European Journal of Clinical Microbiology & Infectious Diseases (2023) 42:747–757
https://doi.org/10.1007/s10096-023-04606-9

ORIGINAL ARTICLE

Genomic characterization of emerging invasive *Streptococcus agalactiae* serotype VIII in Alberta, Canada

Ashley N. Williams[1,2] · Matthew A. Croxen[1,2] · Walter H. B. Demczuk[3] · Irene Martin[3] · Gregory J. Tyrrell[1,2,4]

Received: 9 December 2022 / Accepted: 13 April 2023 / Published online: 21 April 2023

Abstract

Invasive Group B Streptococcus (GBS) can infect pregnant women, neonates, and older adults. Invasive GBS serotype VIII is infrequent in Alberta; however, cases have increased in recent years. Here, genomic analysis was used to characterize fourteen adult invasive serotype VIII isolates from 2009 to 2021. Trends in descriptive clinical data and antimicrobial susceptibility results were evaluated for invasive serotype VIII isolates from Alberta. Isolate genomes were sequenced and subjected to molecular sequence typing, virulence and antimicrobial resistance gene identification, phylogenetic analysis, and pangenome determination. Multilocus sequencing typing identified eight ST42 (Clonal Complex; CC19), four ST1 (CC1), and two ST2 (CC1) profiles. Isolates were susceptible to penicillin, erythromycin, chloramphenicol, and clindamycin, apart from one isolate that displayed erythromycin and inducible clindamycin resistance. All isolates carried genes for peptide antibiotic resistance, three isolates for tetracycline resistance, and one for macrolide, lincosamide, and streptogramin resistance. All genomes carried targets currently being considered for protein-based vaccines (e.g., pili and/or Alpha family proteins). Overall, invasive GBS serotype VIII is emerging in Alberta, primarily due to ST42. Characterization and continued surveillance of serotype VIII will be important for outbreak prevention, informing vaccine development, and contributing to our understanding of the global epidemiology of this rare serotype.

Keywords *Streptococcus agalactiae* · Serotype VIII · Alberta · Invasive disease · Genomics · Emerging infectious disease

Introduction

Streptococcus agalactiae, or Group B Streptococcus (GBS), is a Gram-positive coccus and prominent cause of invasive disease worldwide. Invasive GBS infections occur in pregnant women and neonates as well as older adults and at-risk populations, especially those with underlying conditions [1]. Among adults, GBS infection can present as bacteremia, endocarditis, osteomyelitis, septic arthritis, pneumonia, skin and soft tissue infection, and urinary tract infections [1, 2].

✉ Gregory J. Tyrrell
gjt@ualberta.ca

1 University of Alberta, Edmonton, AB, Canada

2 Alberta Precision Laboratories-Public Health Laboratory (ProvLab), Edmonton, AB, Canada

3 National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada

4 Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada

There are currently 10 recognized serotypes (capsule types) for GBS (Ia, Ib, II, III, IV, V, VI, VII, VIII, and IX) [1]. Serotypes responsible for causing most disease vary between patient demographics and geographic location [3, 4]. Globally, serotypes Ia, III, and V account for most invasive maternal disease cases; serotypes Ia and III for early and late onset infant invasive disease; and among non-pregnant adults serotypes V, Ia, and III are predominant, followed by II and 1b [2, 4]. Currently, there are over 1900 sequence types (STs) of GBS within several clonal complexes (https://pubmlst.org/; accessed April 10, 2022), of which CC1, CC10, CC17, CC19, and CC23 are most associated with colonization and invasion of humans [4].

Serotype VIII GBS has been a rare serotype worldwide but relatively more dominant within Japan [5] and has emerged in several additional countries as a predominant colonizer. In South Korea, serotype VIII was first reported in 2000, representing 2% of isolates collected from both adults and infants in a tertiary care hospital from 1993 to 1996, but has since become a predominant serotype among GBS colonized pregnant

women, representing 20% of isolates collected from 2017 to 2019 [6, 7]. Outside of Japan and South Korea, serotype VIII has been isolated in many countries; however, the levels of maternal colonization and infant disease due to serotype VIII are low, with most studies reporting < 3% of colonizing/invasive isolates as serotype VIII [4, 8–10]. Adult invasive serotype VIII has been detected in several locations such as Canada, Denmark, Japan, Portugal, Taiwan, and the USA [4, 10–15]. A global systematic review and meta-analysis of invasive GBS among non-pregnant adults reported between 1975 and 2018 found serotype VIII accounted for 0.16% of serotyped isolates [2].

From 2014 to 2020 in Alberta, Canada, the most prevalent serotypes of invasive GBS in order of prevalence were III, Ia, Ib, II, V, IV, and VI [14, 16]. During this period, seven cases of invasive serotype VIII were identified (0.45% of the total cases), which was more than double seen in a previous study that identified 3 cases (0.2% of total cases) from 2003 to 2013 [14, 16]. In this work, we characterized these emerging invasive serotype VIII isolates in Alberta collected over the period of January 2003 to September 2021 by determining the multilocus sequence types (ST), antimicrobial resistance and virulence factor carriage, phylogenetic relationships, and pangenome of these isolates. Characterization of the serotype VIII in Alberta will contribute to both our understanding of this emerging rare serotype within the province and to global surveillance efforts.

Materials and methods

Identification of capsular polysaccharide type VIII GBS isolates

Neonatal invasive GBS disease is listed as a notifiable disease in Alberta requiring GBS isolates to be submitted to the Alberta Public Health Laboratory for capsular polysaccharide (CPS) typing and antimicrobial susceptibility testing [14]. For other age groups, clinical microbiology laboratories in Alberta submit invasive GBS isolates to the Public Health Laboratory for CPS typing as a part of an ongoing passive surveillance program. Prior to January 2017, CPS typing was performed using a double immunodiffusion method with CPS-specific antisera developed in rabbits [17]. For 2017 to present, a real-time PCR assay utilizing hydrolysis probes was used for CPS typing [18]. All identified serotype VIII GBS isolates were confirmed as serotype VIII by whole genome sequencing (WGS).

Antimicrobial susceptibility testing

Antimicrobial susceptibility testing for penicillin, erythromycin, clindamycin, vancomycin, and chloramphenicol was conducted by disk diffusion (Oxoid Limited, Nepean, ON, Canada). Interpretations of zone diameter break points were made according to the Clinical and Laboratory Standards Institute (CLSI) M100 Performance Standards for Antimicrobial Susceptibility current at the time [19].

Genomic sequencing

Genomic DNA for sequencing was extracted with a modified protocol for the MagaZorb DNA Mini-Prep Kit (Promega). Briefly, colonies were incubated overnight at 35 °C in Bacto Todd-Hewitt Broth (500 µL; BD Biosciences). Cultures were centrifuged at 6010 RCF for 2 min, supernatant removed, and cells resuspended in 12 mM Tris (500 µL). Centrifugation was repeated and cells resuspended in mutanolysin/hyaluronidase lysis solution (62 µL; 10 µL 3000 U/mL mutanolysin (Sigma), 2 µL 30 mg/mL hyaluronidase (Sigma), and 50 µL 10 mM Tris). Lysozyme (15 µL, 100 mg/mL; Sigma) was added and incubated for 1 h at 37 °C with shaking at 700 rpm (Eppendorf ThermoMixer F1.5). Proteinase K solution (20 µL) and RNase A (20 µL, 20 mg/mL; Qiagen or Invitrogen) were added and tubes incubated at room temperature for 5 min. ATL lysis buffer (200 µL) was added and tubes incubated for 2 h at 56 °C with shaking at 900 rpm (Eppendorf ThermoMixer F1.5). Extracts were centrifuged at 9391 RCF for 2 min, and wash, binding, and elution steps were completed with the KingFisher mL Purification System (Thermo Scientific) with Qiagen Buffer EB. Genomes were sequenced either by NextSeq 500/550 or Illumina MiSeq with paired end reads. Genome coverages ranged from 81 and 99% relative to isolate 3 (accession SAMN31119272). Genome sequences were assembled with Shovill v1.1.0 (https://github.com/tseemann/shovill), annotated with Bakta v1.2.2 (database v3.0) [20], and assembly quality assessed with Quast v5.0.2 (Online Resource 1) [21, 22].

Bioinformatic analysis

In silico MLST was performed on assembled genomes with MLST v2.19.0 (https://github.com/tseemann/mlst) using schema from PubMLST [23] and clonal complexes determined using https://pubmlst.org/. The graph of MLST data was generated with Tableau Desktop v2021.2 and Inkscape v0.92.3. The pangenome for Bakta annotated sequences was calculated using Anvi'o v7.1 using the workflow for microbial pangenomics with default parameters and visualized with anvi-display-pan and Inkscape v0.92.3 [24–28]. Antibiotic resistance genes were identified with the Resistance Gene Identifier (RGI) v5.2.0 and the Comprehensive Antibiotic Resistance Database (CARD) v3.1.4 [29]. Virulence factors were identified with abricate v1.0.1 (https://

github.com/tseemann/abricate) and the core Virulence Factor Database (VFDB) (30-Dec-2021) [30]. Additional virulence factors were identified using blastn 2.12.0 + of full pilus gene sequences (PI-2b, GCF_000012705.1; PI-1b major subunit allele, WP_000777405.1) [31] and partial *alp/rib* genes sequences with a 95% nucleotide identity cutoff to determine gene presence [10].

Phylogenetic relationships were assessed using core single nucleotide polymorphisms (SNPs). Core SNPs were extracted from sequence reads and aligned with Snippy v4.6.0 using the assembled genome of isolate 3 as a reference (https://github.com/tseemann/snippy). Regions of recombination within the alignment were determined with Gubbins v3.1.4 using marginal ancestral reconstruction with IQ-TREE v2.0.3 [32, 33] and masked with maskrc-svg v0.5 (https://github.com/kwongj/maskrc-svg). SNP distance matrices were determined for both the unmasked and masked alignments using snp-dists v0.8.2 (https://github.com/tseemann/snp-dists). A maximum likelihood phylogenetic tree of the masked alignment was constructed with IQ-TREE v2.2.0 using model K3Pu + F + I, 1000 ultrafast bootstraps, and 1000 bootstraps replicates for the SH-like approximate likelihood ratio test (SH-alrt) [33–35]. Trees were visualized with FigTree v1.4.4 (https://github.com/rambaut/figtree) and Inkscape v0.92.3.

Results

Epidemiology and antibiotic resistance profiles of GBS serotype VIII in Alberta

Between January 2003 and September 2021, 14 serotype VIII GBS isolates were identified. All isolates were isolated between 2009 and 2021, with one isolate per year for 2009, 2011, 2013, and 2017; two isolates per year from 2018 to 2020; and four isolates in 2021 (Fig. 1). There were equal numbers of males and females and ages ranged from 24 to 79, with two cases (14%) in individuals under 50, three (21%) in those aged 50–60, and nine (64%) in patients over 60 (Table 1). Two isolates (isolate 6 and isolate 9) were recovered from the same patient 173 days apart. Specimens were isolated from blood ($n = 9$), knee fluid ($n = 2$), right leg ($n = 1$), thumb tissue ($n = 1$), and a toe bone ($n = 1$) (Table 1). All serotype VIII isolates were susceptible to penicillin, erythromycin, clindamycin, chloramphenicol, and vancomycin except for isolate 2, which exhibited erythromycin and inducible clindamycin resistance (Table 1).

ST42 is the most common serotype VIII sequence type in Alberta

The 14 genomes comprised three MLST types: ST1, ST2, and ST42 (Fig. 1; Table 1), and one isolate (isolate 14) was not

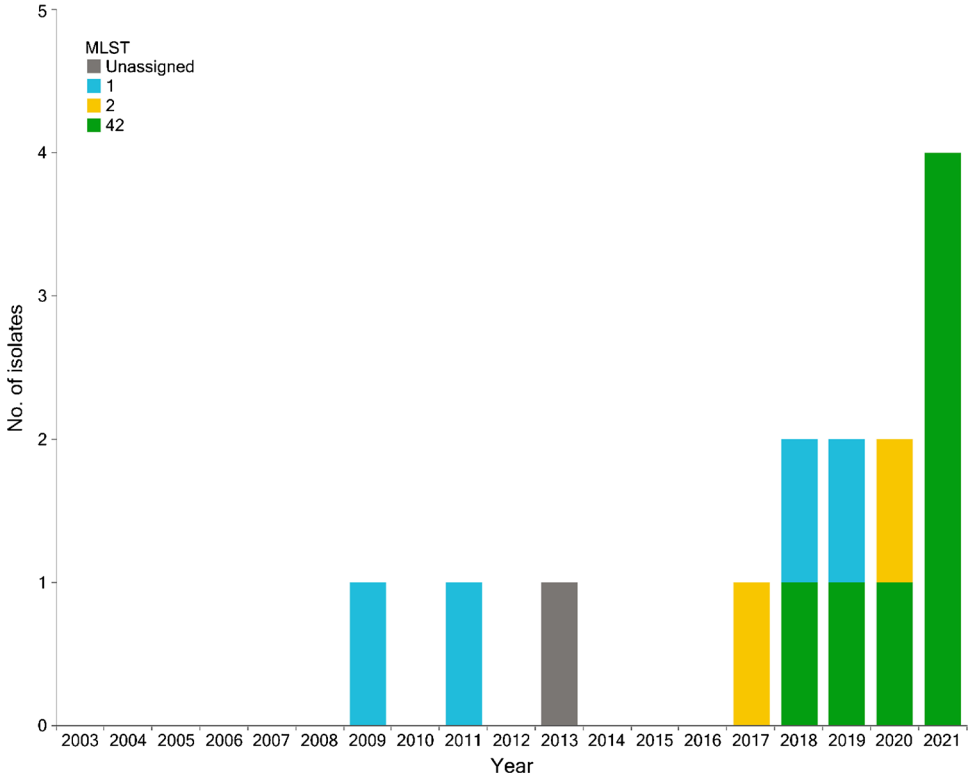

Fig. 1 *Streptococcus agalactiae* serotype VIII cases in Alberta from January 2003 to September 2021. Bars are colored by multilocus sequence type (MLST)

Table 1 Demographic and clinical descriptors of the 14 *Streptococcus agalactiae* serotype VIII cases from January 2009 to September 2021

Isolate	Year	Age	Sex	Source	MLST	Antibiotic (zone size in mm)					
						Pen	Eryth	Clin	DTEST	Chl	Vanc
1	2017	66	F	Blood	2	28	25	22	NEG	25	19
2	2018	76	F	Fluid (knee)	1	28	0	20	POS	23	18
3	2018	72	M	Fluid (knee)	42	28	25	22	NEG	26	19
4	2019	24	M	Blood	42	29	24	21	NEG	24	18
5	2019	71	F	Blood	1	30	23	20	NEG	24	19
6[a]	2020	53	F	Blood	42	32	26	23	NEG	25	30
7	2020	69	M	Blood	2	28	25	23	NEG	24	20
8	2021	64	M	Right leg	42	27	26	22	NEG	25	20
9[a]	2021	53	F	Blood	42	30	26	22	NEG	26	19
10	2021	57	M	tissue thumb)	42	33	29	24	NEG	27	20
11	2021	68	F	Blood	42	31	26	22	NEG	24	19
12	2009	79	F	Blood	1	28	24	21	NEG	22	18
13	2011	49	M	Blood	1	27	24	21	NEG	24	18
14	2013	68	M	Bone (toe)	-	26	23	20	NEG	24	17

Pen penicillin, *Eryth* erythromycin, *Clin* clindamycin, *DTEST* inducible clindamycin resistance; *Chl* chloramphenicol, *Vanc* vancomycin

[a]Isolates 6 and 9 are from the same patient separated in time by 173 days

assigned an ST. Of the three identified sequence types, ST42 was predominant, comprising 7/14 total isolates and all cases identified in 2021. The sequence types were further clustered by clonal complex and fell within CC1 (ST1, ST2) and CC19 (ST42).

Alberta serotype VIII isolates carry few antimicrobial resistance genes

Genomes of serotype VIII isolates were searched for antibiotic resistance ontologies (AROs) (Table 2) [29]. All isolates carried *mprF*, which provides peptide antibiotic resistance. Three isolates (isolates 3, 4, and 14) had strict hits (imperfect hits that meet curated blastp bit score cutoffs) to *tetO*, which is implicated in tetracycline resistance [29]. Isolate 2 was the only isolate with a strict hit to *ermA* (Table 2).

Alberta serotype VIII isolates carry multiple virulence factors including potential protein vaccine targets

Genomes were also surveyed for virulence factors [30]. Virulence factors present in all isolates included capsular operons, exotoxin production (hemolysin/cytolysin, CAMP factor), and exoenzyme production (hyaluronidase) (Table 3). More variation was present among adherence-related genes identified, with 13/14 genomes possessing C5a peptidase, laminin-binding surface protein

Table 2 Antimicrobial resistance genes carried by invasive *Streptococcus agalactiae* serotype VIII isolates

Resistance mechanism	Drug class	AMR gene family	Best hit	No. isolates	Min. % identity[a]	Min. % length of reference sequence[a]
Antibiotic target alteration	Macrolide antibiotic; lincosamide; streptogramin	Erm 23S ribosomal RNA methyltransferase	23S rRNA (adenine(2058)-N(6))-methyltransferase Erm(A)	1	99.59	100
	Peptide antibiotic	Defensin resistant *mprF*	*Streptococcus agalactiae mprF*	14	99.53	99.88
Antibiotic target protection	Tetracycline antibiotic	Tetracycline-resistant ribosomal protection protein	*tetO*	3	98.28	100

Shown are Resistance Gene Identifier hits, which identifies antimicrobial resistance ontologies; only strict hits are shown (no perfect hits identified)

[a]The lowest percent identity or coverage of a serotype VIII gene to the RGI hit among all isolates with that gene

Table 3 Virulence factor genes carried by invasive *Streptococcus agalactiae* serotype VIII in Alberta

Class	Virulence factor	Gene	Description	NCBI accession	No. isolates	Min. % nt identity[a]	Min. % nt coverage[b]
Adherence	C5a peptidase	*scpA*/*scpB*	Streptococcal C5a peptidase	WP_001227855	13	99.94	100
	Lmb	*lmb*	Laminin-binding surface protein	WP_000715197	13	100	100
	PI-1	GBS_RS03565	Major subunit protein	WP_000777402	6	99.88	100
		BP-1b[c]	Major subunit protein variant 1b	WP_000777405.1	8	100	100
		GBS_RS03570	LPXTG cell wall anchor domain-containing protein	WP_000815035	13	98.05	100
		GBS_RS03585	Tip adhesin protein	WP_001868236	13	99.92	100
		srtC1	Class C sortase	WP_000529916	13	99.78	100
		srtC2	Class C sortase	WP_000746885	13	99.65	100
	PI-2b[d]	Entire operon (*orf*, *lep*, *ap1*, *spb1*, *srtC1*, *ap2*, *srt2*)	PI-2b pilus	GCF_000012705.1	14	99.86	42 (100)[e]
Exoenzyme	Hyaluronidase	*hylB*	Hyaluronidase	WP_000403400	14	97.37	100
Exotoxin	CAMP factor	*cfa*/*cfb*	cAMP factor	WP_001101136	14	99.61	100
	Haemolysin/ cytolysin	*acpC*	Acyl carrier protein AcpC	WP_000611493	14	100	100
		cylA	ABC (ATP-binding cassette) transporter CylA	WP_000403526	14	99.78	100
		cylB	ABC (ATP-binding cassette) transporter CylB	WP_000462410	14	99.77	100
		cylD	Malonyl-CoA-ACP transacylase CylD	WP_000859501	14	100	100
		cylE	Haemolysin CylE	WP_000650746	14	99.95	100
		cylF	Putative aminomethyl-transferase CylF	WP_001092618	14	100	100
		cylG	3-ketoacyl-ACP-reductase CylG	WP_000861302	14	99.86	100
		cylI	Putative 3-ketoacyl-ACP synthase CylI	WP_000118217	14	99.91	100
		cylJ	Glycosyltransferase CylJ	WP_000033003	14	98.76	99.92
		cylK	Putative phosphopantetheinyl transferase CylK	WP_001068957	14	98.96	100
		cylX	Acetyl coenzyme A (CoA) carboxylase CylX	WP_000533775	14	99.67	100
		cylZ	3R-hydroxymyristoyl ACP dehydratase CylZ	WP_000164166	14	100	100
Immune modulation	Capsule	GBS_RS06540	Acetyltransferase	WP_000727597	14	99.36	100

Table 3 (continued)

Class	Virulence factor	Gene	Description	NCBI accession	No. isolates	Min. % nt identity[a]	Min. % nt coverage[b]
		GBS_RS06555	Oligosaccharide flippase family protein	WP_001093064	14	99.14	100
		GBS_RS06595	Tyrosine-protein kinase	WP_000197412	14	99.57	99
		GBS_RS06600	Capsular polysaccharide biosynthesis protein CpsC	WP_001033074	14	99.27	98.41
		GBS_RS06605	Tyrosine-protein phosphatase CpsB	WP_000565385	14	99.32	100
		GBS_RS06610	LCP family protein	WP_000064997	14	99.52	100
		neuA	*N*-Acylneuraminate cytidylyltransferase	WP_000802346	14	98.79	95.33
		neuB	*N*-Acetylneuraminate synthase	WP_000262522	14	99.51	100
		neuC	UDP-*N*-acetylglucosamine 2-epimerase	WP_000717643	14	99.39	100

[a]The lowest percent nucleotide (nt) identity of the serotype VIII hit to a VFDB gene among all isolates with that gene

[b]The lowest percent nucleotide (nt) sequence coverage of a serotype VIII gene to the VFDB gene among all isolates with that gene

[c]A variant of the major subunit protein (also called backbone pilin protein) found in some isolates of GBS [36]

[d]Identified with standalone blastn of PI-2b operon described in [31]

[e]All isolates had 100% query coverage, except for isolate 10 where the operon was split between two contigs with 42% and 58% query coverage

(Lmb), and most PI-1 pilus-associated genes. Pilus gene GBS-RS03565 (major subunit protein) was only present in 6/14 isolates (43%). In isolates that did not carry GBS-RS03565, an alternative major subunit protein (BP-1b) was identified by blastn [36]. In addition, pilus island PI-2b, which is absent in the VFDB dataset, was identified by blastn in all isolates, with a split between two contigs in isolate 10.

Isolates were also surveyed by blastn for the presence of members of the Alpha/Rib protein family, which are current targets for vaccine development (Table 4) [37]. Alp1 was present in 43% of isolates, followed by Alp2/3 (29%), Rib (14%), and Alpha (7%). All genomes except isolate 10 carried at least one of the Alpha/Rib gene sequences (Online Resource 2). The type of Alpha/Rib family protein carried corresponded to ST: ST42 = *alp1*, ST1 = *alp2/3*, and ST2 = *rib*.

Table 4 Alpha/Rib family surface proteins carried by invasive *Streptococcus agalactiae* serotype VIII in Alberta

Surface protein	Gene	Gene source	NCBI accession	No. isolates	Min. % nt identity[a]	Min. % nt coverage[b]
Alp1	*alp1*	*Streptococcus agalactiae* 27	AY461799	6	99.61	100
Alp2/3	*alp2/3*	*Streptococcus agalactiae* JM9-130013	AF245663	4	100	100
Alpha	*alpha*	*Streptococcus agalactiae* A909	M97256	1	100	100
Rib	*rib*	*Streptococcus agalactiae* BM110	U58333	2	100	100

Screening for these genes was performed using standalone blastn of partial sequences from [10] with a 95% identity cutoff

[a]The lowest percent nucleotide (nt) identity of a VFDB gene to the serotype VIII hit among all isolates with that gene

[b]The lowest percent nucleotide (nt) sequence coverage of a serotype VIII gene to the VFDB gene among all isolates with that gene

Genome similarity is greatest within Alberta serotype VIII STs

To examine the phylogenetic relationships among serotype VIII isolates, a maximum likelihood tree was generated from a core SNP alignment masked for recombination (Fig. 2). The isolates grouped primarily by ST, with ST42 (CC19) forming a monophyletic group and ST2 nesting within the ST1 group (CC1). Isolate 14, which did not have an assigned ST, formed a sister group to the ST1/ST2 cluster. Isolates 6 and 9, which were sampled from the same patient 173 days apart, also clustered together. Summary statistics of SNP alignments in relation to the reference sequence can be found in Online Resource 3.

To determine relative numbers of SNPs between isolates, pairwise SNP distances were calculated from the unmasked and masked (recombinant regions removed) core SNP alignments (reference = isolate 3; Online Resource 4). There was a decrease in the number of SNPs between the unmasked and masked alignment ranging from 0 to 95% reduction, with a mean of 73% and median of 89%. For the masked alignment, the number of SNPs between different STs ranged from 80 to 490 SNPs overall, with 80–112 SNPs between ST1 and ST2, 381–490 between ST1 and ST42, and 348–436 between ST2 and ST42. Within STs, the number of masked SNPs ranged from 17 to 119 SNPs overall, with 79–119 SNPs between ST1 isolates, 40 SNPs between ST2 isolates, and 17–94 SNPs among ST42 isolates. Isolate 14, which did not fall into an MLST group, shared the fewest numbers of masked SNPs with isolate 1 ($n = 95$ SNPs; ST2), isolate 7 ($n = 96$; ST2), isolate 12 ($n = 94$; ST1), and isolate 13 ($n = 99$; ST1). The fewest number of SNPs were seen between isolates 6 and 9 ($n = 17$ masked and unmasked), which were isolated from the same patient and are both part of ST42.

Alberta serotype VIII isolates share a core genome representing 66% of the pangenome

The core genome of the fourteen serotype VIII isolates was composed of 1662 genes, representing 66% of pangenome (Fig. 3). The accessory genome (present in 2 or more, but less than 14 isolates) was composed of 588 genes (23% of pangenome), and singletons (present in only one isolate) made up 268 genes (11% of pangenome). Singleton gene counts ranged from 0 to 64 genes, with isolate 14, which did not have an assigned ST, carrying the highest number of singleton genes ($n = 64$). Among isolates with MLSTs, most singletons were carried by ST1 and ST2 isolates, with isolate 7 (ST2) carrying the most ($n = 54$ singletons). Among ST42 isolates, isolate 8 carried the largest number of singletons ($n = 23$).

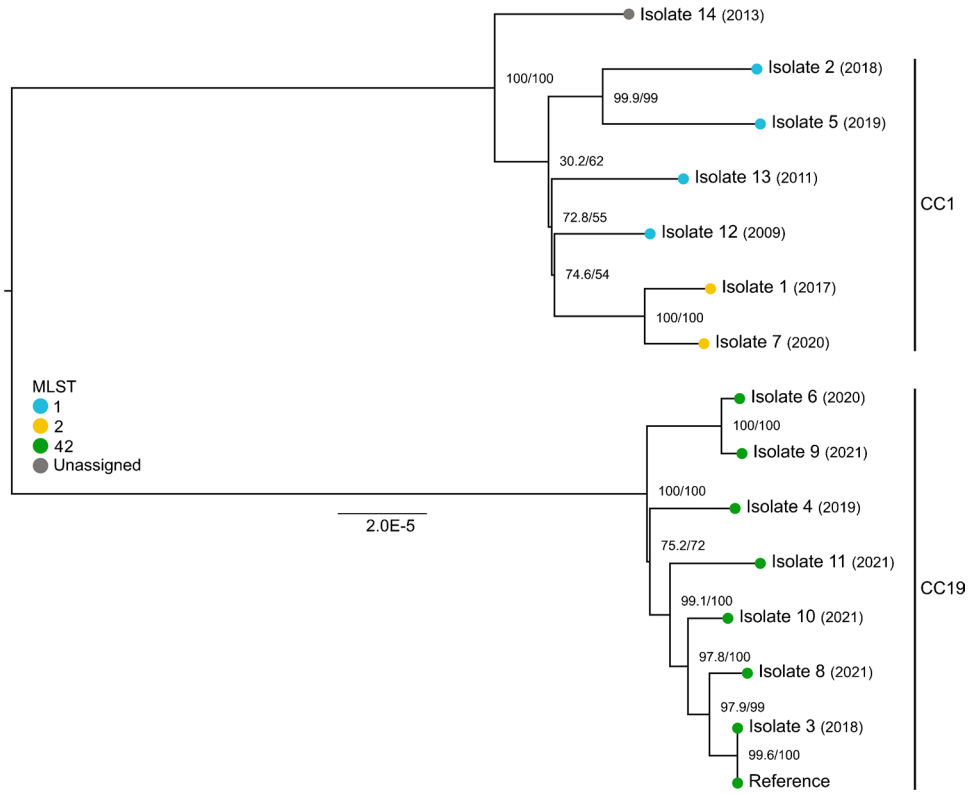

Fig. 2 Core SNP maximum likelihood phylogeny masked for recombination of *Streptococcus agalactiae* serotype VIII isolates. Dates shown are collection years, and isolates are colored by multilocus sequence type (MLST). Tree is midpoint rooted, and branch supports are SH-aLRT support (%)/ ultrafast bootstrap support (%). The reference strain is isolate 3. CC, clonal complex; MLST, multilocus sequence type

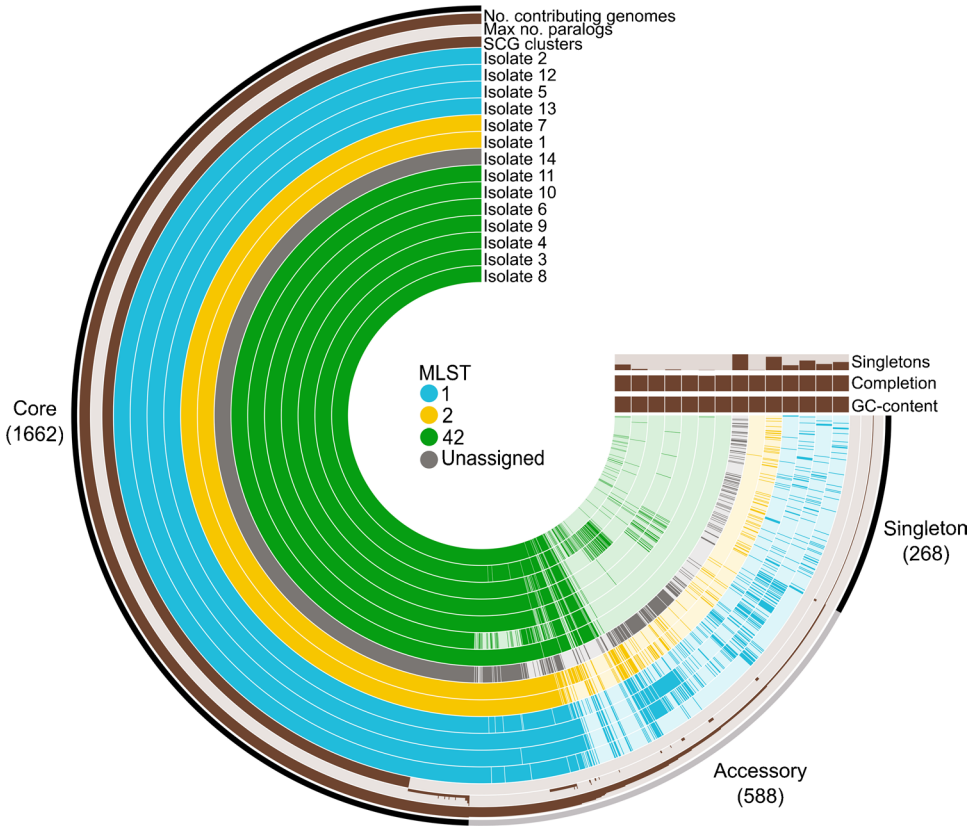

Fig. 3 Pangenome calculation of invasive *Streptococcus agalactiae* serotype VIII in Alberta. One thousand one hundred sixty-two core genes were present in all genomes, 588 accessory genes were present in at least two but less than 14 genomes, and 268 genes were only present in one genome (singletons). Each genome is colored by multilocus sequence type (MLSTs). SCG, single gene cluster (excludes paralogs); MLST, multilocus sequence type

Discussion

In this study, we characterized serotype VIII GBS isolates identified in Alberta from January 2003 to September 2021 [14, 16]. The goal was to characterize this emerging group within Alberta primarily through genomic analyses, as there are no previous genome-based studies focused on this group to the best of our knowledge. We evaluated basic demographics, STs, antimicrobial resistance, virulence factors, and phylogenetic and pangenomic relationships between isolates.

Since 2017, the frequency of serotype VIII GBS infections has increased in Alberta, primarily represented by ST42 (Fig. 1). As of April 2022, 43 isolates with a capsular serotype and/or genotype of VIII were present in PubMLST, the majority of which were part of ST1 (CC1) (34/43 isolates) [23]. ST2 serotype VIII were limited within the database with only four isolates identified in Australia, of which three appeared to be invasive being isolated from blood and joint fluid (PubMLST IDs 5714, 5748, 10570). In general, limited information is available for serotype VIII ST42 (CC19) as most serotype VIII isolates identified have been a part of CC1 [38]. In PubMLST, only seven ST42 isolates were present, six from the USA, and one from the UK, of which only the latter isolate was designated as serotype VIII (PubMLST ID 21210) [39]. To the best of our knowledge, no information about this isolate exists apart from its designation as a clinical GBS isolate. One study characterizing invasive GBS isolates in the USA across multiple states from 2015 to 2017 had a similar pattern of ST distribution among their serotype VIII as Alberta [10]. In the US study, 17 serotype VIII isolates from a total of 6336 invasive GBS were identified. The majority of the isolates were ST42 (10/17 isolates), followed by ST2 ($n=4$), ST1 ($n=1$), and two not assigned a sequence type. In addition, all serotype VIII isolates were identified in adults over 18 years, with eight isolates from individuals 65 years and older. These results are similar to what we have described here in Alberta and may represent the emergence of serotype VIII ST42 infection, especially among older adults in North America.

In our study, antimicrobial resistance was limited, with only one isolate exhibiting erythromycin and inducible clindamycin resistance (Table 1). Alberta has seen increases in resistance to erythromycin and clindamycin among invasive GBS, especially among adult isolates [14]. In a study comparing changes in resistance rates among GBS isolates in Alberta between 2003–2013 and 2014–2020, erythromycin non-susceptibility rose from 36.9 to 50.8% and clindamycin non-susceptibility rose from 21.0 to 45.8% between the two time periods [14, 16]. High rates of erythromycin and clindamycin resistance have been observed in several countries across the globe among both colonizing and invasive isolates

[40]. The most widespread factors conferring resistance to macrolides are *erm* (erythromycin ribosome methylation) genes, including *ermA* and *ermB*, among which *ermB* is most common in GBS [40]. In our study, the one isolate exhibiting erythromycin/inducible clindamycin resistance was found to carry a gene with high nucleotide similarity to *ermA* (Table 2), which has been observed to be associated with this phenotype previously. Among 18 macrolide-resistant colonizing GBS isolates in Brazil, 39% were found to harbor *ermA*, which included isolates exhibiting both macrolide resistance and constitutive/inducible macrolide-lincosamide-streptogramin B resistance [41]. The presence of *erm* genes has been shown to correlate well with resistant phenotypes and are therefore plausibly responsible for the erythromycin and inducible clindamycin resistance of our isolate [42].

A vaccine for GBS has been listed as a priority by the World Health Organization [43]. Vaccines in current development include both polysaccharide-conjugate vaccines and protein-only vaccines that target surface antigens of GBS [44]. While polysaccharide-conjugate vaccine formulations that include serotype VIII have undergone preclinical testing, all formulations currently in clinical trials do not include serotype VIII due to low global burden [44, 45]. In contrast, protein-only vaccines have the potential to be used broadly for protection against all serotypes by targeting universal groups of surface proteins [44]. Attractive choices for protein vaccine development include Alpha/Rib family proteins, which are abundant surface proteins among GBS isolates [44]. A systematic review that included 7193 GBS isolates found that 99% of adult invasive disease and 93% of infant invasive disease isolates carried genes for at least one of Alp1/Epsilon, Alp2/3, Alpha C, or Rib surface protein [4]. All but one of the Alberta serotype VIII isolates possessed at least one Alpha/Rib protein (Table 4; Online Resource 2). Other surface proteins of interest for vaccine target include pili structures, which are also widespread in GBS [4, 46]. Among Alberta serotype VIII isolates, at least one pilus operon was present in each isolate (Table 3). Interestingly, 57% ($n=8$) of the Alberta isolates had an alternative major subunit allele for PI-1 and BP-1b, which was first reported in a study of invasive and colonizing GBS isolates in Canada [36]. Among the 1332 isolates included in the Canadian study, 51 carried the alternative backbone protein (major subunit), among which was one serotype VIII (ST42) [36].

We used core SNP alignments masked for recombination to build a phylogenetic tree of our serotype VIII isolates, which primarily clustered by ST/CC (Fig. 2). Among our serotype VIII was a single isolate, isolate 14, which was not assigned an MLST. Core SNP phylogenetic analysis showed that this isolate formed a sister group to the CC1 cluster, demonstrating sequence similarity to this clonal complex (Fig. 2). The inability to determine the ST of isolate 14 suggests this may be a new ST, perhaps generated through gene duplication or deletion events as a result of recombination, or not represented within the current MLST scheme. Recombination within the genomes of our serotype VIII isolates was evident in the comparison of the masked and unmasked pairwise SNP distances, with an average 73% reduction in SNPs after recombination masking (Online Resource 4).

Multiple studies have characterized the pangenome of GBS [47–50]. The pangenome of the 14 Alberta serotype VIII isolates was composed of 2518 genes, of which 1662 genes were core (Fig. 3). This is in good agreement with previous studies that have calculated core genomes among human-associated *S. agalactiae* isolates by diverse methodologies to be 1806 genes, 1658 genes, and 1472 genes [47–49]. Isolate 14, which had no assigned ST, carried the largest number of singleton genes in our pangenomic analysis ($n=64$) (Fig. 3), further highlighting the uniqueness of this isolate within our collection. Interestingly, several of our isolates did not possess any singleton genes, which is in contrast with studies that have characterized the GBS pangenome as "open," expanding with the addition of each new genome [47, 50]. It is possible that the lack of singletons among some of our isolates, primarily ST42, may represent clonal expansion and as such there would be fewer differences in gene content expected between isolates.

Here, we presented the genomic characterization of invasive serotype VIII GBS within Alberta. Limitations to this study include our small sample size, the use of draft genomes, which may impact gene presence/absence determination, and interpretations of isolate 10 results due to lower depth of coverage and poorer assembly for this isolate. Future directions for this work include the continued surveillance and genomic characterization of invasive serotype VIII in Alberta to determine if the upward trend of invasive infection continues. Globally, invasive serotype VIII GBS is uncommon, and consequently, our understanding of this emerging group is limited. Monitoring and characterization of global isolates is necessary to both inform the epidemiology of this organism and guide vaccine development for the prevention of invasive GBS infection among at-risk adults.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10096-023-04606-9.

Acknowledgements We gratefully thank the diagnostic laboratories in Alberta for their submission of invasive *Streptococcus agalactiae* isolates for capsular polysaccharide typing and antimicrobial susceptibility testing. Further, we thank all the individuals at the Public Health Laboratory (ProvLab) in Edmonton, Alberta, and the National Microbiology Lab in Winnipeg, Manitoba, who contributed to the testing and sequencing of isolates, with special thanks to Sandy Shokoples for providing the DNA extraction protocol.

Author contribution Conceptualization: GJT, MAC, and ANW. Data curation: ANW. Formal analysis: ANW. Funding acquisition: GJT.

Investigation: ANW. Methodology: ANW, MAC, WHBD, and GJT. Project administration: GJT and ANW. Resources: GJT, WHBD, and IM. Software: ANW and MAC. Supervision: GJT and MAC. Validation: https://credit.niso.org/. Visualization: ANW. Writing—original draft: ANW. Writing—review and editing: GJT, MAC, ANW, WHBD, and IM.

Funding This research was partially funded by a grant from the Antimicrobial Resistance-One Health Consortium Major Innovation Fund program of the Ministry of Jobs, Economy and Innovation, Government of Alberta (awarded to GJT).

Data availability Sequence read accessions used in this study are available in the NCBI Sequence Read Archive: SAMN31119270 (isolate 1), SAMN31119271 (isolate 2), SAMN31119272 (isolate 3), SAMN31119273 (isolate 4), SAMN31119274 (isolate 5), SAMN31119276 (isolate 6), SAMN31119275 (isolate 7), SAMN31119277 (isolate 8), SAMN31119278 (isolate 9), SAMN31119279 (isolate 10), SAMN31119280 (isolate 11), SAMN31119267 (isolate 12), SAMN31119268 (isolate 13), and SAMN31119269 (isolate 14).

Code availability Not applicable.

Declarations

Ethics approval No ethics approval was required for this study.

Consent to participate Not applicable.

Consent for publication Not applicable

Conflict of interest The authors declare no competing interests.

References

1. Raabe VN, Shane AL (2019) Group B streptococcus (*Streptococcus agalactiae*). Microbiol Spectr 7:GPP3–0007–2018. https://doi.org/10.1128/microbiolspec.GPP3-0007-2018
2. Navarro-Torné A, Curcio D, Moïsi JC, Jodar L (2021) Burden of invasive group B streptococcus disease in non-pregnant adults: a systematic review and meta-analysis. PLoS One 16:e0258030. https://doi.org/10.1371/journal.pone.0258030
3. Lannes-Costa PS, de Oliveira JSS, da Silva SG, Nagao PE (2021) A current review of pathogenicity determinants of *Streptococcus* sp. J Appl Microbiol 24(131):1600–1620. https://doi.org/10.1111/JAM.15090
4. Bianchi-Jassir F, Paul P, To K-N, Carreras-Abad C, Seale AC, Jauneikaite E et al (2020) Systematic review of group B streptococcal capsular types, sequence types and surface proteins as potential vaccine candidates. Vaccine 7(38):6682–6694. https://doi.org/10.1016/j.vaccine.2020.08.052
5. Lachenauer CS, Kasper DL, Shimada J, Ichiman Y, Ohtsuka H, Kaku M et al (1999) Serotypes VI and VIII predominate among group B streptococci isolated from pregnant Japanese women. J Infect Dis 1(179):1030–1033. https://doi.org/10.1086/314666
6. Choi SJ, Kang J, Uh Y (2021) Recent epidemiological changes in group B streptococcus among pregnant Korean women. Ann Lab Med 1(41):380–385. https://doi.org/10.3343/alm.2021.41.4.380
7. Lee K, Shin JW, Chong Y, Mikamo H (2000) Trends in serotypes and antimicrobial susceptibility of group B streptococci isolated in Korea. J Infect Chemother 1(6):93–97. https://doi.org/10.1007/pl00012158
8. Perme T, Golparian D, Bombek Ihan M, Rojnik A, Lučovnik M, Kornhauser Cerar L et al (2020) Genomic and phenotypic characterisation of invasive neonatal and colonising group B streptococcus isolates from Slovenia, 2001–2018. BMC Infect Dis 16:20. https://doi.org/10.1186/S12879-020-05599-Y
9. Furfaro LL, Nathan EA, Chang BJ, Payne MS (2019) Group B streptococcus prevalence, serotype distribution and colonization dynamics in Western Australian pregnant women. J Med Microbiol 1(68):728–740. https://doi.org/10.1099/JMM.0.000980/CITE/REFWORKS
10. McGee L, Chochua S, Li Z, Mathis S, Rivers J, Metcalf B et al (2021) 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 15(72):1004–1013. https://doi.org/10.1093/CID/CIAA151
11. Omura Y, Kusama Y, Takeuchi N, Ishiwada N (2018) Mediastinal, subcutaneous and multiple muscular abscesses caused by group B streptococcus serotype VIII in a type 2 diabetes mellitus patient. J Infect Chemother 1(24):401–403. https://doi.org/10.1016/j.jiac.2017.12.007
12. Wong SS, Tsui K, Liu QD, Lin L-C, Tsai CR, Chen L-C et al (2011) Serotypes, surface proteins, and clinical syndromes of invasive group B streptococcal infections in northern Taiwan, 1998–2009. J Microbiol Immunol Infect 44:8–14. https://doi.org/10.1016/j.jmii.2011.01.003
13. Ekelund K, Slotved HC, Nielsen HU, Kaltoft MS, Konradsen HB (2003) Emergence of invasive serotype VIII group B streptococcal infections in Denmark. J Clin Microbiol 1(41):4442–4444. https://doi.org/10.1128/JCM.41.9.4442-4444.2003
14. Ma A, Thompson LA, Corsiatto T, Hurteau D, Tyrrell GJ (2021) Epidemiological characterization of group B streptococcus infections in Alberta, Canada: an update from 2014 to 2020. Microbiol Spectr 9:e0128321. https://doi.org/10.1128/SPECTRUM.01283-21/ASSET/B6FB25AC-B3DD-4732-B1F6-63CDAA60B458/ASSETS/IMAGES/LARGE/SPECTRUM.01283-21-F002.JPG
15. Martins ER, Pedroso-Roussado C, Melo-Cristino J, Ramirez M, The Portuguese Group for the Study of Streptococcal Infections (2017) *Streptococcus agalactiae* causing neonatal infections in Portugal (2005–2015): diversification and emergence of a CC17/PI-2b multidrug resistant sublineage. Front Microbiol 8:499. https://doi.org/10.3389/fmicb.2017.00499
16. Alhhazmi A, Hurteau D, Tyrrell GJ (2016) Epidemiology of invasive group B streptococcal disease in Alberta, Canada, from 2003 to 2013. J Clin Microbiol 1(54):1774–1781. https://doi.org/10.1128/JCM.00355-16
17. Lancefield RC, McCarty M, Everly WN (1975) Multiple mouse protective antibodies directed against group B streptococci. Special reference to antibodies effective against protein antigens. J Exp Med 142:165–79. https://doi.org/10.1084/jem.142.1.165
18. Alhhazmi A, Pandey A, Tyrrell GJ (2017) Identification of group B streptococcus capsule type by use of a dual phenotypic/genotypic assay. J Clin Microbiol 23(55):2637–2650. https://doi.org/10.1128/JCM.00300-17
19. Clinical and Laboratory Standards Institute (CLSI) (2020) Performance standards for antimicrobial susceptibility testing, 30th ed. In: CLSI supplement M100. Clinical and Laboratory Standards Institute, Wayne, Pennsylvania
20. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A (2021) Bakta: Rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom 5(7):000685. https://doi.org/10.1099/MGEN.0.000685
21. Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics

15(29):1072–1075. https://doi.org/10.1093/BIOINFORMATICS/BTT086
22. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A (2018) Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 1(34):i142–i150. https://doi.org/10.1093/BIOINFORMATICS/BTY266
23. Jolley KA, Bray JE, Maiden MCJ (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1
24. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS et al (2020) Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 21(6):3–6. https://doi.org/10.1038/s41564-020-00834-3
25. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML et al (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319. https://doi.org/10.7717/PEERJ.1319
26. van Dongen S, Abreu-Goodger C (2012) Using MCL to extract clusters from networks. In: Helden J, Toussaint A, Thieffry D, editors. Bacterial Molecular Networks - Methods in Molecular Biology. New York: Springer, p 281–95. https://doi.org/10.1007/978-1-61779-361-5_15
27. Benedict MN, Henriksen JR, Metcalf WW, Whitaker RJ, Price ND (2014) ITEP: an integrated toolkit for exploration of microbial pan-genomes. BMC Genom 3(15):8. https://doi.org/10.1186/1471-2164-15-8
28. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/NMETH.3176
29. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A et al (2020) CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 8(48):D517–D525. https://doi.org/10.1093/NAR/GKZ935
30. Liu B, Zheng D, Jin Q, Chen L, Yang J (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 8(47):D687–D692. https://doi.org/10.1093/nar/gky1080
31. Périchon B, Guignot J, Szili N, Gao C, Poyart C, Trieu-Cuot P et al (2019) Insights into *Streptococcus agalactiae* PI-2b pilus biosynthesis and role in adherence to host cells. Microbes Infect 29(21):99–103. https://doi.org/10.1016/J.MICINF.2018.10.004
32. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD et al (2015) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43:e15. https://doi.org/10.1093/NAR/GKU1196
33. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A et al (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 1(37):1530–1534. https://doi.org/10.1093/MOLBEV/MSAA015
34. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 1(35):518–522. https://doi.org/10.1093/MOLBEV/MSX281
35. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 8(14):587–589. https://doi.org/10.1038/nmeth.4285
36. Teatero S, Neemuchwala A, Yang K, Gomes J, Athey TBT, Martin I et al (2017) Genetic evidence for a novel variant of the pilus island 1 backbone protein in group B streptococcus. J Med Microbiol 1(66):1409–1415. https://doi.org/10.1099/JMM.0.000588/CITE/REFWORKS
37. Pawlowski A, Lannergård J, Gonzalez-Miro M, Cao D, Larsson S, Persson JJ et al (2022) A group B streptococcus alpha-like protein subunit vaccine induces functionally active antibodies in humans targeting homotypic and heterotypic strains. Cell Rep Med 3:100511. https://doi.org/10.1016/J.XCRM.2022.100511
38. Chen SL (2019) Genomic insights into the distribution and evolution of group B streptococcus. Front Microbiol 28(10):1447. https://doi.org/10.3389/FMICB.2019.01447/FULL
39. Kapatai G, Patel D, Efstratiou A, Chalker VJ (2017) Comparison of molecular serotyping approaches of *Streptococcus agalactiae* from genomic sequences. BMC Genom 1(18):429. https://doi.org/10.1186/S12864-017-3820-5
40. Hayes K, O'Halloran F, Cotter L (2020) A review of antibiotic resistance in group B streptococcus: the story so far. Crit Rev Microbiol 2(46):253–269. https://doi.org/10.1080/1040841X.2020.1758626
41. Dutra VG, Alves VMN, Olendzki AN, Dias CAG, de Bastos AFA, Santos GO et al (2014) *Streptococcus agalactiae* in Brazil: serotype distribution, virulence determinants and antimicrobial susceptibility. BMC Infect Dis 12(14):323. https://doi.org/10.1186/1471-2334-14-323
42. Zeng X, Kong F, Wang H, Darbar A, Gilbert GL (2006) Simultaneous detection of nine antibiotic resistance-related genes in *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization assay. Antimicrob Agents Chemother 1(50):204–209. https://doi.org/10.1128/AAC.50.1.204-209.2006
43. World Health Organization (2017) Group B streptococcus vaccine development technology roadmap: priority activities for development, testing, licensure and global availability of group B streptococcus vaccines [Internet]. Geneva [cited 2022/04/28]. URL: https://www.who.int/publications/i/item/WHO-IVB-17.10
44. Carreras-Abad C, Ramkhelawon L, Heath PT, le Doare K (2020) A vaccine against group B streptococcus: recent advances. Infect Drug Resist 29(13):1263–1272. https://doi.org/10.2147/IDR.S203454
45. Paoletti LC, Julieanne P, Johnson KD, Reinap B, Ross RA, Kasper DL (1999) Synthesis and preclinical evaluation of glycoconjugate vaccines against group B streptococcus types VI and VIII. J Infect Dis 1(180):892–895. https://doi.org/10.1086/314955
46. Nuccitelli A, Rinaudo CD, Maione D (2015) Group B streptococcus vaccine: state of the art. Ther Adv Vaccines 3:76–90. https://doi.org/10.1177/2051013615579869
47. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL et al (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome." Proc Natl Acad Sci U S A 19(102):13950–13955. https://doi.org/10.1073/PNAS.0506758102
48. Lefébure T, Stanhope MJ (2007) Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition. Genome Biol 2(8):R71. https://doi.org/10.1186/gb-2007-8-5-r71
49. Takahashi T, Lee S, Kim S (2021) Genomic characteristics of *Streptococcus agalactiae* based on the pan-genome orthologous group analysis according to invasiveness and capsular genotype. J Infect Chemother 27:814–819. https://doi.org/10.1016/j.jiac.2021.01.008
50. Richards VP, Velsko IM, Alam MT, Zadoks RN, Manning SD, Pavinski Bitar PD et al (2019) Population gene introgression and high genome plasticity for the zoonotic pathogen *Streptococcus agalactiae*. Mol Biol Evol 36:2572–2590. https://doi.org/10.1093/molbev/msz169