A 15-year single-centre clinical and genomic analysis of late- and very late-onset Group B Streptococcus infection.

Carl Britto[1], Léa Cavalli[2], Madikay Senghore[2], Alexander McAdam[3], William P. Hanage[2], Ying-Jie Lu[4], and Richard Malley[4]

[1]Division of Critical Care, Department of Anaesthesiology, Critical Care and Pain Medicine Boston Children's Hospital.
[2]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
[3]Infectious Diseases Diagnostic Laboratory, Department of Laboratory Medicine, Boston Children's Hospital.
[4]Division of Infectious Diseases, Boston Children's Hospital and Harvard Medical School.

Abstract

Background: Group B Streptococcus (GBS) is a major cause of meningitis in infants, often prompting intensive care. The factors leading to severe infant disease and the differences between Late-Onset Disease (LOD) and Very-Late-Onset Disease (VLOD) are poorly understood.

Methods: We characterized invasive GBS isolates from patients at Boston Children's Hospital over 15 years using whole-genome sequencing. We compared isolate diversity with samples representing both global and national contexts, and examined vaccine coverage and antimicrobial resistance. Logistic regression and linear mixed models evaluated the relationships between clinical presentations or pathogen virulence factors and disease severity or age of onset, accounting for population structure and multiple testing.

Results: In the 87 patients studied, 44.3% needed ICU care, and 18.6% had meningitis. Among GBS, we identified 5 serotypes and 6 clonal complexes, with hypervirulent clones CC17/cpsIII and CC23/cpsIa being predominant. All isolates contained vaccine candidate targets and were susceptible to penicillin and vancomycin, though many were resistant to erythromycin and clindamycin. ICU admission correlated with specific hematological abnormalities, but meningitis did not. No known virulence factors were associated with ICU or meningitis. LOD and VLOD did not differ bacteriologically or clinically, though the PI-2A1 pilus was linked with a higher likelihood of VLOD.

Conclusions: Our study offers insights into GBS LOD and VLOD. The frequent ICU admissions and cases of meningitis underscore significant morbidity. Clinically and bacteriologically, LOD and VLOD are similar, with potential differences linked to varying immune status with age. This highlights the need for further research into the host-pathogen interactions underlying infection.

1 Introduction

Streptococcus agalactiae, commonly known as Group B Streptococcus (GBS), is a pathobiont that can transition from an asymptomatic colonizer of the vaginal or gastrointestinal tract to a virulent pathogen, causing severe invasive infections. [1] It is a leading cause of sepsis and meningitis among neonates and infants globally. [1,2] In the United States, the Centers for Disease Control and Prevention (CDC) reports that GBS infections occur in approximately 0.4 out of every 1,000 infants under three months of age, with potential outcomes including mortality or long-term complications such as developmental delays, hearing loss, or cerebral palsy. [3, 4]

Neonatal GBS disease is classified as early-onset disease (EOD) when it is diagnosed within the first seven days of life. After this period, infant infections are categorized as late-onset disease

(LOD) if diagnosed between seven days and three months, or as very late-onset disease (VLOD) when occurring between three months and one year of age. While both EOD and LOD are well-documented and associated with significant morbidity and mortality, VLOD remains comparatively understudied. It is unclear whether LOD and VLOD are genuinely distinct entities, though reports suggest they may be clinically and bacteriologically similar. [5,6] Additionally, the mechanisms of acquisition of both LOD and VLOD are not fully understood. EOD results from vertical transmission during birth via a colonized vaginal canal and is effectively prevented through intrapartum antibiotic prophylaxis (IAP) for GBS-positive pregnant women. In contrast, LOD and VLOD incidences have not been affected by IAP implementation in Europe and North America, allowing these later-onset infections to surpass EOD in frequency, despite historically being less common. [7,8] This suggests that LOD and VLOD might be acquired through similar pathways, such as postpartum acquisition from maternal sources (including mucosal colonization and breast milk), as well as hospital and community infections [9]. Despite these similar potential acquisition niches, changes in immune status during infancy may alter susceptibility to different strains. A deeper understanding of the clinical manifestations, acquisition pathways, and bacterial etiologies of LOD and VLOD is needed for their effective prevention and treatment. Notably, VLOD is not addressed in U.S. guidelines for managing GBS infections in infants, which indicates potential differences in clinical management. [10] Furthermore, several vaccine candidates are under development, with ongoing research to identify correlates of protection. [11,12]

The substantial clinical burden of GBS meningitis, especially its long-term neurodevelopmental impacts, is well-documented. [13–20] Global trends in molecular epidemiology identify serotype III, which belongs to the hypervirulent clonal complex 17 (CC17), as the most prevalent serotype in cases of meningitis and LOD. [7,13,21] Despite this, the specific molecular mechanisms that underlie neurotropism remain unclear. Data scarcity also limits the understanding of the molecular basis of other clinical features of acute infection in infants beyond meningitis, such as ICU admission rates, hematological indices, and the onset of VLOD. Understanding these factors is essential for the development of effective immunoprevention strategies to manage invasive disease.

In this study, we characterize the clinical and molecular characteristics of the neonatal invasive GBS disease cases recorded at Boston Children's Hospital (BCH), Massachusetts, between 2007 and 2021. Our dataset notably provides comprehensive clinical details for both LOD, and the less-researched VLOD. We employ whole-genome sequencing to examine the diversity of GBS isolates, estimate the potential coverage of vaccine candidates, and assess antimicrobial resistance. Furthermore, we examine a twin pair providing insights into possible VLOD transmission pathways. Finally, we investigate the association between clinical presentations or molecular factors and well-defined clinical outcomes among infants, such as age at disease onset, risk of ICU admission, and the occurrence of meningitis.

2 Methods

2.1 Ethics statement

Ethical approval was obtained from the BCH IRB (R31579/CN007). The study adhered to the STROBE guidelines, as well as ethical principles of the Helsinki declaration. [22]

2.2 Data Collection

We obtained 100 GBS isolates from patients at Boston Children's Hospital (BCH) between 2007 and 2021. Located in Boston, Massachusetts, BCH is a tertiary referral pediatric hospital serving a catchment area encompassing the majority of New England's 72,000 square miles. The hospital admits a diverse range of patients, including neonates with suspected sepsis. A scarcity of EOD was observed, which aligns with the absence of labor and delivery services at BCH. Consequently, our analysis primarily focuses on LOD and infections occurring later in infancy, notably VLOD. *[Carl/Rick to include statement about why we also observe some adult patients in our cohort]* Isolates were obtained from either the blood or cerebral spinal fluid (CSF) of patients. The presence of bacteria in the blood indicated bacteremia, while its presence in the CSF, characterized by pleocytosis, indicated meningitis. Each isolate was stored for whole-genome sequencing, and the associated demographic, antimicrobial sensitivity profile, laboratory, and clinical information were extracted from hospital records. The recorded variables included gestational age, sex, length of stay, vitals, clinical presentation, hematological indices, medications, and interventions. *[Carl/Rick: length of stay, vitals, clinical presentation/symptoms, medications, and interventions are not included in the current data sheet for publication - would you like to include them, and describe them in*

the paper? Or exclude them ?] The hematological indices recorded included white blood cell (WBC) count, hemoglobin level, platelet count, and absolute neutrophil count (ANC). Hematological data were entirely missing for two patients, and ANC data were missing for an additional patient. We treated the missing data as random and excluded these patients in statistical analyses involving the missing variables. For all patients except one preterm infant (Patient#86), gestational age coincided with age in days since birth. We also documented the level of care received by each patient, distinguishing between Intensive Care Unit (ICU) and inpatient pediatric unit admissions. To exclude precautionary admissions, ICU admissions were defined as stays exceeding 48 hours.

2.3 Whole-genome sequencing and Antibiotic susceptibility testing

[Carl/Rick: Please add details on the experimental methods used for sequencing and AST]

2.4 Genomic analysis

Trimmomatic (v.0.39) was used to remove sequencing adapters from raw reads, followed by FastQC (v0.12.1) and Kraken (v.1.0) on trimmed reads to verify low adapter content and identify contamination with the miniKraken database (minikraken_20171019_8GB), respectively. [23–25] We eliminated 4 isolates that had less than 50% of reads classified as 'Streptococcus agalactiae' (GBS). Srst2 (v.0.2.0) and the Sanger GBS-Typer (v.1.0.12) tool were used on trimmed reads to obtain the serotype and Sequence Type (ST) of each isolate, and to identify the presence of genes encoding surface proteins and antibiotic resistance. [26, 27] The ST of each isolates was then translated into its corresponding clonal complex (CC) based on known ST-CC matches available in the PubMLST Streptococcus agalactiae typing database. [28] Unicycler (v0.4.8) was used for genome assembly, followed by QUAST to check the quality of assemblies. [29, 30] All our isolates passed the QC criteria, since they had fewer than 250 contigs, N50 $\geq$ 30,000, genome lengths between 1.7m and 2.4m, and a GC content between 30% and 40%. Abricate (v.1.0.1) was used on assembled genomes to detect virulence factor genes with the vfdb databases and the SIP gene with a custom database. [31] Snippy (v.4.6.0) was used to obtain a whole genome SNP alignment, which was input into RAxML (v8.2.12) to infer a Maximum likelihood (ML) phylogenetic tree. [32, 33] Inference was conducted under the generalized time-reversible model with a Gamma distribution (GTR + Γ) to model site-specific rate variation and with 100 bootstrap replicates to support results. [34] A Streptococcus pyogenes (Group A Streptococcus, GAS) strain (ENA accession: SRR1104967) was used as an outgroup to root the inferred phylogenetic tree of the GBS isolates. [35] The resulting tree was visualized and annotated using the R package ggtree (v3.8.2). [36] GBS was isolated from both the blood and CSF in 9 patients, yielding identical genome sequences. We therefore removed duplicate strains from the genomic analysis, leaving us with a total of 87 samples for analysis. For the pair of VLOD twins identified in our dataset, the SNP-distances between their isolates were calculated with *snp-dists*. [37] Finally, the BCH LOD strains (n=48) were contextualized at the national and global level using publicly available LOD genomes from the US CDC (n=267, project accession: PRJNA355303) and other global cohorts: Netherlands (n=175, project accession: PRJEB14124), Malawi (n=47, project accession: PRJEB8986), Canada (n=34, project accession: PRJNA295774), Ireland (n=16, project accession: PRJEB26339). [4, 21, 38, 39] A global phylogeny was built with RAxML to compare the phylogeographic distribution of samples. [40] The code for bioinformatic steps is available in GitHub at: https://github.com/Leacavalli/BCH-GBS.

2.5 Analysis of Clinical and Molecular Correlates of Severe Disease

We compared the clinical presentation of LOD and VLOD by examining ICU admission rates and the incidence of meningitis. We also analyzed hematological indices of infant infections across age of disease onset (LOD vs. VLOD), level of care (ICU vs. inpatient pediatric unit), and meningitis status. The hematological indices analyzed included WBC count, hemoglobin level, ANC, and platelet counts, along with the occurrence of leukopenia, leukocytosis, and neutropenia. To classify leukopenia, leukocytosis, and neutropenia, we binarized the WBC and ANC values using infant-specific thresholds: neutropenia was defined as ANC $< 1.50 \times 10^3$ cells/mm^3, leukocytosis as WBC count $\geq 10.50 \times 10^3$ cells/mL, and leukopenia as WBC count $< 3.50 \times 10^3$ cells/mL. Logistic regression was employed to calculate odds ratios and 95% confidence intervals for each hematological indicator. The Bonferroni correction was applied for multiple testing, setting a significance threshold of 0.007 (0.05/7) for the seven hematological indicators.

We conducted a genetic association study to identify molecular risk factors associated with clinical outcomes. The outcomes of interest included the age of disease onset and severity indicators

such as ICU admission, meningitis, and abnormal hematological indices, including leukopenia, leukocytosis, and neutropenia. Due to limitations in statistical power arising from sample size, the study concentrated on known virulence genes instead of conducting a genome-wide SNP analysis or evaluating all genes in the pangenome. We used a linear mixed model in Pyseer to examine gene-phenotype associations incorporating patristic distances from the RAxML phylogeny to account for population structure. [41] A Bonferroni correction was applied to address multiple testing, establishing a significance threshold of 0.003 (0.05/17) to evaluate associations with 17 virulence factors. Data analysis and visualizations were performed using R (V.4.3.1), and the corresponding code is available on GitHub at https://github.com/Leacavalli/BCH-GBS. [42]

3 Results

3.1 Molecular Surveillance of Invasive GBS

Our final dataset comprised 96 clinical isolates collected from 87 patients with invasive GBS infections between 2007 and 2021. In 72 patients, GBS was isolated from the blood only; in 6 patients, it was found in the cerebrospinal fluid (CSF) only; and in 9 cases, it was isolated from both sources. In total, 70 infants, defined as children under one year old, were diagnosed with invasive GBS infection. Two patients were diagnosed within 7 days of birth (EOD), 48 between 7 days and 3 months (LOD), and 20 from 3 to 12 months (VLOD) (Figure 1, Table S1). The remaining 17 isolates were collected from 10 older children and 7 adults. Throughout the study period, serotype III was the most prevalent in infants (n=45, 64.3%) followed by serotype Ia (n=15, 21.4%), which was also predominant among older children aged 1 to 18 years (n = 5, 50%) (Figure 1, Table S1). In this sample, no isolates with capsular serotypes VI –IX and no non-typeable isolates were recovered.

We identified six distinct Clonal Complexes (CCs) among the invasive GBS isolates: CC1 (N=4, 4.6%), CC12 (N=8, 9.2%), CC17 (N=41, 47.1%), CC19 (N=9, 10.3%), CC23 (N=21, 24.1%) and C459 (N=4, 4.6%) (Figures 1, Table S2). These were further divided into 18 known Sequence Types (STs), with ST-17 being the most frequent (n=33, 37.9%), followed by ST-23 (n=17, 19.5%) and ST-19 (n=5, 5.7%) (Table S2). Additionally, a new MLST profile that had not been previously reported in the pubMLST database was discovered: adhP(4) pheS(1) atr(4) glnA(4) sdhA(3) glcK(3) tkt(2). This new ST differed from ST-8 by one housekeeping gene (glnA), and from ST-10 by two housekeeping genes (adhP and glnA), classifying it within CC12 (Table S3). With one exception, each ST was associated with a single serotype, and each CC represented by a dominant serotype (Figure 2). Specifically, CC17 was exclusively serotype III and CC23 was exclusively serotype Ia. Finally, we found evidence of capsule switching in CC19 and CC12. In CC19, ST-19 and ST-335 were serotype III while ST-28 was serotype II. CC1/ST-1 was serotype V, and ST-459 was serotype IV. All CC12 isolates were serotype Ib, including the newly identified sequence type, with the exception of one ST-10/cps II isolate.

The distribution of clonal complexes (CCs) and serotypes varied by age at infection, with infants differing significantly from older patients. Specifically, CCs differed between LOD and older children, LOD and adults, and VLOD and adults, with Fisher's test FDR-adjusted p-values of 0.00123 for each comparison (refer to Figure 1 and Table S1). There were also serotype differences between LOD and older children (p=0.00176) and between LOD and adults (p=0.03630). Although CC17/cps-III was more prevalent in LOD isolates (65%) compared to VLOD isolates (45%), and CC23/cps-Ia showed the opposite trend (LOD: 15%, VLOD: 30%), no significant difference was observed in the CC distribution between LOD and VLOD (p=0.0945). Additionally, no clonal expansion over time was detected (X-squared = 73.483, p-value = 0.3648), and the serotype distribution remained stable as well (X-squared = 75.915, p-value = 0.2937) (Figure 1). These results indicate a stable diversity within the invasive GBS population over the study period, potentially reflecting the absence of selective pressures from clinical interventions targeting specific clones or capsular types.

Finally, the genetic diversity of LOD isolates in our study was consistent with that observed in globally and nationally representative GBS populations (Figure **??**). The isolates exhibited polyphyletic structures similar to those seen across broader geographical scales, encompassing various clonal complexes representative of distinct genotypes. Furthermore, there was no evidence of genetic isolation by geographic location within each clonal complex, suggesting a history of recombination and frequent international migration rather than localized transmission clusters.

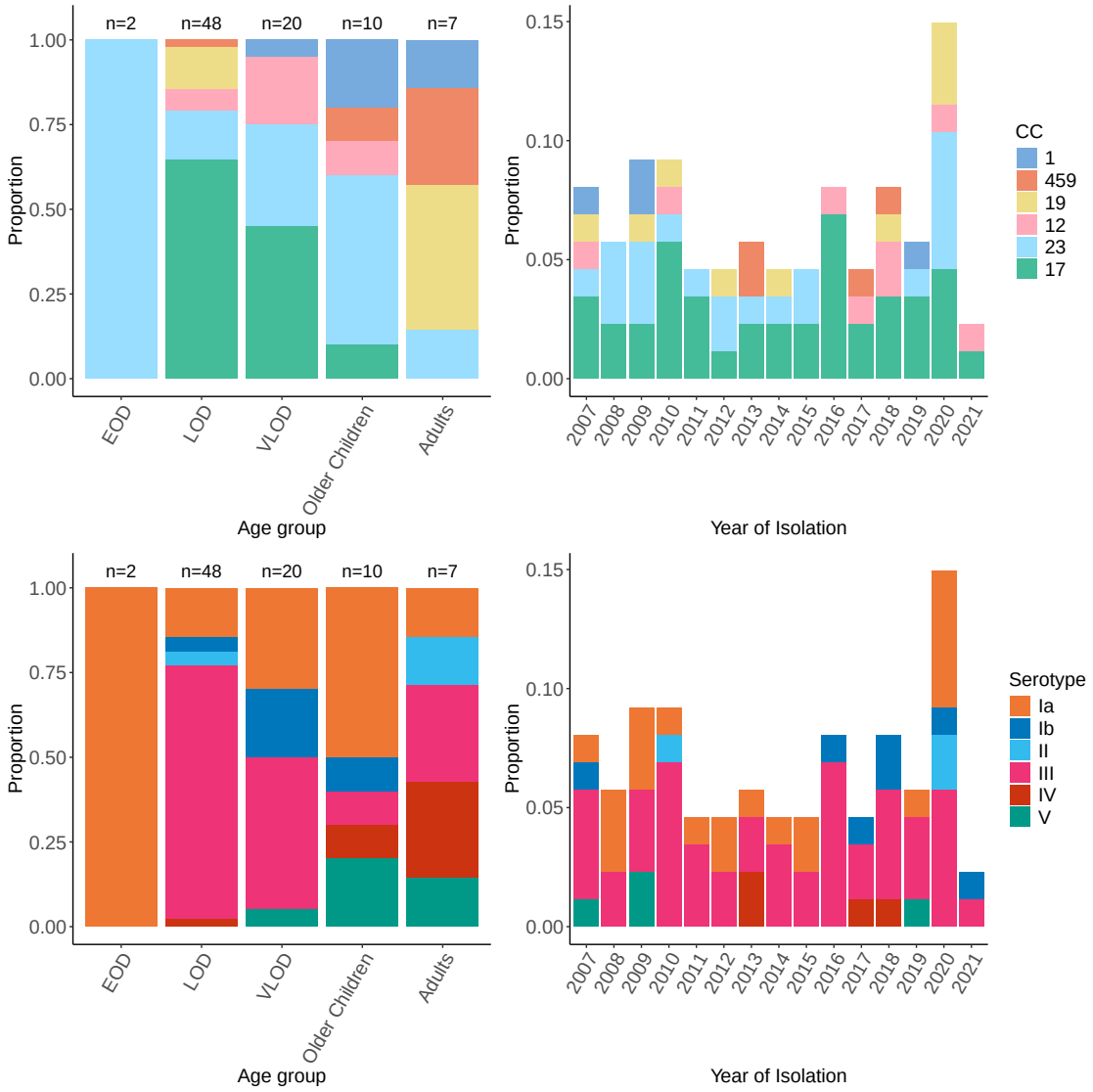

Figure 1. Distribution of group B Streptococcus (GBS) clonal complexes (CCs) and serotypes across age groups and over time. The top panel displays the proportions of CCs across different age groups (left) and annually from 2007 to 2021 (right), while the bottom panel presents the corresponding distributions for serotypes. The age groups are defined as follows: early-onset disease (EOD) is diagnosed within 7 days of birth, late-onset disease (LOD) between 7 days and 3 months, very late-onset disease (VLOD) between 3 months and 1 year, older children range from 1 to 18 years, and adults are 18 years and older.

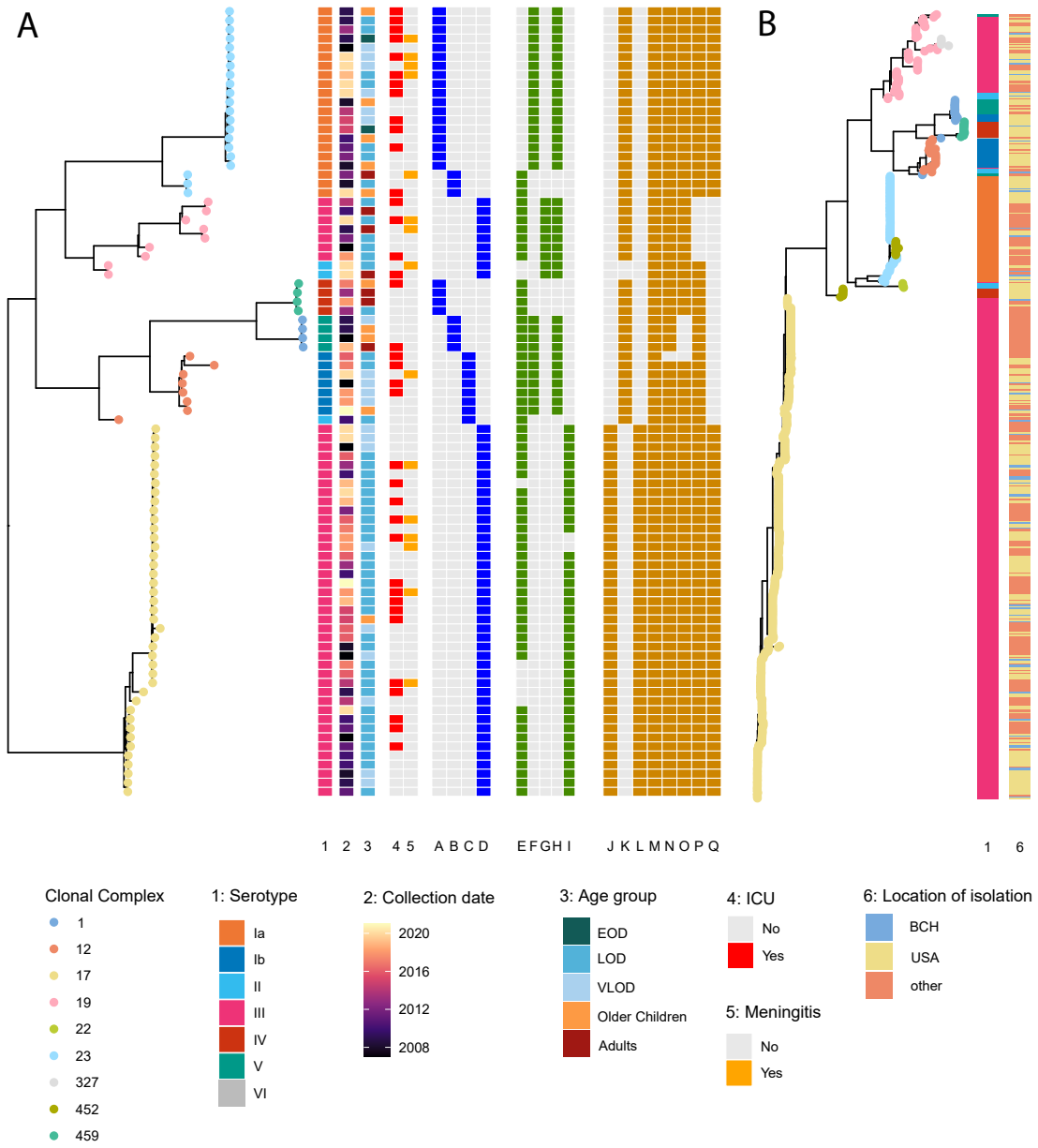

Figure 2. Phylogenies of Boston Children's Hospital (BCH) and Global Group B Streptococcus (GBS) isolates. Panel A presents the phylogenetic tree of Group B Streptococcus (GBS) isolates from Boston Children's Hospital (BCH), categorized by age groups: early-onset disease (EOD, diagnosed within 7 days of birth), late-onset disease (LOD, 7 days to 3 months), very late-onset disease (VLOD, 3 months to 1 year), older children (1 to 18 years), and adults (18 years and older). Various virulence factors and surface proteins are detailed, including Alpha-like proteins (ALP1, ALP23, Alpha, RIB), pilus islands (PI-1, PI-2a1, PI-2a2, PI-2b), and other factors such as the hypervirulence gene cluster A (HVGA), serine-rich repeat proteins (SRR1, SRR2), Sip, laminin-binding protein (lmb), C5a peptidase (scpB), hyaluronidase (hylB), and fibrinogen-binding protein (fbsB). Panel B extends the context by situating BCH LOD isolates (in blue on heatmap 6) within a broader phylogenetic framework, including national and global LOD isolates. It features isolates from the USA (yellow on heatmap 6), gathered through the CDC's ABCs program, and from other international sources such as Ireland, Malawi, Canada, and The Netherlands (orange on heatmap 6).

3.2 Coverage of genes encoding vaccine targets

In assessing the potential coverage of the two main candidate GBS vaccines, the hexavalent vaccine (GBS6 by Pfizer) targeting capsular polysaccharides (cps) of serotypes Ia, Ib, II, III, IV and V would theoretically cover 100% of infant cases in our study (Figure 1, Table S4) [11]. The other leading vaccine candidate (GBS-NN2 by Minervax) contains fusion proteins from the alpha-like protein (Alp) family of GBS. [12] All infant isolates in our study contained genes encoding at least one Alp protein: rib (n=46, 66%), alphaC (n=7, 10%), alp1 (n=15, 21%), and alp23 (n=2, 3%) (Figure 2, Table S4). The rib gene was exclusively found in CC17 (n=40) and CC19 (n=6) isolates, alphaC in CC12 (n=7), alp1 in CC23 (n=14) and CC459 (n=1), and alp23 in CC23 (N=1) and CC1 (n=1) (Figure 2). Although GBS maternal vaccines are primarily designed to prevent EOD and LOD, our findings suggest that GBS6 and GBS-NN2 could also effectively cover the VLOD cases commonly seen in our cohort, provided they induce long-lasting maternal antibodies (Figure 1, Table S4). The capsular types and proteins targeted by these vaccines are present in all isolates from older children and adults in our dataset, indicating their potential as suitable candidates for adult vaccination in outbreaks or other high-risk scenarios. [43] For example, this includes the 2015 outbreak in Singapore caused by ST283/cpsIII in non-pregnant adults linked to the consumption of raw fish. [44]

We also examined the distribution of other immunogenic proteins encoded by isolates in our study cohort. All infant isolates encoded at least one of the three pilus islands: PI-1 (n=49, 70%), PI-2a (n=27, 39%) and PI-2b (n=38, 54%), as well as the sip gene (Table S4). These immunogenic proteins have the potential to serve as universal vaccine targets. In contrast, not all infant isolates carried genes encoding C5a peptidase (n=68, 97%), Lmb (n=69, 99%), and FbsB (n=55, 79%) (Table S4). The FbsB protein was exclusively present in hypervirulent clones CC17 (n=40, 100%) and CC23 (n=10, 100%) isolates, while absent from all other clonal complexes (Figure 2).

3.3 Linking Phenotypic Antibiotic Resistance to Genotypic Markers

Phenotypic antimicrobial resistance (AMR) testing showed all isolates were susceptible to penicillin, the preferred treatment for GBS infections, and to vancomycin, which is recommended for those with penicillin allergies and clindamycin resistance. Thirty-three (38%) isolates exhibited resistance to erythromycin, twenty-three (26%) were resistant to clindamycin, and three (3%) had intermediate resistance to clindamycin. Genes encoding macrolide resistance (ermA, ermB, ermT, and mefA-msrD) were identified in all erythromycin-resistant isolates and in 4 (7%) erythromycin-susceptible isolates. Genes linked to lincosamide resistance (ermA, ermB, and ermT) were found in all clindamycin-resistant isolates, one of the three clindamycin-intermediate isolates, and seven of the 61 that were clindamycin-susceptible. Despite overlap between phenotypic and genotypic resistance profiles, the presence of some false negatives suggests limitations in phenotypic resistance testing, which may be due to the subjective nature of Minimum Inhibitory Concentration (MIC) breakpoints. This underscores the value of in-silico resistance gene detection and the necessity of ongoing phenotypic resistance testing and regular updates to online databases.

We also assessed the presence of genetic resistance markers associated with resistance to other antibiotics (Table S5, Fig S1). Tetracyclin resistance predicted in 78 isolates (90%) in our study from the presence of tet-M ($n = 76$, 87%), tet-O ($n = 2$, 2%), or both ($n = 6$, 7%). The highest rates of tetracycline resistance was observed in CC1 ($n = 4/4$, 100%) and CC17 ($n = 39/40$, 98%) compared to other CCs (CC12: $n = 6/8$, 75%; CC459: $n = 2/4$, 50%; CC19: $n = 8/9$, 89%; CC23: $n = 18/21$, 86%). Aminoglycoside resistance was predicted in 4 isolates (5%) due to the presence of the aph(3')-III and ant(6) resistance genes, all belonging to CC17/ST-17. No SNP variants conferring reduced penicillin susceptibility were found in the rRNA 23-S1 and 23S-3 genes. However, gyrA or parC SNP variants, typically associated with fluoroquinolone resistance, were detected in three isolates: two CC12 isolates carried the S81L mutation in GyrA and S79F in ParC, and one CC17 isolate carried the S79F mutation in ParC alone.

In 2010, the GBS AIP guidelines for pregnant individuals were revised, notably removing erythromycin due to concerns about resistance. In light of this change during the study period, we evaluated the temporal changes in genetic elements that confer resistance (see Figure S1). We found that isolates carrying the aph(3')III/ant(6) genes only appeared from 2014 onwards. In contrast, the mef(A)-msr(D) genes, which confer resistance to erythromycin, were absent after 2013. Meanwhile, the frequency of erm genes, associated with resistance to both clindamycin and erythromycin, increased from 29% to 41%.

3.4 A Case of Twins with VLOD

In 2020, a pair of twins presented to the emergency department at BCH with symptoms of fever and seizures. Both were diagnosed with meningitis after GBS was isolated from their CSF. Due to the severity of symptoms, one twin was admitted to the Intensive Care Unit (ICU) and treated with ceftriaxone, vancomycin, and penicillin, while the other received care in the inpatient unit with ceftriaxone and penicillin G. [*Rick/Carl: Do you think the clinical outcomes (E.g. Recovery, complications) of the pair of twins (S92, S93, S94) would be of interest to the reader ? If so, please include a sentence or two.*]

These cases were classified as VLOD, with symptoms emerging 107 days after birth. Isolates from both twins were identified as the CC23/cpsIa strain, which is the second most common strain in neonatal invasive GBS disease, following the hypervirulent CC17/cpsIII strain. Phylogenetic analysis revealed the isolates were closely related, differing by only 5 single nucleotide polymorphisms (SNPs) (Figure S2). Consistent with this, they encoded identical virulence factors (ALP1, PI2A1, SRR1), carried the same antibiotic resistance markers (TET-M and TET-SM, conferring resistance to tetracyclines), and were deemed susceptible to penicillin, vancomycin, erythromycin, and clindamycin in antimicrobial susceptibility testing.

The genetic similarity of the isolates implies a common source of infection. Neither twin had syndromic or anatomical anomalies, nor any known immunodeficiency that might elevate their risk of infection. Furthermore, both had negative neonatal GBS screening results. This, combined with the simultaneous onset of symptoms, suggests that early intestinal colonization followed by translocation and invasion, or transmission between twins, was unlikely. Instead, it supports the hypothesis of simultaneous acquisition from an external source, such as community or enteral transmission, similar to other cases reported in LOD twins. [45]

3.5 Clinical and Molecular Correlates of Infant Severe Disease

Among the infant cases, 31 (44.3%) required ICU admission due to risks of clinical decompensation, mental status changes or escalating respiratory support. Analysis indicated that higher counts of white blood cells (WBC), platelets, and absolute neutrophils (ANC) were significantly correlated with reduced odds of ICU admission (WBC: OR=0.84, 95% CI=0.75-0.93, p-value=0.001, platelet: OR= 0.99, 95% CI=0.99-1, p-value=0.003; ANC: OR= 0.83, 95% CI= 0.73-0.94, p-value= 0.004; Table S6). In line with this, leukopenia was associated with significantly higher odds of ICU admission (OR= 6.5, 95% CI= 1.84-22.98, p-value= 0.004), while leukocytosis showed a protective effect (OR= 0.1, 95% CI= 0-0.5, p-value= 0.002). [*Carl/Rick: are these hemotological indices expected to be significantly associated with the odds of ICU admission ? Does it reflect any specific guidelines ?*] Meningitis was documented in thirteen (18.6%) infants, but no hematological indices were associated with meningitis after multiple testing corrections (Table S6). There were no significant differences in the odds of ICU admission and meningitis odds between LOD and VLOD patients (ICU admission: OR= 3, 95% CI=0.9-9.6, p-value=0.063; meningitis: OR= 0.8 , 95% CI=0.2-3, p-value=0.743). Similarly, no significant differences were observed in any hematological indices (Hb, WBC, platelet, ANC counts, or leukocytosis, leukopenia, neutropenia) between LOD and VLOD patients (Table S6).

We initially found that RIB, PI2B, SRR2, Sip.3a, HVGA, and FsbB were significantly associated with an increased odds of disease in infants compared to older patients (older children and adults), while ALP23, SRR1, and Sip1A were linked to decreased odds (Table S7). However, these associations were not statistically significant after adjusting for population structure, suggesting they were influenced by differences in GBS strain distribution across age groups rather than true causal relationships between the virulence factors and disease onset age. Indeed, clonal complex was significantly associated with infant disease ($\chi^2 = 25.71, p = 0.0001$), with CC17 being linked to increased odds (OR: 120.0, 95 % CI: 5.92 ;2434.31, p-value= 0.0018). Among infants, no virulence factors were associated with ICU admission, meningitis, or abnormal hematological indices (leukopenia, leukocytosis, neutropenia) after accounting for population structure and correcting for multiple tests (Tables S8–S11). Furthermore, these clinical outcomes were not significantly correlated with any particular clonal complex: ICU admission ($\chi^2 = 6.58, p = 0.25$), meningitis ($\chi^2 = 2.39, p = 0.79$), neutropenia ($\chi^2 = 6.09, p = 0.30$), leukopenia ($\chi^2 = 8.72, p = 0.12$), and leukocytosis ($\chi^2 = 3.75, p = 0.59$). Odds of LOD were 0.60 times those of VLOD in the presence of pilus island PI-2A1 (95% CI: 0.42 - 0.85, $p = 0.002$). Although this effect remained significant (p=0.0058) after controlling for population structure, it did not quite meet the stringent significance threshold of 0.003 required by multiple testing correction, despite being close (Table S13). We also evaluated the association between PI-2A1 and continuous age among infants, hoping to

increase statistical power, but found that the effect remained similar (Table S14). Nevertheless, given the conservative nature of the Bonferroni correction and the small size of our cohort, this finding may still indicate a genuine biological effect.

4 Discussion

Our study offers insights into infant GBS disease in a setting not usually captured by national surveillance, focusing on LOD and the relatively understudied VLOD. The diversity observed in clonal complexes and serotypes within our cohort reflects global trends, with hypervirulent clones CC17/cpsIII and CC23/cpsIa being predominant in cases of infant invasive disease. VLOD accounted for one third of the cases in our cohort, showing no significant clinical differences compared to LOD. Additionally, a VLOD case involving twins indicated simultaneous external acquisition, consistent with proposed routes of LOD acquisition. [9, 45, 46] The strain diversity was similar between LOD and VLOD cases, with CC17/cpsIII prevailing in LOD and CC23/cpsIa in VLOD. Interestingly, the presence of PI-2A variant 1 was associated with higher odds of VLOD, which persisted when population structure was taken into account. PI-2A stands out as particularly variable among the three GBS pilus islands, with its backbone proteins BP-2A and AP1-2A comprising seven alleles. [47] In the surface protein typing tool we used, PI-2A1 and PI-2A2 were chosen to maximize PI-2A coverage rather than reflect distinct biological entities with known phenotypes. [48] Nonetheless, they are genomically distinct and could hold bacterial relevance. Given the conservation of BP-2A and AP1-2A alleles, genomic variations among PI-2A subvariants likely do not affect their functionality, suggesting that PI-2A1 might not have unique virulence characteristics. Instead, the observed trend may result from antigenic variation among the subvariants, due to immune selection pressures. For example, previous exposure or waning maternal antibodies could enhance susceptibility distinct antigenic alleles as infants age, similarly to patterns observed with a pilus locus in pneumococci. [49] This highlights the complex interactions between host immunity and bacterial adaptation, warranting further investigation into the distribution of pilus variants across age groups. Aside from these possible age-related changes in exposure and immune landscape, the similarities between LOD and VLOD cases suggest they are likely clinically and bacteriologically similar.

All invasive disease isolates encoded targets of both maternal vaccine candidates currently in development: GBS6 and GBS-NN2. [11, 50] While these vaccines aim to prevent EOD and LOD, the observed potential coverage and the maternal antibodies they induce —reported to persist in infants for over three months — also suggest they could help prevent VLOD. [50, 51] Moreover, the broad coverage of these vaccines in older pediatric and adult patients highlights their value for adult vaccination during outbreaks or in high-risk settings, which is significant given the rise in invasive disease among non-pregnant adults since 1990. [43, 52] Despite this promise, the global diversity of CPS types and Alp sequences raises concerns about universal vaccine coverage. Gaps in coverage due to non-vaccine types and non-typeable isolates underscore the need for post-vaccine GBS population monitoring to guide and inform future clinical interventions. [53–55] Several other surface proteins have been identified as potential vaccine candidates due to their immunogenic nature observed in preclinical research. [56] The presence of Sip and at least one pilus island in all our isolates supports their potential as targets in a universal GBS vaccine. [14] In contrast, C5a, Lmb, FbsB and Srr1/2 proteins were absent from several of our isolates, which may limit their use as vaccine targets.

With regards to antimicrobial resistance, our findings also align with broader trends. We detected no resistance to penicillin and vancomycin, the first and last line antibiotics recommended in AIP, respectively. [57] Similarly, no genetic markers for resistance were detected for other beta-lactams such as ampicillin or cefazolin, also recommended for AIP in certain cases. [57] Conversely, approximately 29% and 40% of isolates were non-susceptible to clindamycin and erythromycin, respectively, supporting the latter's removal from U.S. AIP guidelines in 2010 due to resistance concerns [57] Although not statistically significant, the absence of mef(A)/msr(D) genetic markers and a rise in the prevalence of erm genes after 2013 might have biological and clinical significance. The absence of mef(A)/msr(D), which confers resistance to erythromycin but not clindamycin, may reflect reduced selective pressure for erythromycin resistance following the 2010 revision of AIP guidelines. [57] The ongoing presence of ermA, ermB, and ermT genes, conferring resistance to both erythromycin and clindamycin, aligns with clindamycin's continued use in treating GBS infections. This rise in macrolide and lincosamide resistance is widely reported and is believed to result from clindamycin overuse in penicillin-allergic patients. [57, 58] These findings underscore the importance of clindamycin susceptibility testing in maternal carriage isolates and the use of

vancomycin for resistant strains. Finally, tetracycline exhibits the highest resistance rate in GBS (90% in our study) despite having never been used to treat this infection and rarely been prescribed in the last 20 years. This resistance stems from the widespread use of tetracycline in the 1950s to prevent and treat various infections, leading to lasting changes in the bacterial population and reportedly contributing to the emergence of neonatal GBS in the 1960s. This highlights the need to consider both the direct risks and broader impacts of antibiotics on natural human and animal flora when prescribing them.

Our study highlights the significant morbidity associated with invasive GBS disease in infants, with 44.3% requiring ICU admission and 18.6% experiencing meningitis. Infants requiring ICU admission exhibited distinct hematological profiles, marked by decreased white blood cell, platelet, and neutrophil counts, as well as leukopenia. Despite these findings, we did not identify specific virulence genes linked to ICU admission or these hematological abnormalities. Similarly, while the long-term impacts of meningitis are well-documented, precise host or pathogen risk factors remain unidentified. [13–20] Although previous studies frequently associate meningitis with the CC17/cpsIII strain, this correlation was not observed in our cohort, nor did we find any virulence factors connected to meningitis. [7, 13, 21] The absence of molecular risk factors tied to severe disease may result from the limited power of our study due to a small sample size, which also restricted our capacity to explore associations within the broader accessory genome. Alternatively, it might suggest the influence of host factors not captured in our study or complex host-pathogen interactions in determining disease severity and outcomes. Factors such as the patient's immune status, potential delays in initiating antimicrobial therapy, and other comorbidities likely contribute to clinical manifestations. Although retrospective in nature and reliant on routinely collected data, our study provides valuable insights into the clinical presentation of invasive GBS disease in infants.

References

[1] S. Shabayek and B. Spellerberg, "Group b streptococcal colonization, molecular characteristics, and epidemiology," *Frontiers in microbiology*, vol. 9, p. 437, 2018.

[2] World Health Organization (WHO), "Group b streptococcus infection causes an estimated 150,000 preventable stillbirths and infant deaths every year," 2017. Accessed: 04 December 2024.

[3] Center for Disease Control and Prevention (CDC), "ABCs bact facts interactive data dashboard," 2024. Accessed: 04 December 2024.

[4] L. McGee, S. Chochua, Z. Li, S. Mathis, J. Rivers, B. Metcalf, A. Ryan, N. Alden, M. M. Farley, L. H. Harrison, *et al.*, "Multistate, population-based distributions of candidate vaccine targets, clonal complexes, and resistance features of invasive group b streptococci within the united states, 2015–2017," *Clinical Infectious Diseases*, vol. 72, no. 6, pp. 1004–1013, 2021.

[5] J. B. Cantey, C. Baldridge, R. Jamison, and L. A. Shanley, "Late and very late onset group b streptococcus sepsis: one and the same?," *World Journal of Pediatrics*, vol. 10, pp. 24–28, 2014.

[6] A. W. Bartlett, B. Smith, C. R. George, B. McMullan, A. Kesson, M. M. Lahra, and P. Palasanthiran, "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*, vol. 36, no. 1, pp. 20–24, 2017.

[7] C. R. Phares, R. Lynfield, M. M. Farley, J. Mohle-Boetani, L. H. Harrison, S. Petit, A. S. Craig, W. Schaffner, S. M. Zansky, K. Gershman, *et al.*, "Epidemiology of invasive group b streptococcal disease in the united states, 1999-2005," *Jama*, vol. 299, no. 17, pp. 2056–2065, 2008.

[8] S. A. Nanduri, S. Petit, C. Smelser, M. Apostol, N. B. Alden, L. H. Harrison, R. Lynfield, P. S. Vagnone, K. Burzlaff, N. L. Spina, *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 pediatrics*, vol. 173, no. 3, pp. 224–233, 2019.

[9] F. Miselli, I. Frabboni, M. Di Martino, I. Zinani, M. Buttera, A. Insalaco, F. Stefanelli, L. Lugli, and A. Berardi, "Transmission of group b streptococcus in late-onset neonatal disease: A narrative review of current evidence," *Therapeutic Advances in Infectious Disease*, vol. 9, p. 20499361221142732, 2022.

[10] K. M. Puopolo, R. Lynfield, J. J. Cummings, I. Hand, I. Adams-Chapman, B. Poindexter, D. L. Stewart, S. W. Aucott, J. P. Goldsmith, M. Mowitz, *et al.*, "Management of infants at risk for group b streptococcal disease," *Pediatrics*, vol. 144, no. 2, 2019.

[11] J. Absalon, N. Segall, S. L. Block, K. J. Center, I. L. Scully, P. C. Giardina, J. Peterson, W. J. Watson, W. C. Gruber, K. U. Jansen, *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*, vol. 21, no. 2, pp. 263–274, 2021.

[12] C. Banks, B. J. Lindbom, G. Kitson, M. Darsley, and P. B. Fischer, "Preclinical development of a novel group b streptococcus (gbs) vaccine candidate for maternal immunization based upon the alpha-like protein family of gbs surface proteins (alp)," *Birth Defects Research*, vol. 115, no. 9, pp. 933–944, 2023.

[13] M. N. van Kassel, G. de Boer, S. A. Teeri, D. Jamrozy, S. D. Bentley, M. C. Brouwer, A. van der Ende, D. van de Beek, and M. W. Bijlsma, "Molecular epidemiology and mortality of group b streptococcal meningitis and infant sepsis in the netherlands: a 30-year nationwide surveillance study," *The lancet microbe*, vol. 2, no. 1, pp. e32–e40, 2021.

[14] E. Horváth-Puhó, M. N. van Kassel, B. P. Gonçalves, B. de Gier, S. R. Procter, P. Paul, A. van der Ende, K. K. Søgaard, S. J. Hahné, J. Chandna, *et al.*, "Mortality, neurodevelopmental impairments, and economic outcomes after invasive group b streptococcal disease in early infancy in denmark and the netherlands: a national matched cohort study," *The Lancet Child & Adolescent Health*, vol. 5, no. 6, pp. 398–407, 2021.

[15] J. Bramugy, H. Mucasse, S. Massora, P. Vitorino, C. Aerts, I. Mandomando, P. Paul, J. Chandna, F. Seedat, J. E. Lawn, *et al.*, "Short-and long-term outcomes of group b streptococcus invasive disease in mozambican children: results of a matched cohort and retrospective observational study and implications for future vaccine introduction," *Clinical Infectious Diseases*, vol. 74, no. Supplement_1, pp. S14–S23, 2022.

[16] J. Chandna, W.-H. Liu, Z. Dangor, S. Leahy, S. Sridhar, H. B. John, H. Mucasse, Q. Bassat, A. Bardaji, A. Abubakar, *et al.*, "Emotional and behavioral outcomes in childhood for survivors of invasive group b streptococcus disease in infancy: findings from 5 low-and middle-income countries," *Clinical Infectious Diseases*, vol. 74, no. Supplement_1, pp. S35–S43, 2022.

[17] P. Paul, J. Chandna, S. R. Procter, Z. Dangor, S. Leahy, S. Santhanam, H. B. John, Q. Bassat, J. Bramugy, A. Bardají, *et al.*, "Neurodevelopmental and growth outcomes after invasive group b streptococcus in early infancy: A multi-country matched cohort study in south africa, mozambique, india, kenya, and argentina," *EClinicalMedicine*, vol. 47, 2022.

[18] L. M. Harden, S. Leahy, S. G. Lala, P. Paul, J. Chandna, S. Lowick, S. Mbatha, T. Jaye, B. Laughton, A. Ghoor, *et al.*, "South african children: a matched cohort study of neurodevelopmental impairment in survivors of invasive group b streptococcus disease aged 5 to 8 years," *Clinical Infectious Diseases*, vol. 74, no. Supplement_1, pp. S5–S13, 2022.

[19] H. B. John, A. Arumugam, M. Priya, N. Murugesan, N. Rajendraprasad, G. Rebekah, P. Paul, J. Chandna, J. E. Lawn, and S. Santhanam, "South indian children's neurodevelopmental outcomes after group b streptococcus invasive disease: a matched-cohort study," *Clinical Infectious Diseases*, vol. 74, no. Supplement_1, pp. S24–S34, 2022.

[20] B. P. Gonçalves, S. R. Procter, P. Paul, J. Chandna, A. Lewin, F. Seedat, A. Koukounari, Z. Dangor, S. Leahy, S. Santhanam, *et al.*, "Group b streptococcus infection during pregnancy and infancy: estimates of regional and global burden," *The Lancet Global Health*, vol. 10, no. 6, pp. e807–e819, 2022.

[21] C. Chaguza, D. Jamrozy, M. W. Bijlsma, T. W. Kuijpers, D. van de Beek, A. van der Ende, and S. D. Bentley, "Population genomics of group b streptococcus reveals the genetics of neonatal disease onset and meningeal invasion," *Nature communications*, vol. 13, no. 1, p. 4215, 2022.

[22] W. M. Association *et al.*, "World medical association declaration of helsinki: ethical principles for medical research involving human subjects," *Jama*, vol. 310, no. 20, pp. 2191–2194, 2013.

[23] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: a flexible trimmer for illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, 2014.

[24] D. E. Wood and S. L. Salzberg, "Kraken: ultrafast metagenomic sequence classification using exact alignments," *Genome biology*, vol. 15, pp. 1–12, 2014.

[25] S. Andrews *et al.*, "Fastqc: a quality control tool for high throughput sequence data," 2010.

[26] M. Inouye, H. Dashnow, L.-A. Raven, M. B. Schultz, B. J. Pope, T. Tomita, J. Zobel, and K. E. Holt, "Srst2: Rapid genomic surveillance for public health and hospital microbiology labs," *Genome medicine*, vol. 6, pp. 1–16, 2014.

[27] V. Dyster, H. Hung, and K. Pepper, "Gbs-typer-sanger-nf." https://github.com/sanger-bentley-group/GBS-Typer-sanger-nf, 2022.

[28] K. A. Jolley, J. E. Bray, and M. C. Maiden, "Open-access bacterial population genomics: Bigsdb software, the pubmlst. org website and their applications," *Wellcome open research*, vol. 3, 2018.

[29] R. R. Wick, L. M. Judd, C. L. Gorrie, and K. E. Holt, "Unicycler: resolving bacterial genome assemblies from short and long sequencing reads," *PLoS computational biology*, vol. 13, no. 6, p. e1005595, 2017.

[30] A. Gurevich, V. Saveliev, N. Vyahhi, and G. Tesler, "Quast: quality assessment tool for genome assemblies," *Bioinformatics*, vol. 29, no. 8, pp. 1072–1075, 2013.

[31] T. Seemann, "Abricate: Mass screening of contigs for antimicrobial and virulence genes.." https://github.com/tseemann/abricate, 2016.

[32] T. Seemann, "Snippy: Rapid haploid variant calling and core genome alignment.." https://github.com/tseemann/snippy, 2014.

[33] A. Stamatakis, "Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312–1313, 2014.

[34] Z. Yang, "Maximum likelihood phylogenetic estimation from dna sequences with variable rates over sites: approximate methods," *Journal of Molecular evolution*, vol. 39, pp. 306–314, 1994.

[35] T. B. Athey, S. Teatero, A. Li, A. Marchand-Austin, B. W. Beall, and N. Fittipaldi, "Deriving group a streptococcus typing information from short-read whole-genome sequencing data," *Journal of clinical microbiology*, vol. 52, no. 6, pp. 1871–1876, 2014.

[36] G. Yu, D. K. Smith, H. Zhu, Y. Guan, and T. T.-Y. Lam, "ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data," *Methods in Ecology and Evolution*, vol. 8, no. 1, pp. 28–36, 2017.

[37] T. Seemann, F. Klötzl, and J. Jonathan Terhorst, "snp-dists." https://github.com/tseemann/snp-dists.

[38] S. Teatero, E. Ramoutar, A. McGeer, A. Li, R. G. Melano, J. Wasserscheid, K. Dewar, and N. Fittipaldi, "Clonal complex 17 group b streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool," *Scientific reports*, vol. 6, no. 1, p. 20047, 2016.

[39] M. Meehan, M. Eogan, N. McCallion, R. Cunney, J. E. Bray, K. A. Jolley, A. Unitt, M. C. Maiden, O. B. Harrison, and R. J. Drew, "Genomic epidemiology of group b streptococci spanning 10 years in an irish maternity hospital, 2008–2017," *Journal of Infection*, vol. 83, no. 1, pp. 37–45, 2021.

[40] L. S. Katz, T. Griswold, S. S. Morrison, J. A. Caravas, S. Zhang, H. C. den Bakker, X. Deng, and H. A. Carleton, "Mashtree: a rapid comparison of whole genome sequence files," *Journal of Open Source Software*, vol. 4, no. 44, 2019.

[41] J. A. Lees, M. Galardini, S. D. Bentley, J. N. Weiser, and J. Corander, "Pyseer: a comprehensive tool for microbial pangenome-wide association studies," *Bioinformatics*, vol. 34, no. 24, pp. 4310–4312, 2018.

[42] R. C. Team *et al.*, "R: A language and environment for statistical computing," *Foundation for Statistical Computing, Vienna, Austria*, 2013.

[43] C. L. Trotter, M. Alderson, Z. Dangor, M. Ip, K. Le Doare, E. Nakabembe, S. R. Procter, M. Sekikubo, and P. Lambach, "Vaccine value profile for group b streptococcus," *Vaccine*, vol. 41, pp. S41–S52, 2023.

[44] S. Tan, Y. Lin, K. Foo, H. F. Koh, C. Tow, Y. Zhang, L. W. Ang, L. Cui, H. Badaruddin, P. L. Ooi, *et al.*, "Group b streptococcus serotype iii sequence type 283 bacteremia associated with consumption of raw fish, singapore," *Emerging infectious diseases*, vol. 22, no. 11, p. 1970, 2016.

[45] R. Elling, M. Hufnagel, A. De Zoysa, F. Lander, K. Zumstein, M. Krueger, and P. Henneke, "Synchronous recurrence of group b streptococcal late-onset sepsis in twins," *Pediatrics*, vol. 133, no. 5, pp. e1388–e1391, 2014.

[46] A. Furuta, A. Brokaw, G. Manuel, M. Dacanay, L. Marcell, R. Seepersaud, L. Rajagopal, and K. Adams Waldorf, "Bacterial and host determinants of group b streptococcal infection of the neonate and infant," *Frontiers in Microbiology*, vol. 13, p. 820365, 2022.

[47] I. Margarit, C. D. Rinaudo, C. L. Galeotti, D. Maione, C. Ghezzo, E. Buttazzoni, R. Rosini, Y. Runci, M. Mora, S. Buccato, *et al.*, "Preventing bacterial infections with pilus-based vaccines: the group b streptococcus paradigm," *The Journal of infectious diseases*, vol. 199, no. 1, pp. 108–115, 2009.

[48] B. Metcalf, "GBS_Scripts_Reference." https://github.com/BenJamesMetcalf/GBS_Scripts_Reference.

[49] G. Regev-Yochay, W. P. Hanage, K. Trzcinski, S. L. Rifas-Shiman, G. Lee, A. Bessolo, S. S. Huang, S. I. Pelton, A. J. McAdam, J. A. Finkelstein, *et al.*, "Re-emergence of the type 1 pilus among streptococcus pneumoniae isolates in massachusetts, usa," *Vaccine*, vol. 28, no. 30, pp. 4842–4846, 2010.

[50] M. Gonzalez-Miro, A. Pawlowski, J. Lehtonen, D. Cao, S. Larsson, M. Darsley, G. Kitson, P. B. Fischer, and B. Johansson-Lindbom, "Safety and immunogenicity of the group b streptococcus vaccine alpn in a placebo-controlled double-blind phase 1 trial," *Iscience*, vol. 26, no. 3, 2023.

[51] S. A. Madhi, A. S. Anderson, J. Absalon, D. Radley, R. Simon, B. Jongihlati, R. Strehlau, A. M. Van Niekerk, A. Izu, N. Naidoo, *et al.*, "Potential for maternally administered vaccine for infant group b streptococcus," *New England Journal of Medicine*, vol. 389, no. 3, pp. 215–227, 2023.

[52] L. K. F. Watkins, L. McGee, S. J. Schrag, B. Beall, J. H. Jain, T. Pondo, M. M. Farley, L. H. Harrison, S. M. Zansky, J. Baumbach, *et al.*, "Epidemiology of invasive group b streptococcal infections among nonpregnant adults in the united states, 2008-2016," *JAMA internal medicine*, vol. 179, no. 4, pp. 479–488, 2019.

[53] F. Bianchi-Jassir, P. Paul, K.-N. To, C. Carreras-Abad, A. C. Seale, E. Jauneikaite, S. A. Madhi, N. J. Russell, J. Hall, L. Madrid, *et al.*, "Systematic review of group b streptococcal capsular types, sequence types and surface proteins as potential vaccine candidates," *Vaccine*, vol. 38, no. 43, pp. 6682–6694, 2020.

[54] S. Bellais, A. Six, A. Fouet, M. Longo, N. Dmytruk, P. Glaser, P. Trieu-Cuot, and C. Poyart, "Capsular switching in group b streptococcus cc17 hypervirulent clone: a future challenge for polysaccharide vaccine development," *The Journal of infectious diseases*, vol. 206, no. 11, pp. 1745–1752, 2012.

[55] N. J. Croucher, J. A. Finkelstein, S. I. Pelton, P. K. Mitchell, G. M. Lee, J. Parkhill, S. D. Bentley, W. P. Hanage, and M. Lipsitch, "Population genomics of post-vaccine changes in pneumococcal epidemiology," *Nature genetics*, vol. 45, no. 6, pp. 656–663, 2013.

[56] A. Dobrut and M. Brzychczy-Włoch, "Immunogenic proteins of group b streptococcus—potential antigens in immunodiagnostic assay for gbs detection," *Pathogens*, vol. 11, no. 1, p. 43, 2021.

[57] J. R. Verani, L. McGee, S. J. Schrag, *et al.*, "Prevention of perinatal group b streptococcal disease: revised guidelines from cdc, 2010," 2010.

[58] B. L. Pineles, K. E. Goodman, L. Pineles, and A. D. Harris, "Appropriate antibiotic use for group b streptococcus prophylaxis among penicillin-allergic patients in academic and nonacademic hospitals," in *Open Forum Infectious Diseases*, vol. 9, p. ofac514, Oxford University Press US, 2022.