infection. 2006;8(7):1714-22.

604. Florindo C, Damiao V, Silvestre I, Farinha C, Rodrigues F, Nogueira F, et al. Epidemiological surveillance of colonising group B Streptococcus epidemiology in the Lisbon and Tagus Valley regions, Portugal (2005 to 2012): emergence of a new epidemic type IV/clonal complex 17 clone. Eurosurveillance. 2014;19(23):20825.

605. 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.

606. Gizachew M, Tiruneh M, Moges F, Tessema B. Streptococcus agalactiae maternal colonization, antibiotic resistance and serotype profiles in Africa: a meta-analysis. Annals of clinical microbiology and antimicrobials. 2019;18(1):1-14.

607. Campo CH, Martínez MF, Otero JC, Rincón G. Vagino-rectal colonization prevalence by Streptococcus agalactiae and its susceptibility profile in pregnant women attending a third-level hospital. Biomédica. 2019;39(4):689-98.

608. Matani C, Trezzi M, Matteini A, Catalani C, Messeri D, Catalani C. Streptococcus agalactiae: prevalence of antimicrobial resistance in vaginal and rectal swabs in Italian pregnant women. Infez Med. 2016;24(3):217-21.

609. Martínez MA, Ovalle A, Durán C, Reid I, Urriola G, Garay B, et al. Serotipos y susceptibilidad antimicrobiana de Streptococcus agalactiae. Revista médica de Chile. 2004;132(5):549-55.

610. Tazi A, Gueudet T, Varon E, Gilly L, Trieu-Cuot P, Poyart C. Fluoroquinolone-resistant group B streptococci in acute exacerbation of chronic bronchitis. Emerging infectious diseases. 2008;14(2):349.

611. Wu C-J, Lai J-F, Huang I-W, Hsieh L-Y, Wang H-Y, Shiau Y-R, et al. Multiclonal emergence of levofloxacin-resistant group B Streptococcus, Taiwan. Journal of Antimicrobial Chemotherapy. 2017;72(12):3263-71.

612. Sociedade Brasileira de Pediatria (SBP). Available at http://wwwsbpcombr/publicacoes/. 2018.

613. MELO SCCSd, SANTOS NCdS, OLIVEIRA Md, SCODRO RBdL, Cardoso RF, Pádua RAF, et al. Antimicrobial susceptibility of Streptococcus agalactiae isolated from pregnant women. Revista do Instituto de Medicina Tropical de São Paulo. 2016;58.

614. Hua C, Yu H, Zhuang J, Li X, Xu H, Luo Q, et al. An analysis of 181 cases with blood stream infection caused by Streptococcus agalactiae in children from 2011 to 2015: a multi-center retrospective study. Zhonghua er ke za zhi= Chinese Journal of Pediatrics. 2016;54(8):577-81.

615. Malek-Jafarian M, Hosseini F-S, Ahmadi A-R. Pattern of infection and antibiotic activity among Streptococcus agalactiae isolates from adults in Mashhad, Iran. Reports of biochemistry & molecular biology. 2015;3(2):89.

616. Gizachew M, Tiruneh M, Moges F, Adefris M, Tigabu Z, Tessema B. Newborn colonization and antibiotic susceptibility patterns of Streptococcus agalactiae at the University of Gondar Referral Hospital, Northwest Ethiopia. BMC Pediatrics. 2018;18(1):378; doi: 10.1186/s12887-018-1350-1.

617. Assefa S, Desta K, Lema T. Group B streptococci vaginal colonization and drug susceptibility pattern among pregnant women attending in selected public antenatal care centers in Addis Ababa, Ethiopia. BMC pregnancy and childbirth. 2018;18(1):1-9.

618. Uh Y, Jang I, Hwang G, Yoon K, Song W. Emerging erythromycin resistance among group B streptococci in Korea. European Journal of Clinical Microbiology and Infectious Diseases. 2001;20(1):52-4.

619. Borchardt SM, DeBusscher JH, Tallman PA, Manning SD, Marrs CF, Kurzynski TA, et al. Frequency of antimicrobial resistance among invasive and colonizing Group B streptococcal isolates. BMC infectious diseases. 2006;6(1):1-8.

620. Desjardins M, Delgaty K, Ramotar K, Seetaram C, Toye B. Prevalence and mechanisms of erythromycin resistance in group A and group B Streptococcus: implications for reporting susceptibility results. Journal of clinical microbiology. 2004;42(12):5620-3.

621. Garland SM, Cottrill E, Markowski L, Pearce C, Clifford V, Ndisang D, et al. Antimicrobial resistance in group B streptococcus: the Australian experience. Journal of medical microbiology. 2011;60(2):230-5.

622. Read TD, Massey RC. Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: a new direction for bacteriology. Genome medicine. 2014;6(11):1-11.
623. Power RA, Parkhill J, de Oliveira T. Microbial genome-wide association studies: lessons from human GWAS. Nature reviews genetics. 2017;18(1):41-50.
624. Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA. Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet. 2006;38(1):82-5; doi: 10.1038/ng1695.
625. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nature Reviews Methods Primers. 2021;1(1):59; doi: 10.1038/s43586-021-00056-9.
626. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years of GWAS discovery: biology, function, and translation. The American Journal of Human Genetics. 2017;101(1):5-22.
627. Lees JA, Kremer PH, Manso AS, Croucher NJ, Ferwerda B, Serón MV, et al. Large scale genomic analysis shows no evidence for pathogen adaptation between the blood and cerebrospinal fluid niches during bacterial meningitis. Microbial genomics. 2017;3(1).
628. Li Y, Metcalf BJ, Chochua S, Li Z, Walker H, Tran T, et al. Genome-wide association analyses of invasive pneumococcal isolates identify a missense bacterial mutation associated with meningitis. Nature communications. 2019;10(1):1-11.
629. Kulohoma BW, Cornick JE, Chaguza C, Yalcin F, Harris SR, Gray KJ, et al. Comparative genomic analysis of meningitis-and bacteremia-causing pneumococci identifies a common core genome. Infection and immunity. 2015;83(10):4165-73.
630. Lilje B, Rasmussen RV, Dahl A, Stegger M, Skov RL, Fowler Jr VG, et al. Whole-genome sequencing of bloodstream Staphylococcus aureus isolates does not distinguish bacteraemia from endocarditis. Microbial genomics. 2017;3(11).
631. Young BC, Earle SG, Soeng S, Sar P, Kumar V, Hor S, et al. Panton–Valentine leucocidin is the key determinant of Staphylococcus aureus pyomyositis in a bacterial GWAS. Elife. 2019;8:e42486.
632. Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, et al. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. Proceedings of the national academy of sciences. 2013;110(29):11923-7.
633. Lees JA, Ferwerda B, Kremer PH, Wheeler NE, Serón MV, Croucher NJ, et al. Joint sequencing of human and pathogen genomes reveals the genetics of pneumococcal meningitis. Nature communications. 2019;10(1):1-14.
634. Laabei M, Recker M, Rudkin JK, Aldeljawi M, Gulay Z, Sloan TJ, et al. Predicting the virulence of MRSA from its genome sequence. Genome research. 2014;24(5):839-49.

635. Lees JA, Croucher NJ, Goldblatt D, Nosten F, Parkhill J, Turner C, et al. Genome-wide identification of lineage and locus specific variation associated with pneumococcal carriage duration. Elife. 2017;6:e26255.

636. Chaguza C, Yang M, Cornick JE, Du Plessis M, Gladstone RA, Kwambana-Adams BA, et al. Bacterial genome-wide association study of hyper-virulent pneumococcal serotype 1 identifies genetic variation associated with neurotropism. Communications biology. 2020;3(1):1-12.

637. Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K, Ali S, et al. Genome-wide analysis of multi-and extensively drug-resistant Mycobacterium tuberculosis. Nature genetics. 2018;50(2):307-16.

638. Chewapreecha C, Marttinen P, Croucher NJ, Salter SJ, Harris SR, Mather AE, et al. Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. PLoS genetics. 2014;10(8):e1004547.

639. Farhat MR, Shapiro BJ, Kieser KJ, Sultana R, Jacobson KR, Victor TC, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. Nature genetics. 2013;45(10):1183-9.

640. Suzuki M, Shibayama K, Yahara K. A genome-wide association study identifies a horizontally transferred bacterial surface adhesin gene associated with antimicrobial resistant strains. Scientific reports. 2016;6(1):1-10.

641. Hicks ND, Carey AF, Yang J, Zhao Y, Fortune SM. Bacterial genome-wide association identifies novel factors that contribute to ethionamide and prothionamide susceptibility in Mycobacterium tuberculosis. MBio. 2019;10(2):e00616-19.

642. Davies MR, McIntyre L, Mutreja A, Lacey JA, Lees JA, Towers RJ, et al. Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. Nature genetics. 2019;51(6):1035-43.

643. Obolski U, Gori A, Lourenço J, Thompson C, Thompson R, French N, et al. Identifying genes associated with invasive disease in S. pneumoniae by applying a machine learning approach to whole genome sequence typing data. Scientific reports. 2019;9(1):1-9.

644. Martins ER, Andreu A, Correia P, Juncosa T, Bosch J, Ramirez M, et al. Group B streptococci causing neonatal infections in barcelona are a stable clonal population: 18-year surveillance. J Clin Microbiol. 2011;49(8):2911-8; doi: 10.1128/jcm.00271-11.

645. Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, Van Der Linden M, et al. Rapid pneumococcal evolution in response to clinical interventions. science. 2011;331(6016):430-4.

646. Rosini R, Margarit I. Biofilm formation by Streptococcus agalactiae: influence of environmental conditions and implicated virulence factors. Frontiers in cellular and infection microbiology. 2015;5:6.

647. Landwehr-Kenzel S, Henneke P. Interaction of Streptococcus agalactiae and cellular innate immunity in colonization and disease. Frontiers in immunology. 2014;5:519.

648. Santi I, Grifantini R, Jiang S-M, Brettoni C, Grandi G, Wessels MR, et al. CsrRS regulates group B Streptococcus virulence gene expression in response to environmental pH: a new perspective on vaccine development. Journal of bacteriology. 2009;191(17):5387-97.

649. Di Xia F, Mallet A, Caliot E, Gao C, Trieu-Cuot P, Dramsi S. Capsular polysaccharide of Group B Streptococcus mediates biofilm formation in the presence of human plasma. Microbes and infection. 2015;17(1):71-6.

650. Patras KA, Derieux J, Al-Bassam MM, Adiletta N, Vrbanac A, Lapek JD, et al. Group B Streptococcus biofilm regulatory protein A contributes to bacterial physiology and innate immune resistance. The Journal of infectious diseases. 2018;218(10):1641-52.

651. Dammann AN, Chamby AB, Catomeris AJ, Davidson KM, Tettelin H, Van Pijkeren J-P, et al. Genome-Wide fitness analysis of group B Streptococcus in human amniotic fluid reveals a transcription factor that controls multiple virulence traits. PLoS pathogens. 2021;17(3):e1009116.

652. Buscetta M, Papasergi S, Firon A, Pietrocola G, Biondo C, Mancuso G, et al. FbsC, a novel fibrinogen-binding protein, promotes Streptococcus agalactiae-host cell interactions. Journal of Biological Chemistry. 2014;289(30):21003-15.

653. Périchon B, Szili N, Du Merle L, Rosinski-Chupin I, Gominet M, Bellais S, et al. Regulation of PI-2b pilus expression in hypervirulent Streptococcus agalactiae ST-17 BM110. PLoS One. 2017;12(1):e0169840.

654. Ragan MA, Beiko RG. Lateral genetic transfer: open issues. Philosophical Transactions of the Royal Society B: Biological Sciences. 2009;364(1527):2241-51.

655. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. The microbial pan-genome. Current opinion in genetics & development. 2005;15(6):589-94.

656. Gautreau G, Bazin A, Gachet M, Planel R, Burlot L, Dubois M, et al. PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. PLoS computational biology. 2020;16(3):e1007732-e; doi: 10.1371/journal.pcbi.1007732.

657. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Applied and environmental microbiology. 2013;79(24):7696-701.

658. Livingstone PG, Morphew RM, Whitworth DE. Genome Sequencing and Pan-Genome Analysis of 23 Corallococcus spp. Strains Reveal Unexpected Diversity, With Particular Plasticity of Predatory Gene Sets. Frontiers in Microbiology. 2018;9(3187); doi: 10.3389/fmicb.2018.03187.

659. Chen SL. Genomic Insights Into the Distribution and Evolution of Group B Streptococcus. Frontiers in microbiology. 2019;10:1447-; doi: 10.3389/fmicb.2019.01447.

660. Poyart C, Réglier-Poupet H, Tazi A, Billoët A, Dmytruk N, Bidet P, et al. Invasive group B streptococcal infections in infants, France. Emerging infectious diseases. 2008;14(10):1647.

661. Morozumi M, Wajima T, Takata M, Iwata S, Ubukata K. Molecular characteristics of group B streptococci isolated from adults with invasive infections in Japan. Journal of clinical microbiology. 2016;54(11):2695-700.

662. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic acids research. 2007;36(suppl_1):D480-D4.

663. Voth DE, Broederdorf LJ, Graham JG. Bacterial Type IV secretion systems: versatile virulence machines. Future microbiology. 2012;7(2):241-57.

664. Nobelmann B, Lengeler JW. Molecular analysis of the gat genes from Escherichia coli and of their roles in galactitol transport and metabolism. Journal of bacteriology. 1996;178(23):6790-5.

665. 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):9539; doi: 10.1038/s41598-020-66214-3.

666. Six A, Krajangwong S, Crumlish M, Zadoks RN, Walker D. Galleria mellonella as an infection model for the multi-host pathogen Streptococcus agalactiae reflects hypervirulence of strains associated with human invasive disease. Virulence. 2019;10(1):600-9.

667. Samen U, Eikmanns BJ, Reinscheid DJ, Borges F. The surface protein Srr-1 of Streptococcus agalactiae binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. Infection and immunity. 2007;75(11):5405-14.

668. Rajagopal L. Understanding the regulation of Group B Streptococcal virulence factors. Future microbiology. 2009;4(2):201-21; doi: 10.2217/17460913.4.2.201.

669. 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.

670. Martins E, Andreu A, Correia P, Juncosa T, Bosch J, Ramirez M, et al. Group B streptococci causing neonatal infections in Barcelona are a stable clonal population: 18-year surveillance. Journal of clinical microbiology. 2011;49(8):2911-8.

671. Lu B, Wu J, Chen X, Gao C, Yang J, Li Y, et al. Microbiological and clinical characteristics of Group B Streptococcus isolates causing materno-neonatal infections: high prevalence of CC17/PI-1 and PI-2b sublineage in neonatal infections. Journal of medical microbiology. 2018;67(11):1551-9.

672. Martins ER, Pedroso-Roussado C, Melo-Cristino J, Ramirez M, Oliveira H, Vaz T, et al. 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.

673. Ali MM, Woldeamanuel Y, Asrat D, Fenta DA, Beall B, Schrag S, et al. Features of Streptococcus agalactiae strains recovered from pregnant women and newborns attending different hospitals in Ethiopia. BMC infectious diseases. 2020;20(1):848-; doi: 10.1186/s12879-020-05581-8.

674. Cheng Z, Qu P, Ke P, Yang X, Zhou Q, Lan K, et al. Antibiotic Resistance and Molecular Epidemiological Characteristics of Streptococcus agalactiae Isolated from Pregnant Women in Guangzhou, South China. Canadian Journal of Infectious Diseases and Medical Microbiology. 2020;2020.

675. Zschiedrich CP, Keidel V, Szurmant H. Molecular mechanisms of two-component signal transduction. Journal of molecular biology. 2016;428(19):3752-75.

676. Ahn S-J, Qu M-D, Roberts E, Burne RA, Rice KC. Identification of the Streptococcus mutans LytST two-component regulon reveals its contribution to oxidative stress tolerance. BMC microbiology. 2012;12(1):1-12.

677. Seepersaud R, Bensing BA, Yen YT, Sullam PM. The accessory Sec protein Asp2 modulates GlcNAc deposition onto the serine-rich repeat glycoprotein GspB. Journal of bacteriology. 2012;194(20):5564-75.

678. Lycklama a Nijeholt JA, Driessen AJ. The bacterial Sec-translocase: structure and mechanism. Philosophical Transactions of the Royal Society B: Biological Sciences. 2012;367(1592):1016-28.

679. Fiester SE, Nwugo CC, Penwell WF, Neary JM, Beckett AC, Arivett BA, et al. Role of the carboxy terminus of SecA in iron acquisition, protein translocation, and virulence of the bacterial pathogen Acinetobacter baumannii. Infection and immunity. 2015;83(4):1354-65.

680. Soares MP, Weiss G. The Iron age of host–microbe interactions. EMBO reports. 2015;16(11):1482-500.

681. Hoffmann T, Schütz A, Brosius M, Völker A, Völker U, Bremer E. High-salinity-induced iron limitation in Bacillus subtilis. Journal of bacteriology. 2002;184(3):718-27; doi: 10.1128/jb.184.3.718-727.2002.

682. Bryan JD, Shelver DW. Streptococcus agalactiae CspA is a serine protease that inactivates chemokines. Journal of bacteriology. 2009;191(6):1847-54.

683. Sokol CL, Luster AD. The chemokine system in innate immunity. Cold Spring Harbor perspectives in biology. 2015;7(5):a016303.

684. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell. 2013;152(6):1237-51.

685. Chandrangsu P, Rensing C, Helmann JD. Metal homeostasis and resistance in bacteria. Nature Reviews Microbiology. 2017;15(6):338-50.

686. Juttukonda LJ, Skaar EP. Manganese homeostasis and utilization in pathogenic bacteria. Molecular microbiology. 2015;97(2):216-28.

687. Grohmann E, Muth Gn, Espinosa M. Conjugative plasmid transfer in gram-positive bacteria. Microbiology and molecular biology reviews. 2003;67(2):277-301.

688. Castillo F, Benmohamed A, Szatmari G. Xer site specific recombination: double and single recombinase systems. Frontiers in microbiology. 2017;8:453.

689. Kostriken R, Morita C, Heffron F. Transposon Tn3 encodes a site-specific recombination system: identification of essential sequences, genes, and actual site of recombination. Proceedings of the National Academy of Sciences. 1981;78(7):4041-5.

690. Schneider E, Hunke S. ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains. FEMS Microbiology Reviews. 1998;22(1):1-20; doi: 10.1111/j.1574-6976.1998.tb00358.x.

691. Lamagni TL, Keshishian C, Efstratiou A, Guy R, Henderson KL, Broughton K, et al. Emerging trends in the epidemiology of invasive group B streptococcal disease in England and Wales, 1991–2010. Clinical infectious diseases. 2013;57(5):682-8.

692. Nuccitelli A, Rinaudo CD, Maione D. Group B Streptococcus vaccine: state of the art. Ther Adv Vaccines. 2015;3(3):76-90; doi: 10.1177/2051013615579869.

693. 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); doi: https://doi.org/10.1099/mgen.0.000783.

694. Perovic O, Crewe-Brown HH, Khoosal M, Karstaedt AS. Invasive group B streptococcal disease in nonpregnant adults. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 1999;18(5):362-4; doi: 10.1007/pl00015020.

695. Jackson LA, Hilsdon R, Farley MM, Harrison LH, Reingold AL, Plikaytis BD, et al. Risk factors for group B streptococcal disease in adults. Annals of internal medicine. 1995;123(6):415-20; doi: 10.7326/0003-4819-123-6-199509150-00003.

696. High KP, Edwards MS, Baker CJ. Group B Streptococcal Infections in Elderly Adults. Clinical Infectious Diseases. 2005;41(6):839-47; doi: 10.1086/432804.

697. Farley MM. Group B streptococcal infection in older patients. Spectrum of disease and management strategies. Drugs & aging. 1995;6(4):293-300; doi: 10.2165/00002512-199506040-00004.

698. Sunkara B, Bheemreddy S, Lorber B, Lephart PR, Hayakawa K, Sobel JD, et al. Group B Streptococcus infections in non-pregnant adults: the role of immunosuppression. International Journal of Infectious Diseases. 2012;16(3):e182-e6; doi: https://doi.org/10.1016/j.ijid.2011.11.008.

699. Patras KA, Nizet V. Group B streptococcal maternal colonization and neonatal disease: molecular mechanisms and preventative approaches. Frontiers in Pediatrics. 2018;6:27.

700. Andrews JI, Diekema DJ, Hunter SK, Rhomberg PR, Pfaller MA, Jones RN, et al. Group B streptococci causing neonatal bloodstream infection: antimicrobial susceptibility and serotyping results from SENTRY centers in the Western Hemisphere. American journal of obstetrics and gynecology. 2000;183(4):859-62.

701. Lambertsen L, Ekelund K, Skovsted I, Liboriussen A, Slotved H-C. Characterisation of invasive group B streptococci from adults in Denmark 1999 to 2004. European journal of clinical microbiology & infectious diseases. 2010;29(9):1071-7.

702. Luan S-L, Granlund M, Sellin M, Lagergård T, Spratt BG, Norgren M. Multilocus sequence typing of Swedish invasive group B streptococcus isolates indicates a neonatally associated genetic lineage and capsule switching. Journal of clinical microbiology. 2005;43(8):3727-33.

703. Ali MM, Asrat D, Fenta DA, Chaka TE, Woldeamanuel Y. Group B Streptococcus colonization rate and serotype distribution among pregnant women and their newborns at Adama Hospital Medical College, Ethiopia. Scientific Reports. 2020;10(1):9301; doi: 10.1038/s41598-020-66474-z.

704. Imperi M, Gherardi G, Berardi A, Baldassarri L, Pataracchia M, Dicuonzo G, et al. Invasive neonatal GBS infections from an area-based surveillance study in Italy. Clinical microbiology and infection. 2011;17(12):1834-9.

705. 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):1-7.

706. 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.

707. To KN, Powell O, Jamrozy 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. J Microbiol Methods. 2021;190:106322; doi: 10.1016/j.mimet.2021.106322.

708. Bolduc GR, Baron MJ, Gravekamp C, Lachenauer CS, Madoff LC. The alpha C protein mediates internalization of group B Streptococcus within human cervical epithelial cells. Cellular microbiology. 2002;4(11):751-8; doi: 10.1046/j.1462-5822.2002.00227.x.

709. Joubrel C, Tazi A, Six A, Dmytruk N, Touak G, Bidet P, et al. Group B streptococcus neonatal invasive infections, France 2007-2012. Clin Microbiol Infect. 2015;21(10):910-6; doi: 10.1016/j.cmi.2015.05.039.

710. 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. 2016;1(7):16067; doi: 10.1038/nmicrobiol.2016.67.

711. Chrispin Chaguza DMJ, Merijn W. Bijlsma, Taco W. Kuijpers, Diederik van de Beek, Arie van der Ende, Stephen D. Bentley. GWAS of Group B Streptococcus identifies genetic variation

influencing neonatal disease onset and meningeal invasion. (manuscript submitted for publication).

712. Prestinaci F, Pezzotti P, Pantosti A. Antimicrobial resistance: a global multifaceted phenomenon. Pathog Glob Health. 2015;109(7):309-18; doi: 10.1179/2047773215Y.0000000030.
713. De Francesco M, Caracciolo S, Gargiulo F, Manca N. Phenotypes, genotypes, serotypes and molecular epidemiology of erythromycin-resistant Streptococcus agalactiae in Italy. European journal of clinical microbiology & infectious diseases. 2012;31(8):1741-7.
714. Morozumi M, Wajima T, Kuwata Y, Chiba N, Sunaoshi K, Sugita K, et al. Associations between capsular serotype, multilocus sequence type, and macrolide resistance in Streptococcus agalactiae isolates from Japanese infants with invasive infections. Epidemiology & Infection. 2014;142(4):812-9.
715. Martins ER, Melo-Cristino J, Ramirez M. Evidence for rare capsular switching in Streptococcus agalactiae. Journal of bacteriology. 2010;192(5):1361-9.
716. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature. 2000;405(6784):299-304; doi: 10.1038/35012500.
717. Burmeister AR. Horizontal Gene Transfer. Evol Med Public Health. 2015;2015(1):193-4; doi: 10.1093/emph/eov018.
718. Roberts AP, Mullany P. A modular master on the move: the Tn916 family of mobile genetic elements. Trends in microbiology. 2009;17(6):251-8.
719. Emaneini M, Mirsalehian A, Beigverdi R, Fooladi AAI, Asadi F, Jabalameli F, et al. High Incidence of Macrolide and Tetracycline Resistance among Streptococcus Agalactiae Strains Isolated from Clinical Samples in Tehran, Iran. Maedica (Bucur). 2014;9(2):157-61.
720. Bennett JE, Dolin R, Blaser MJ. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases E-Book. Elsevier health sciences; 2019.
721. Hopwood DA. Streptomyces in nature and medicine: the antibiotic makers. Oxford University Press; 2007.
722. Ginny C Oong PT. Chloramphenicol. (StatPearls [Internet].):Treasure Island (FL): StatPearls Publishing; 2022 Jan.
723. UK G. Research and analysis Chloramphenicol eye drops containing borax and boric acid buffers: review of the use in children under 2 years.
724. Chan CX, Beiko RG, Ragan MA: Large-scale detection of recombination in nucleotide sequences. In: AIP Conference Proceedings. vol. 971: American Institute of Physics; 2008: 101-4.
725. Slotved HC, Elliott J, Thompson T, Konradsen HB. Latex assay for serotyping of group B Streptococcus isolates. Journal of clinical microbiology. 2003;41(9):4445-7; doi: 10.1128/jcm.41.9.4445-4447.2003.

726. Chen X, Zhang Y, Zhang Z, Zhao Y, Sun C, Yang M, et al. PGAweb: a web server for bacterial pan-genome analysis. Front Microbiol. 2018;9:1910.

727. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proceedings of the National Academy of Sciences of the United States of America. 2005;102(39):13950-5; doi: 10.1073/pnas.0506758102.

728. 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.

729. Heath PT. Status of vaccine research and development of vaccines for GBS. Vaccine. 2016;34(26):2876-9.

730. Brueggemann AB, Pai R, Crook DW, Beall B. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. PLoS Pathog. 2007;3(11):e168.

731. Jannati E, Roshani M, Arzanlou M, Habibzadeh S, Rahimi G, Shapuri R. Capsular serotype and antibiotic resistance of group B streptococci isolated from pregnant women in Ardabil, Iran. Iranian journal of microbiology. 2012;4(3):130.

732. Eskandarian N, Ismail Z, Neela V, Van Belkum A, Desa M, Nordin SA. Antimicrobial susceptibility profiles, serotype distribution and virulence determinants among invasive, non-invasive and colonizing Streptococcus agalactiae (group B streptococcus) from Malaysian patients. European Journal of Clinical Microbiology & Infectious Diseases. 2015;34(3):579-84.

733. Okura M, Auger J-P, Shibahara T, Goyette-Desjardins G, Van Calsteren M-R, Maruyama F, et al. Capsular polysaccharide switching in Streptococcus suis modulates host cell interactions and virulence. Scientific reports. 2021;11(1):1-14.