

Evidence of a novel sublineage of *Streptococcus agalactiae* in elephants from zoo populations in Germany

Linda Fenske^{1,*}, Elita Jauneikaite^{2,3}, Maria Getino², Yu Wan^{2,4}, Alexander Goesmann¹ and Tobias Eisenberg⁵

Abstract

Streptococcus agalactiae research primarily centres on investigating human and bovine infections, although this pathogen also can be carried and cause infections in a wider range of animal species. Moreover, infections with *S. agalactiae* are posing significant health implications, and recent studies furthermore are highlighting a potential zoonotic risk. Despite the relatively frequent isolation of *S. agalactiae* from elephants, only a few reports document infections in wild and zoo populations. We performed a comparative genomic analysis of 24 elephant isolates from three different zoos in Germany to achieve a comprehensive characterization. Elephant isolates showed pronounced phylogenetic divergence from isolates of other host species, while also forming clusters based on zoo of origin and their genotypes (MLST profiles). Capsular serotypes could not be predicted for the majority of the isolates ($n=20/24$). Several genes, exclusively associated with the elephant host, may underlie the pathogen's capacity to improve its survival and virulence across varied ecological niches. This study not only deepens our understanding of *S. agalactiae* across diverse species and environments but also represents the first whole-genome sequencing characterization of *S. agalactiae* isolates from elephants, helping to expand our knowledge about infections in animals.

Impact Statement

This study provides the first whole-genome characterization of *Streptococcus agalactiae* isolates from elephants, revealing their distinct phylogenetic divergence from group B streptococcus (GBS) isolates of other host species. Our findings suggest that elephant-derived GBS may represent a unique sublineage, with potential adaptations to their host and environment. The absence of identifiable capsular serotypes in most isolates and the presence of host-associated genes highlight the need for further research on the pathogen's evolution, virulence and zoonotic potential. Expanding genomic studies to include isolates from broader geographic regions will be crucial for understanding the role of GBS in exotic animals and its potential impact on both wildlife and public health.

DATA SUMMARY

Raw whole-genome sequencing data for elephant isolates used in this study are available in the European Nucleotide Archive (ENA; <https://www.ebi.ac.uk/ena/browser/home>) project accession number PRJEB28328, and individual GBS isolate accession numbers have been listed in Table S1, available in the online Supplementary Material. Reads and assemblies from Oxford Nanopore MinION sequencing for two GBS isolates, 161002207-5 and 161002207-6 have been deposited to ENA project accession number PRJNA1244849.

Received 31 March 2025; Accepted 03 August 2025; Published 04 September 2025

Author affiliations: ¹Bioinformatics and Systems Biology, Justus Liebig University Giessen, Giessen, Germany; ²NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Department of Infectious Disease, Imperial College London, London, UK; ³Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK; ⁴David Price Evans Global Health and Infectious Diseases Research Group, University of Liverpool, Liverpool, UK; ⁵Department of Veterinary Medicine, Hessian State Laboratory, Giessen, Germany.

***Correspondence:** Linda Fenske, linda.fenske@computational.bio.uni-giessen.de

Keywords: capsule typing; elephants; MLST; penicillin-binding proteins; prophage regions; *Streptococcus agalactiae*; virulence; whole-genome sequencing.

Abbreviations: AMR, antimicrobial resistance; GBS, group B streptococcus; MLST, multilocus sequence typing; ONT, Oxford Nanopore Technologies; PBP, penicillin-binding protein.

Two supplementary figures and three supplementary tables are available with the online version of this article.

001489 © 2025 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

INTRODUCTION

Streptococcus agalactiae, commonly known as group B streptococcus (GBS), is a facultative anaerobic, non-motile, chain-forming, Gram-positive, catalase-negative bacterium that can cause infections in a wide range of hosts. As a human pathogen, GBS plays a significant role as the primary cause of neonatal infections, usually transmitted from mother to newborn during childbirth, provoking conditions such as pneumonia, meningitis or septicaemia [1]. Though GBS is also known to be carried asymptotically within the urogenital tract and the gut, it can also cause infections in adults, especially the elderly. Examples of such infections are urinary tract infections, diabetic foot infections, osteomyelitis and, in rare cases, even more severe conditions like toxic shock syndrome, necrotizing fasciitis, disseminated intravascular coagulopathy and renal dysfunction [2, 3]. GBS is also well known as a pathogen in dairy farming, where it was first reported as a causative agent for the majority of mastitis cases in 1927 [4]. The implementation of the five-point hygiene plan in the 1960s, which included measures such as rapid identification and treatment of infected cows, whole-herd antibiotic dry cow therapy, post-milking teat disinfection, culling chronically infected cows and routine disinfection of milking equipment, significantly reduced the prevalence of GBS infections but has not eliminated its relevance in the dairy industry today. Although GBS research predominantly focuses on human and bovine diseases, GBS has also been reported to cause infections in a variety of animal species including dolphins [5], camels [6], rats [7], fish, seals and other aquatic species [8, 9] and elephants [10].

One of the key virulence factors in GBS is the capsular polysaccharide, which is involved in host immune response evasion [11]. Based on the antigenic properties of the polysaccharide capsule, a serotyping system for GBS was developed [12]. Currently, ten major serotypes (Ia, Ib and II-IX) are recognized, encoded by the capsular locus (*cps*), which comprises 16–18 genes [13, 14]. It is known that the prevalence and distribution of serotypes vary across geographic regions, host species and clinical presentations [15], and some of the serotypes are also associated with different virulence potential [16, 17]. Most studies on GBS focus on human isolates, and though studies in animals have been published, there is a lack of detailed genomic description of GBS isolated from elephants. The earliest report of GBS in elephants dates back to 1997, when the bacterium was identified in pododermatitis lesions of African elephants [18], followed by its detection in Asian elephants in 2008 [19]. Since then, only one other study reported GBS isolated from African and Asian elephants in zoos in Germany [10]. The latter study employed molecular techniques to characterize and compare the identified GBS isolates from elephants to those in other animals and humans, highlighting the potential genetic diversity of GBS causing infections in elephants, including the main finding that most of the GBS from elephants were non-typable using the standard sera or capsular locus typing by multiplex PCR. The non-typable state of GBS isolates is of interest as disease-causing GBS are typically encapsulated, though some studies have reported isolates from humans lacking a capsular locus [20], and GBS isolates from animals have been reported as non-typable [21] or not clearly typable due to previously unknown serotype variants. This could be due to the multiplex PCR designed for human-specific GBS serotypes that does not account for the potential capsular loci diversity of GBS from animals [22].

In the present study, we conducted a comparative genomic analysis of 24 GBS isolates obtained from African and Asian elephants between 2010 and 2023 from zoos in Germany. We aimed to characterize in detail the genomes of these GBS isolates by determining their virulence factors, genotypes based on multilocus sequence typing (MLST), capsular serotypes and potential host-specific genes. This provides valuable insights into the evolution and pathogenicity of GBS populations in less studied host species such as elephants. To our knowledge, this is the first study characterizing GBS isolates from elephants using whole-genome sequencing (WGS).

METHODS

Bacterial isolates

A total of 24 GBS isolates from African and Asian elephants in German zoos were whole-genome sequenced and analyzed. Of these, 23 isolates were obtained during routine bacteriological investigations from elephants in two different zoos (A+B; ~200 km apart) between 2010 and 2016. Twelve of these isolates were obtained from the previous study by Eisenberg *et al.* [10], indicated in Table S1. One isolate (IHIT53690) was obtained from an elephant in Zoo C (~400 km apart from Zoo A; ~200 km apart from Zoo B) in 2023 as part of a routine investigation carried out by the Institute of Hygiene and Infectious Diseases of Animals, Giessen. Metadata of all draft genomes used in this study are provided in Table S1.

Identification by MALDI-TOF MS

Bacterial isolates were selected from the culture plates and then transferred to steel targets according to the manufacturer's instructions (Bruker Biotyper, Bruker Daltonics, Bremen, Germany). Isolates were prepared using the direct smear method and analysed by MALDI-TOF MS using Biotyper (v3.3.1.0).

WGS, raw reads processing and assembly

GBS isolates were streaked on Columbia blood agar plates (Oxoid, Basingstoke, UK) and incubated at 37 °C, 5% CO₂ overnight. Genomic DNA was extracted using the GenElute bacterial Genomic DNA kit (Sigma-Aldrich, Burlington, MA, USA) following the manufacturer's instructions for Gram-positive bacteria with modifications as follows: GBS was lysed in 180 µl G+lysis solution with added 20 µl mutanolysin (Sigma-Aldrich, USA; prepared at 3,000 U ml⁻¹) and 20 µl lysozyme (Sigma-Aldrich, USA; prepared at 100 mg ml⁻¹) added prior to incubation at 37 °C for 1 h. Subsequent steps followed the manufacturer's instructions. DNA was quantified using a NanoDrop spectrophotometer (Thermo, Waltham, MA, USA).

For short-read sequencing, multiplexed DNA library preparation was conducted according to the Illumina protocol, and WGS was performed on a HiSeq X Ten system (Illumina, USA) with 150-cycle paired-end mode. Raw reads were trimmed and filtered with fastp (v0.23.2) [23], and reads were checked for quality with FastQC (v0.11.9) (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads were used for *de novo* assembly into contiguous sequences using Unicycler (v0.5.0) [24]. The draft assemblies were purged from possible errors with Polypolish (v0.5.0) [25] and POLCA from the MaSuRCA (v4.1.0) (Maryland Super Read Cabog Assembler) genome assembly and analysis toolkit [25].

Two isolates (161002207-5, 161002207-6) were selected for long-read sequencing using an Oxford Nanopore Technologies (ONT, UK) MinION Mk1B device. Genomic DNA libraries were prepared using the Rapid Barcoding Sequencing Kit (SQK-RBK004; ONT, UK) and loaded into an R9.4.1 flow cell (FLO-MIN106; ONT, UK). Basecalling was conducted using Guppy (v6.5.7) (ONT, UK) under its super-accuracy mode. Long reads were checked for quality using Filtlong (v0.2.1) (<https://github.com/rrwick/Filtlong>) and assembled using Tricycler (v0.5.5) [26] followed by short-read polishing using Polypolish and pypolca (v0.3.1) [27, 28] as recommended in Ryan Wick's guide to bacterial genome assembly [29]. Polished assemblies were annotated using Bakta (v1.7.0) [30]. Contamination and completeness of all genome assemblies were estimated with CheckM2 (v1.0.1) [31], and furthermore, a taxonomic verification with GTDB-Tk (v2.2.3) was conducted [32].

Population analysis

MLSTs were assigned using mlst (v2.23.0) (<https://github.com/tseemann/mlst>) utilizing the PubMLST *S. agalactiae* database [33] (<https://pubmlst.org/organisms/streptococcus-agalactiae>). Comparative genome analyses for determination of the pan and core genome, as well as singleton genes, were performed with EDGAR (v3.2) [34]. For the creation of the phylogenetic tree, the core genes of all genomes were computed. In the following step, alignments of each core gene set are generated using MUSCLE, and the alignments are concatenated to one huge alignment. The tree was constructed with FastTree using the approximately maximum likelihood method. For visualization of the core genome phylogeny, iTOL (v7) was used [35]. For phylogenetic placement of the elephant-derived isolates, a subset of GBS genomes from different host species was selected and included for comparison. Up to 23 representative genomes for each of the different host species were selected from a database of confirmed GBS genomes [36]. These included genomes from rats ($n=5$), dogs ($n=4$), dolphins ($n=1$), fish ($n=23$), frogs ($n=2$), seals ($n=4$), camels ($n=16$), bovines ($n=17$) and humans ($n=23$) (Table S1). To validate the findings of the phylogenetic analysis, a target-free split k -mer analysis and single-linkage clustering were conducted using SKA (v1.0) [37]. Specifically, k -mer files ($k=15$) were generated with the fasta subcommand under default parameters; pairwise SNP distances were calculated, and SKA clusters were defined at a threshold of 10 SNPs, if they met the minimum identity cutoff of 0.9.

Antimicrobial resistance genes and mobile genetic elements

Antimicrobial resistance (AMR) genes were detected using AMRFinderPlus [38], the Resistance Gene Identifier software [39] and abriTAMR (<https://github.com/MDU-PHL/abritamr>). Putative resistance markers to β -lactam antibiotics were analysed using EDGAR. Therefore, orthologous genes encoding the penicillin-binding proteins PBP1A, PBP1B, PBP2A, PBP2B and PBP2X were identified in the draft genomes. The amino acid sequences of these penicillin-binding protein (PBP) genes were visualized and manually arranged in Jalview [40] based on sequence similarities and subsequently compared to the penicillin-susceptible strain 2603 V/R (GenBank: NC_004116). For reconstruction and typing of potential plasmids, MOB-suite was used (v3.1.8) [41], and the PHASTEST web tool was used to identify prophage regions within the draft genomes [42]. For 12 isolates, capsular serotyping and antimicrobial susceptibility testing results from Eisenberg *et al.* [10] were incorporated in Table S2.

Analysis of the capsular locus genes

A combination of GBS-SBG [43], srst2 (v0.2.0) [44], seq_typing (v2.3.0) (https://github.com/B-UMMI/seq_typing) and KMA (v1.4.12) [45] was used to try to determine the capsular serotype based on the capsular loci genes present. As the majority of isolates remained untypable despite utilizing various tools, an in-depth analysis of the *cps* locus was conducted. The *cps* loci of the elephant-derived GBS isolates were extracted with *in silico* PCR utilizing SnapGene (v8.0.2) (<https://www.snapgene.com/>). The primers used were introduced in a previous study [22]. If no binding sites for these primers were detected within the genomes, alternative primers flanking the *cps* locus at more distant regions were used [20].

Genome-wide association studies

A gene-enrichment analysis, focused on GBS genomes from elephants, was performed using Scoary (v1.6.16) [46] from the gene presence/absence matrix generated by Panaroo (v1.5.0) [47]. The latter was generated from the GFF3 files annotated with Bakta. The host groups were defined as binary phenotypes with a value of 1 assigned to genomes from the elephant host group and 0 to those GBS genomes from all other species.

RESULTS

Sequencing statistics

All 24 genomes were classified as *Streptococcus agalactiae* according to the GTDB taxonomy. The estimated completeness of all draft genomes was above 99.99% with a contamination rate lower than 0.07%. The combined lengths of the assembled contigs range from 1,863,022 to 2,041,196 bp with a G+C content between 35.4 and 35.5 mol%. To evaluate how the genome size and G+C content of the elephant isolates used here integrate into the overall context of GBS genomes, we compared the elephant draft genomes to all GBS genomes currently included in the bacterial web repository BakRep [48]. The BakRep v1 contains 10,359 genomes classified as *S. agalactiae* (as of October 2024). After filtering available genomes for completeness of >95% and contamination rate <1%, 9,984 genomes remained for comparison. As none of the genomes included in the repository listed the elephant as the host species, no further subdivision of the data was made. The GBS genomes included in BakRep had a genome size ranging from 1,798,114 to 2,667,783 bp (only two genomes were larger than 2.5 Mbp, i.e. 2.53 and 2.67 Mbp, respectively, in the whole collection) with a G+C content ranging from 34.9 to 40.3 mol% (Fig. 1). This highlights that the GBS draft genomes from elephants have a mean genome size of 1,930,136 (SD±515,960.30) that is below the genomes contained in BakRep with a mean genome size of 2,067,641 (SD±6,982.26).

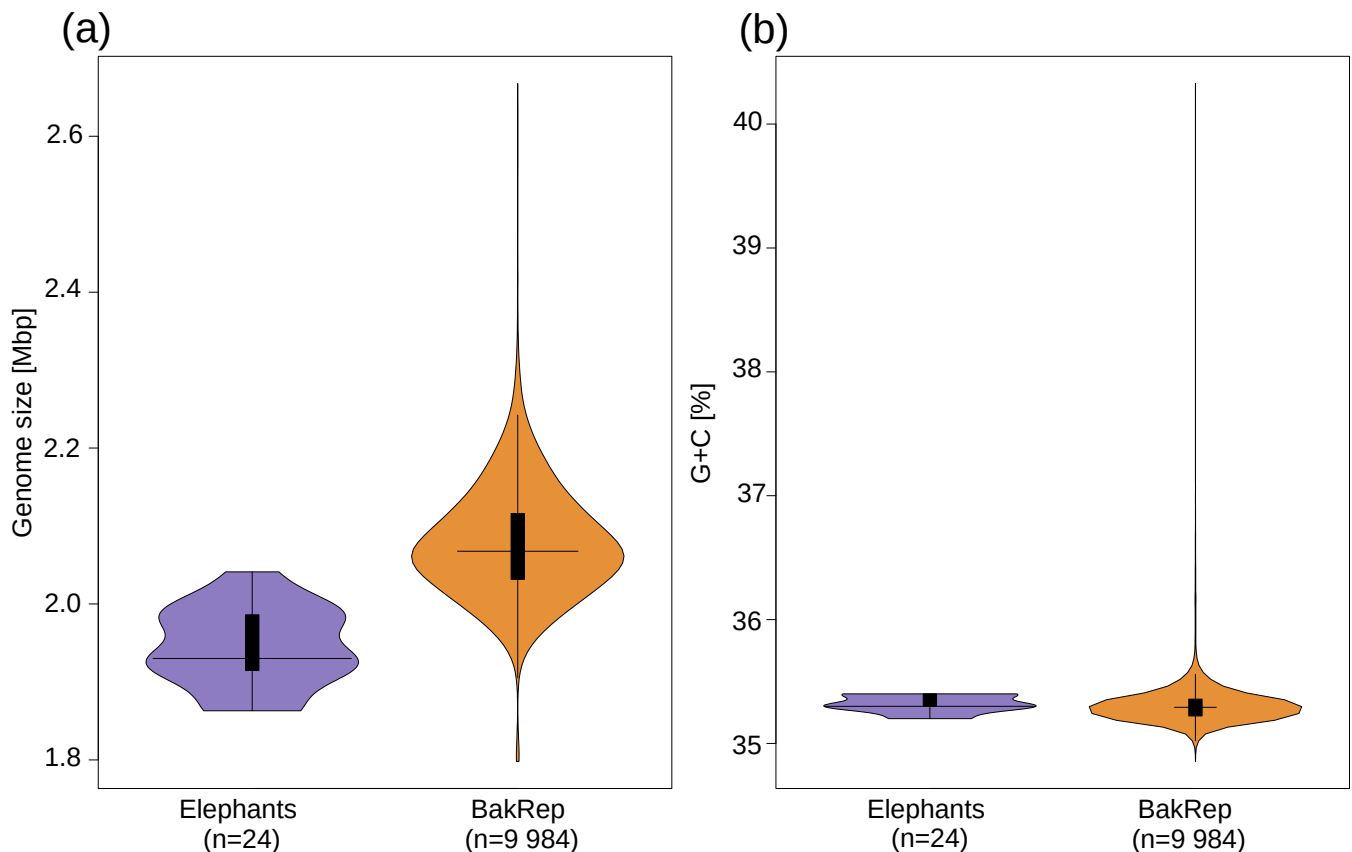


Fig. 1. Comparison between genome size (a) and G+C content (b). GBS isolates from elephants used in this study ($n=24$, purple) and all GBS genomes included in BakRep ($n=9,984$, orange). Horizontal black lines in the middle indicate medians; bold black bars represent interquartile ranges; vertical black lines represent outliers.

Characteristics of GBS genomes from elephants: genotypes, capsular types and AMR

Out of 24 GBS isolates from elephants, 13 (54.17%) were assigned to ST2019, 10 isolates were assigned to sequence type ST2304 and 1 isolate was assigned to ST2305. The latter two STs were first reported in this study and submitted to the *S. agalactiae* MLST database (Table 1). Interestingly, all ST2019 GBS isolates ($n=13$) came from Zoo A, all ST2304 GBS ($n=11$) came from Zoo B and the one GBS ST2305 isolate came from Zoo C; of note, ST2305 was a single-locus variant (SLV) of ST2304. None of the identified STs could be assigned to a clonal complex.

We knew from a previous study [10] that 12 of our elephant-derived GBS isolates were reported as non-typable for their capsular serotype using molecular capsular typing methods. The remaining 12 isolates were not routinely serotyped in the laboratory. After WGS, 4 out of 24 (16.7%) GBS isolates were assigned to serotype Ia. For all other isolates, no clear assignment could be made even after using various accepted GBS capsular typing methods (Table 2).

To analyse the *cps* locus in greater detail, *cps* loci of all elephant-derived isolates were extracted. Binding sites for the primer sequences used by Crestani *et al.* [22] could only be found in the isolates 151006965-2, 151010127, 161002207-6, 161002208-5 and IHIT53690. All these isolates, except IHIT53690, matched the serotype Ia reference sequence (GenBank: LT671983.1) with high similarity, aligning with the results from other applied methods. The region extracted using primer sequences from Creti *et al.* [20] exhibited a deletion in the remaining 19 isolates, similar to the one previously reported by Creti *et al.* (Fig. 2). For IHIT53690, no clear serotype assignment could be made due to several gaps and mismatches (Fig. S2).

Analysis of AMR determinants showed that all 24 GBS isolates likely remain susceptible to most antibiotics due to no acquired resistance genes detected in the genomes, bar *mprF* gene. In PBP1A, all isolates shared a V742A substitution and a deletion of four amino acid residues (739–742) previously described but not associated with reduced susceptibility to β -lactams [49]. In PBP1B, all isolates had an A95D substitution, already detected in GBS from Japan associated with reduced penicillin resistance [50], and in 13 out of 24 (54.17%) isolates, a V64I substitution was identified (not previously described). In the PBP2A, 13 out of 24 (54.17%) isolates had the E18K substitution, and 10 out of 24 (41.76%) had the S394G substitution (not previously described). All isolates showed the V80A substitution in the PBP2B, which was previously detected in other studies [49–51], while 13 out of 14 (54.17%) isolates had a D572N substitution in the PBP2X gene (not previously described). Notably, the 13 isolates with the E18K substitution in PBP2A and the D572 substitution in PBP2X all originated from Zoo B, while the 10 isolates with the S394G substitution came from Zoo A. However, phenotypic testing done in the previous study by Eisenberg *et al.* [10] confirmed high susceptibility to penicillin G and other β -lactams, including those with identified PBP mutations. Additionally, 16 of the 24 tested isolates exhibited low-level phenotypic resistance to gentamicin (Table S2). All substitutions and results of previously done antimicrobial susceptibility testing are listed in Table S2.

Table 1. MLST results for 24 isolates of *S. agalactiae* from elephants analyzed in this study utilizing the MLST scheme of Jones *et al.* [74]. The newly identified STs in this study are highlighted in bold

Isolate	Allele							ST
	<i>adhP</i>	<i>pheS</i>	<i>atr</i>	<i>glnA</i>	<i>sdhA</i>	<i>glcK</i>	<i>tkt</i>	
10-7-D-02041; 141000096/2; 141014875; 151001913- 2; 151002450-1; 151002450-2; 151003337-1; 151003337-2; 151003441-1; 151003441-2; 151003441-3; 151008126; 151010216	20	13	15	1	15	9	5	2019
161002207-5; 161002207-6; 161002208-5; 151004802; 151005628- 1; 151005628-2; 151006965-1; 151006965-2; 151006967; 151010127	16	17	4	2	4	9	5	2304
IHIT53690	16	23	4	2	4	9	5	2305

Table 2. Serotyping results for 24 isolates of *S. agalactiae* from elephants obtained with different *cps* typing tools

Isolate	Tool		
	GBS-SBG (best match)	Seq_typing (best match)	KMA (best match)
161002207-5	NT	NT	VI (score: 15264; coverage: 1.30%)
161002207-6	Ia (coverage: 99.98%)	Ia (sequence covered: 95.74%)	Ia (score: 1111688; coverage: 85.85%)
161002208-5	III-3 (coverage: 78.85%)	Ia (sequence covered: 97.03%)	Ia (score: 1305804; coverage: 87.78%)
10-7-D-02041	NT	NT	VI (score: 12116; coverage: 1.28%)
141000096/2	NT	NT	VI (score: 13035; coverage: 1.27%)
141014875	NT	NT	VI (score: 13505; coverage: 1.28%)
151001913-2	NT	NT	II (score: 12165; coverage: 2.35%)
151002450-1	NT	NT	VI (score: 15747; coverage: 1.29%)
151002450-2	NT	NT	VI (score: 19190; coverage: 1.28%)
151003337-1	NT	NT	VI (score: 14669; coverage: 3.12%)
151003337-2	NT	NT	VI (score: 12245; coverage: 1.26%)
151003441-1	NT	NT	VI (score: 16389; coverage: 1.29%)
151003441-2	NT	NT	VI (score: 11571; coverage: 1.29%)
151003441-3	NT	NT	VI (score: 16338; coverage: 1.34%)
151004802	NT	NT	VI (score: 14957; coverage: 1.29%)
151005628-1	NT	NT	VI (score: 17042; coverage: 1.29%)
151005628-2	NT	NT	VI (score: 15494; coverage: 1.34%)
151006965-1	NT	NT	VI (score: 13236; coverage: 1.29%)
151006965-2	Ia (coverage: 99.99%)	Ia (sequence covered: 97.03%)	Ia (score: 1127495; coverage: 85.73%)
151006967	NT	NT	VI (score: 12073; coverage: 1.27%)
151008126	NT	NT	VI (score: 18202; coverage: 1.29%)
151010127	Ia (coverage: 100%)	Ia (sequence covered: 96.65%)	Ia (score: 1139070; coverage: 86.66%)
151010216	NT	NT	VI (score: 15068; coverage: 1.27%)
IHIT53690	VII (coverage: 56.55%)	IX (sequence covered: 65.85%)	IV (score: 995859; coverage: 80.59%)

Mobile genetic elements: prophages identified in GBS isolates

There were neither complete nor partial plasmid sequences identified in any of the 24 isolates. The prophage analysis suggested that 13 out of 24 (54.17%) isolates had predicted intact phages: 7 out of 24 (29.17%) isolates had *Streptococcus* phage T12 (NC_028700), 2 out of 24 (8.33%) isolates had *Streptococcus* prophage 315.1 (NC_004584.1) and 1 out of 24 (4.17%) had an intact *Staphylococcus* phage SPbeta-like (NC_029119.1), while another isolate carried a questionable version of the same phage. Two out of twenty-four (8.33%) isolates had Bacteriophage Shelly (NC_041909.1), 1 out of 24 (4.17%) had *Bacillus* phage G (NC_023719.1) and 1 out of 24 (4.17%) isolates had incomplete *Streptococcus* phage 5093 (NC_012753.1) (Fig. 3).

Pan-genome analysis and phylogenetic analysis of GBS genomes from elephants

Analysis of the gene content of the 24 elephant-derived GBS isolates revealed a pan-genome comprising 2,352 genes, with a core genome of 1,629 genes. When the representative GBS genomes from other host species (as described in the 'Method' section) were included in the analysis, the pan-genome expanded to 5,485 genes, while the core genome decreased to 1,325 genes. Within the phylogenetic analysis, GBS isolates clustered into two broader clusters: one comprising lineages previously reported only in a single host species, such as camel-associated CC609, bovine-associated CC67 and the elephant sequence types identified in this study, and another cluster comprising lineages reported in multiple host species, including human-associated CC23, fish-associated CC7 and rat-associated CC10, and similar consistent with prior findings (Fig. 4) [52]. To further explore the genetic relationships among GBS isolates from elephants, a phylogenetic analysis was conducted exclusively on the 24 elephant-derived isolates (Fig. 5).

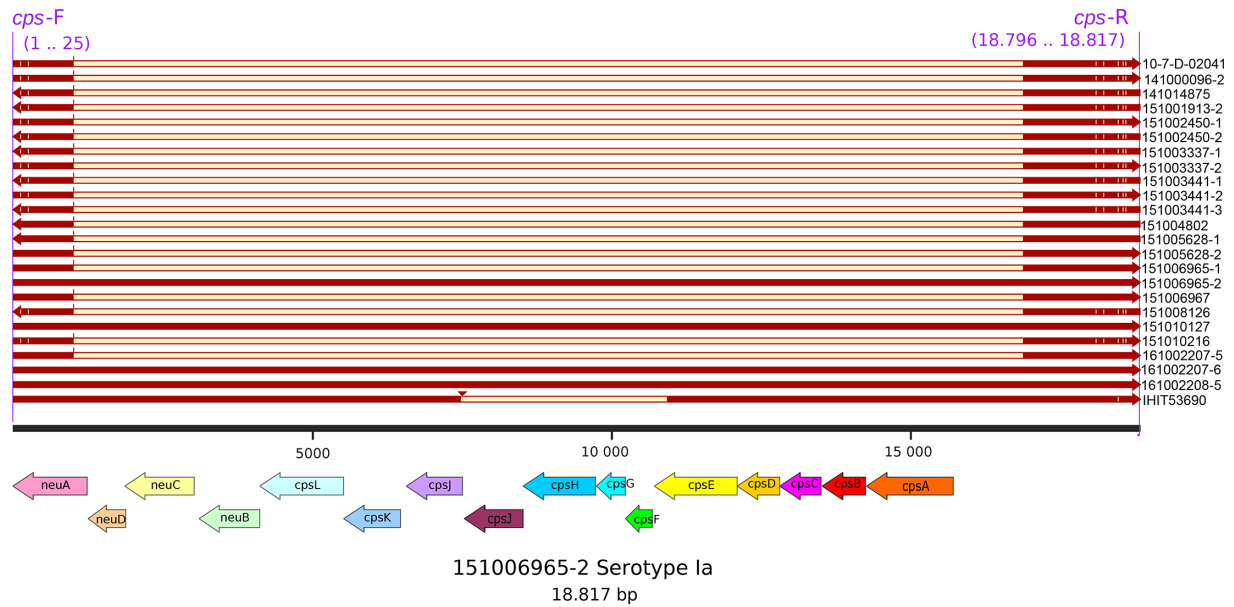


Fig. 2. Multiple-sequence alignment of the extracted *cps* locus of the elephant-derived GBS isolates. The regions between the primers (*cps*-F and *cps*-R) are indicated in dark red, while the deletion is shown in beige. Coloured arrows at the bottom depict the genetic organization of the serotype Ia gene sequence from isolate 151006965-1, which was used as reference. Primer binding sites are indicated at the top in purple. The figure was created with SnapGene (v8.0.2), and annotation of the genes was performed with Bakta (v1.7.0).

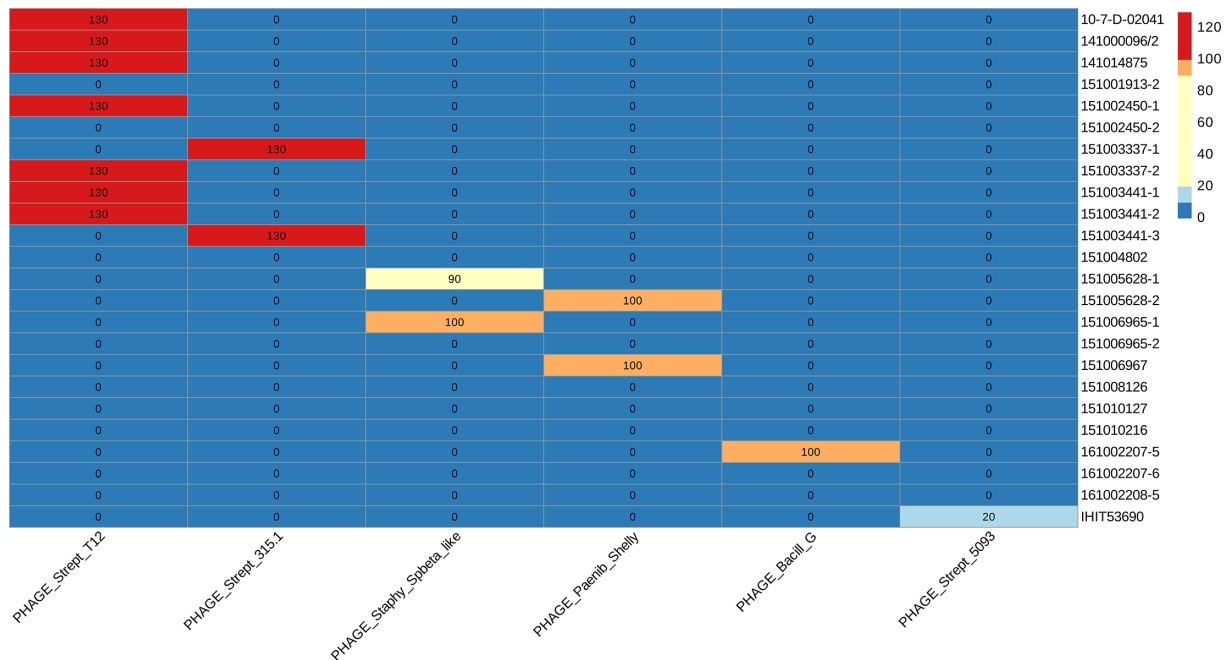


Fig. 3. Heatmap of the identified prophage regions. The score (<70: incomplete; 70-90: questionable; >90: intact) for each isolate is indicated in the boxes. The heatmap was created with R from the results generated with PHASTEST.

The isolates generally clustered by geographical origin, corresponding to the zoo from which the host was sampled. An exception was the most recent isolate from an African elephant in Zoo C (collected in 2023), which shared a common ancestor with four GBS isolates from Asian elephants in Zoo A. Notably, a clear clustering pattern emerged based on both the zoo of origin and the collection year. In Zoo B, all isolates were from African elephants and shared the same sequence type (ST2019). These isolates lacked an identifiable capsular serotype and were primarily isolated in 2015. However, this clade also included two isolates from

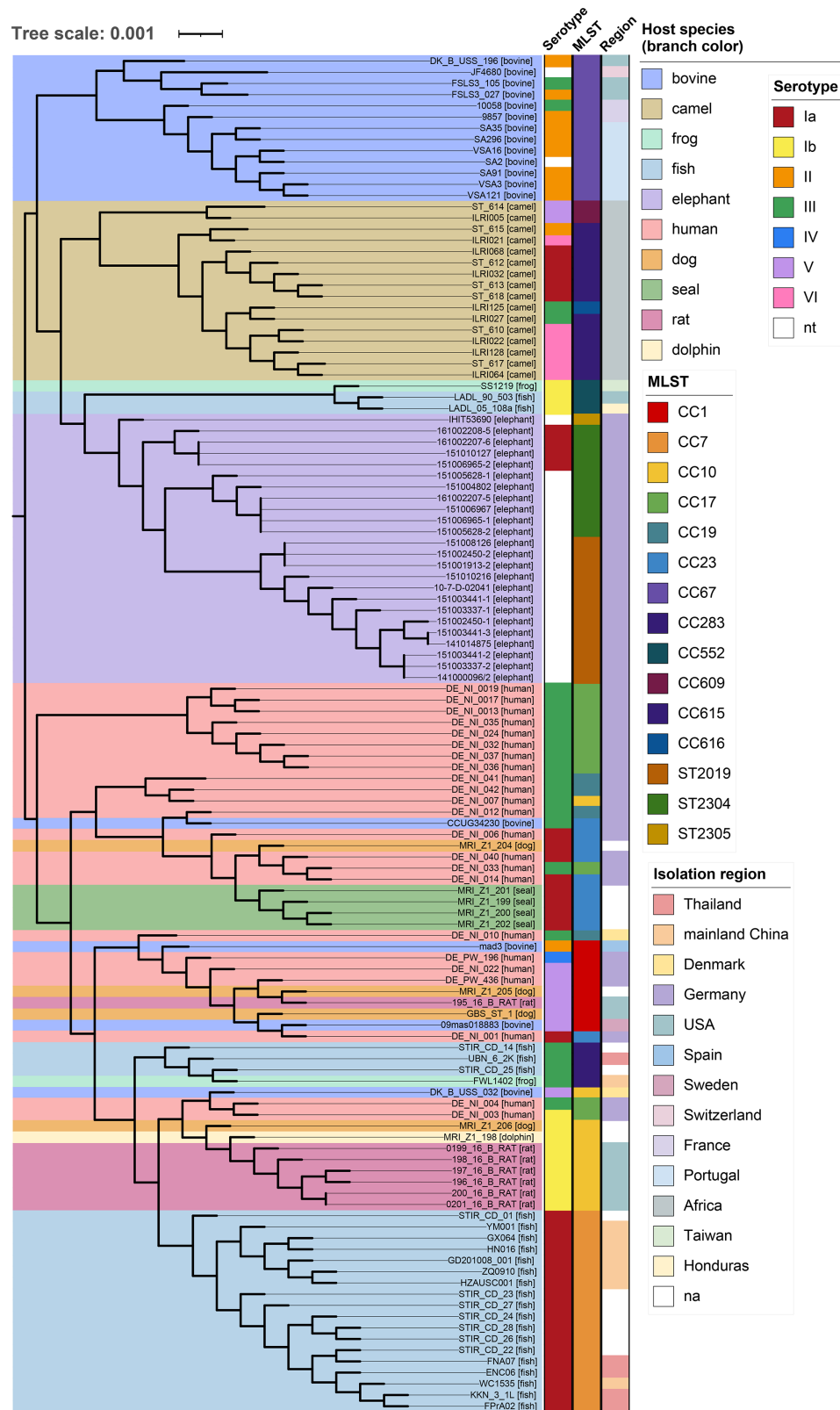


Fig. 4. Phylogenetic analysis of different GBS isolates ($n=121$). The tree is colour-coded according to the respective host species, with the host species also indicated in square brackets. The scale bar corresponds to the number of amino acid substitutions per site, with branch lengths reflecting the relative genetic divergence among genomes. Annotation of the tree was done using iTOL.

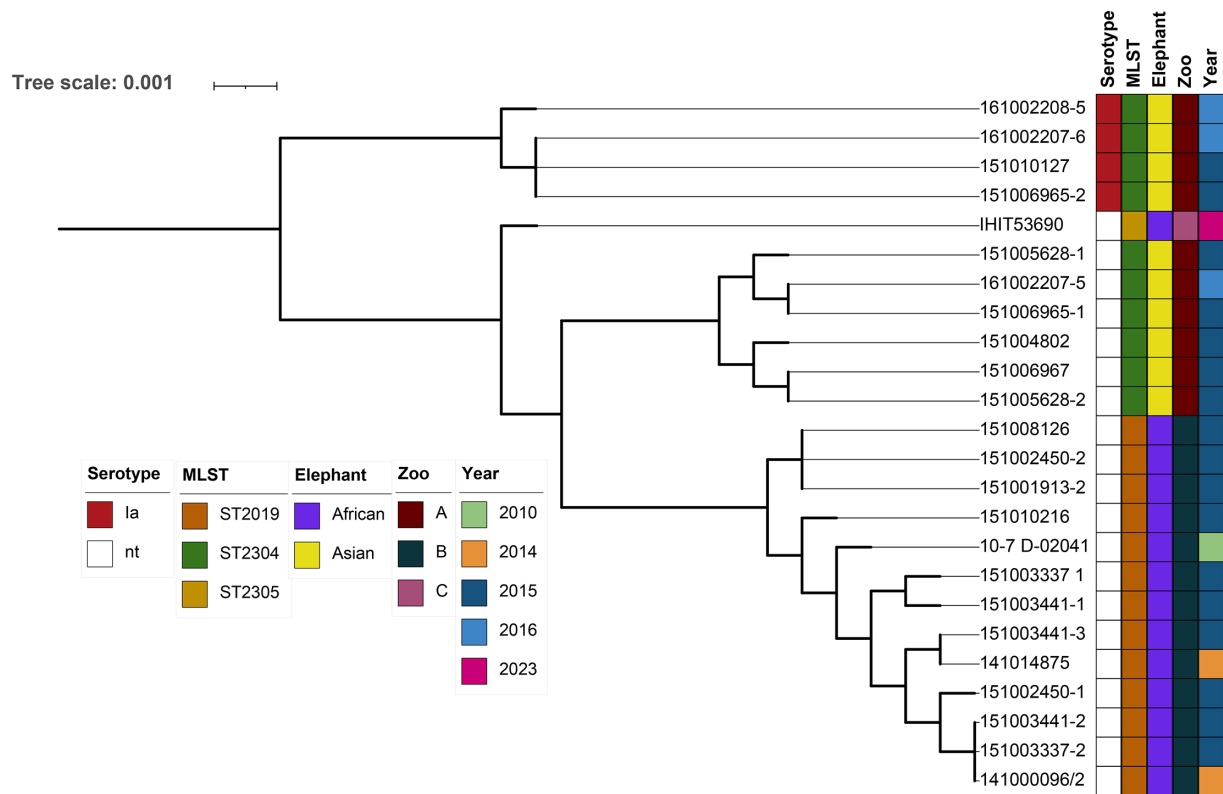


Fig. 5. Core genome phylogenetic tree of the 24 elephant isolates. Coloured blocks showing capsular serotype, MLST, elephant species, zoo of origin and isolation year, as indicated by the legend. The scale bar corresponds to the number of amino acid substitutions per site, with branch lengths reflecting the relative genetic divergence among genomes. The tree was annotated using iTOL.

2014 and one from 2010 that clustered very closely together, suggesting potential persistence of this GBS strain within the zoo population over multiple years.

With regard to Zoo A, isolates 161002208-5, 151010127, 151006965-2 and 161002207-6 formed a distinct clade separate from the other Zoo A isolates. Notably, these were the only isolates with an identifiable capsular serotype, all classified as serotype Ia. As expected, major clades aligned with MLST types: ST2019 isolates clustered together, while ST2304 clustered with the closely related ST2305 isolate, IHIT53690, as ST2305 is an SLV of ST2304 (Fig. 5).

Comparison dataset cluster attribution and genome-wide association investigation

SKA identified a total of 96 clusters among all GBS genomes ($n=121$), with the elephant-derived isolates distributed across six distinct clusters (Table S3). However, the elephant isolates did not cluster with any isolates from other host species (Fig. S1). The median SKA SNP distance was 7,147.5 SNPs across all GBS isolates and 2,344.4 SNPs specifically for the elephant-derived isolates. Within Zoo A, the median SNP distance was 182.5 SNPs, while in Zoo B, it was 5.2 SNPs.

A genome-wide association study was conducted using a pan genome approach to investigate genes that are unique to GBS isolated from elephants. We have found that the *maR* gene was significantly associated with the elephant host, found in all elephant-derived isolates but not in any of the isolates from the other host species. Two additional proteins uniquely associated with all elephant isolates were the Maf family protein and the three acyl-CoA thioester hydrolase/BAAT C-terminal domain-containing proteins. The Lin0465 protein was found in all but one of the elephant isolates (IHIT53690) and was absent in all but two isolates from other host species (P -value: 8.17×10^{-15}).

DISCUSSION

S. agalactiae is a known multi-host pathogen, primarily associated with infections in humans and bovines. However, it is far more frequently isolated as a harmless colonizer, with many strains representing asymptomatic carriage rather than disease. Despite this, GBS is also relatively frequently isolated in endangered species such as wild and zoo elephants. However, to date, no studies have investigated these GBS isolates from elephants using WGS.

The phylogenetic analysis demonstrates that the core genome of GBS isolates from elephants is distinct from GBS genomes isolated from different host species, supporting the niche specificity and potentially host-restricted lineage of GBS [52, 53].

There were challenges in defining the elephant capsule type. In our study, serotype Ia was defined only in four GBS isolates, with the rest of the isolates classed as non-typable due to a deletion of multiple genes in the capsular locus. Recently, it has been reported that a human GBS isolate had deletions of multiple capsular genes [20], the same as identified in our study. Authors of the mentioned study have proposed that a recombination event led to the loss of the whole capsule locus, and large capsular recombination events are known to happen in the GBS population [54, 55]. Given that elephants are a rarely reported host for GBS, as indicated by the limited studies available and, especially, only a small number of elephant GBS genomes available, further genetic and experimental data would be required to confirm if elephant GBS isolates are more likely to be acapsular as part of their species-specific adaptations. It is important to note that GBS isolates investigated here came from infections in elephants, and carriage isolates are not sought after; hence, we are not able to comment on the potential diversity of disease-causing vs asymptomatic or carriage strains of GBS in elephant populations. Notably, current development of GBS vaccines is primarily focused on human GBS vaccines based on ten capsular serotypes or specific surface proteins [56], and current vaccines for fish are also based on serotypes [57]. There is a concern that serotype-based vaccines, when implemented widely, will drive the evolution of non-vaccine serotype and acapsular GBS clones to emerge, similar to what has happened to pneumococcal clones [58]. The risk of this happening may increase if there is horizontal transfer of capsular operons between strains or even across different host species that occur more frequently [22, 59].

Eleven of 24 GBS isolates from elephants had a novel ST2304 ($n=10$) and ST2305 ($n=1$), highlighting the genetic diversity in the GBS isolated from animals that is still unexplored. Notably, isolate IHIT53690 from Zoo C differed only by one allele variant from the isolates originating from Zoo B. The remaining GBS isolates from elephants were classified as ST2019 and were found only in Zoo B. Within the PubMLST database, only one other isolate of ST2019 has been reported, which was also reported as non-typable for its capsular serotype and has been isolated from elephants in the UK (personal communication, Dr Jauneikaite, manuscript in preparation). One of the limitations of our study is that there is potential confounding between host species and geographical location where the elephants were present at the time of sampling, as GBS isolates were available only from one species of elephants present in each of the zoo, potentially highlighting the genotype specificity linked to either elephant species or the zoo. Only one elephant species was present, and it is closely associated with the zoo of origin. Consequently, it is not possible to determine whether the observed STs are influenced primarily by host species or geographic and environmental factors. Furthermore, the dataset spanning over a time period of 13 years with irregular sampling intervals adds another layer of complexity to the interpretation of phylogenetic relationships and strain persistence. These limitations should be taken into account when interpreting findings related to host specificity, transmission patterns and the evolutionary trend of GBS in elephants and other species.

We have found the *mprF* gene present in all elephant-derived GBS genomes. This gene has been reported to play a role in bacterial resistance to cationic antimicrobial peptides, including agents like gentamicin or daptomycin, the latter classified as an antibiotic of last resort used in cases of multidrug-resistant infections [60]. It is known that GBS displays intrinsic resistance to gentamicin due to the low permeability of its cell wall to large molecules such as aminoglycosides [61]. The presence of the *mprF* gene may further contribute to reduced susceptibility, as it has been reported to affect the bacterial cell membrane and decrease susceptibility to certain antibiotic classes. In *Staphylococcus aureus*, for example, *mprF* is known to modulate susceptibility to cationic antibiotics, including the glycopeptide vancomycin, the aminoglycoside gentamicin and moenomycin [62, 63].

PBPs, the enzymes essential for the synthesis of peptidoglycan as components of the cell wall in Gram-positive bacteria, have also been investigated in these isolates and found no mutations that had previously been reported to be linked with decreased susceptibility to penicillin or β -lactam antibiotics. Previous reports indicate that there are specific point mutations in PBPs, primarily PBP2x, that are linked to reduced susceptibility to β -lactam antibiotics [49, 64]. We have found the A95D mutation in PBP1B and the V80A mutation in PBP2B in all elephant isolates, which has been previously reported, although they have only been tentatively linked to penicillin-non-susceptible GBS [49, 50, 64]. None of the other substitutions in PBPs identified in this study have been associated with β -lactam resistance to date, and from limited phenotypic susceptibility data available to us, we were able to show that 12 GBS isolates investigated in our study did not show any evidence for resistance to penicillin [10].

Bacteriophages, including prophages, play a crucial role in the evolution of bacterial genomes and are often associated with enhanced infectivity, pathogenicity and virulence across various bacterial species [65, 66]. Seven GBS isolates in our study carried intact prophage regions with most gene hit counts for *Streptococcus* phage T12. This phage is a prototypic temperate phage first associated with *Streptococcus pyogenes* that carries the *speA* gene coding for erythrogenic toxin A [67]. However, there are no reports of phage T12 in GBS, and none of the elephant isolates in the pan-genome analysis contained the *speA* gene or any related toxin.

When examining the presence of host-specific genes, only one gene was consistently found across all elephant GBS isolates in comparison to other GBS isolates. This gene encodes the MprF enzyme, responsible for the unique synthesis of a cationic glycolipid, Lys-Glc-DAG, which aids in the invasion of human endothelial cells, suggesting a potential role in enhancing

bacterial entry into host cells and promoting disease progression [68]. However, it remains to be confirmed whether it fulfils the same function in elephants. The *mar* gene found exclusively in all elephant isolates encodes the MarR family transcriptional regulator, which plays a key role in controlling the oxidative stress response, a vital environmental sensing function, particularly for pathogenic bacteria [69]. A recent study postulated that the broad host range and ability of GBS to colonize different tissues may be due to the ability of its regulatory systems to recognize and respond to external stimuli, including oxidative and aerobic stress [70]. Maf-like proteins have been proposed to belong to a family of house-cleaning nucleotide hydrolyzing enzymes, which prevent the incorporation of noncanonical nucleotides into cellular DNA [71]. This could also offer an advantage in adapting to different hosts and environmental conditions. Lin0465, which encodes an intracellular PfpI protease in *Listeria*, plays a role in stress response [72, 73]. A study postulated that lin0464, lin0465 and their homologues contribute more to the environmental fitness of these bacteria rather than to their virulence [73]. In summary, this set of genes unique to the elephant isolates examined here collectively confers advantages that enhance adaptability to various hosts and environmental conditions. By enabling GBS to respond effectively to diverse physiological and ecological challenges, these genes likely play a role in its ability to adapt to a host like the elephant and its survivability in different environments.

Conclusion

GBS plays an important role in bacterial infections not only in human and bovine diseases, but also in less often studied host species like elephants. Comparative genomic analysis revealed that, in several aspects, the elephant GBS isolates differ from those of other host species, despite the geographical limitations of our study. Capsular typing of the elephant-derived isolates revealed a high proportion of non-typable strains due to a deletion in the *cps* locus, highlighting their genetic differentiation from serotypes predominantly characterized in human GBS isolates. To further clarify whether elephants may represent a distinct sublineage of GBS, future studies should include more isolates from elephants across diverse geographic regions. All of this could contribute to a better understanding of the zoonotic potential and pathogenic properties of GBS.

Funding information

This work was funded by The Rosetrees Trust and The Stoneygate Trust Imperial College Research Fellowship (M683) awarded to Dr Elita Jauneikaite. This work was also supported by the Justus Liebig University Giessen, Germany, as well as the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) (FKZ 031A532–031A540, W-de.NBI-010). Furthermore, the Hessian Ministry of Agriculture and Environment, Viticulture, Forestry, Hunting and Homeland supports the Hessian State Laboratory.

Acknowledgements

E.J., M.G. and Y.W. are affiliated with the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College London in partnership with the UK Health Security Agency, in collaboration with Imperial Healthcare Partners, University of Cambridge and University of Warwick. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, the Department of Health and Social Care or the UK Health Security Agency. Y.W. is a research fellow funded by the David Price Evans endowment to the University of Liverpool. Microbiological work and sample preparation for whole-genome sequencing of GBS isolates were undertaken at the Colebrook Laboratory, a facility supported by the NIHR Imperial Biomedical Research Centre (BRC). Part of the bioinformatics analysis was performed on equipment purchased as part of the MRC CARP fellowship award MR/T005254/1. We would also like to thank Sophie Aurich from the Institute of Hygiene and Infectious Diseases of Animals, Giessen, for providing isolate IHIT53690.

Author contributions

T.E. and E.J. designed and supervised the study. E.J., M.G. and Y.W. performed the WGS and analyzed the data. L.F. conducted bioinformatic analysis and interpreted the data. Bioinformatic analyses were supervised by A.G. L.F. wrote the first draft of the manuscript. L.F. and E.J. wrote the next versions of the manuscript draft. A.G. and E.J. acquired funding for the study. All authors critically checked and contributed to the final version of the manuscript.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Le Doare K, Heath PT. An overview of global GBS epidemiology. *Vaccine* 2013;31 Suppl 4:D7–12.
2. Farley MM, Strasbaugh LJ. Group B streptococcal disease in nonpregnant adults. *Clin Infect Dis* 2001;33:556–561.
3. Sendi P, Johansson L, Norrby-Teglund A. Invasive group B streptococcal disease in non-pregnant adults. *Infection* 2008;36:100–111.
4. Ruegg PL. A 100-year review: mastitis detection, management, and prevention. *J Dairy Sci* 2017;100:10381–10397.
5. Evans JJ, Pasnik DJ, Klesius PH, Al-Ablani S. First report of *Streptococcus agalactiae* and *Lactococcus garvieae* from a wild bottlenose dolphin (*Tursiops truncatus*). *J Wildl Dis* 2006;42:561–569.
6. Seligsohn D. Pure white gold: subclinical mastitis in dairy camels in Kenya with a special focus on *Streptococcus agalactiae*. Acta Universitatis Agriculturae Sueciae, Swedish University of Agricultural Sciences 2021. <https://pub.epsilon.slu.se/23725/> [accessed 8 September 2021].
7. Shuster KA, Hish GA, Selles LA, Chowdhury MA, Wiggins RC, et al. Naturally occurring disseminated group b streptococcus infections in postnatal rats. *Comp Med Februar* 2013;63:55–61.
8. Delannoy CMJ, Crumlish M, Fontaine MC, Pollock J, Foster G, et al. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. *BMC Microbiol* 2013;13:41.
9. Numberger D, Siebert U, Fulde M, Valentin-Weigand P. Streptococcal infections in marine mammals. *Microorganisms* 2021;9:350.
10. Eisenberg T, Rau J, Westerhüs U, Knauf-Witzens T, Fawzy A, et al. *Streptococcus agalactiae* in elephants – a comparative study with

- isolates from human and zoo animal and livestock origin. *Vet Microbiol* 2017;204:141–150.
11. Avci FY, Kasper DL. How bacterial carbohydrates influence the adaptive immune system. *Annu Rev Immunol* 2010;28:107–130.
 12. Wilkinson HW, Moody MD. Serological relationships of type I antigens of group B streptococci. *J Bacteriol* 1969;97:629–634.
 13. Cieslewicz MJ, Chaffin D, Glusman G, Kasper D, Madan A, et al. Structural and genetic diversity of group B streptococcus capsular polysaccharides. *Infect Immun* 2005;73:3096–3103.
 14. Kapatai G, Patel D, Efstratiou A, Chalker VJ. Comparison of molecular serotyping approaches of *Streptococcus agalactiae* from genomic sequences. *BMC Genomics* 2017;18:429.
 15. Bianchi-Jassir F, Paul P, To K-N, Carreras-Abad C, Seale AC, et al. Systematic review of Group B Streptococcal capsular types, sequence types and surface proteins as potential vaccine candidates. *Vaccine* 2020;38:6682–6694.
 16. Burcham LR, Spencer BL, Keeler LR, Runft DL, Patras KA, et al. Determinants of Group B streptococcal virulence potential amongst vaginal clinical isolates from pregnant women. *PLoS One* 2019;14:e0226699.
 17. Liu Y, Liu J. Group B streptococcus: virulence factors and pathogenic mechanism. *Microorganisms* 2022;10:2483.
 18. Keet DF, Grobler DG, Raath JP, Gouws J, Carstens J, et al. Ulcerative pododermatitis in free-ranging African elephant (*Loxodonta africana*) in the Kruger National Park; 1997. <https://repository.up.ac.za/handle/2263/20739> [accessed 25 July 2024].
 19. Aupperle H, Reischauer A, Bach F, Hildebrandt T, Göritz F, et al. Chronic endometritis in an Asian elephant (*Elephas maximus*). *J Zoo Wildl Med* 2008;39:107–110.
 20. Creti R, Imperi M, Pataracchia M, Alfarone G, Recchia S, et al. Identification and molecular characterization of a *S. agalactiae* strain lacking the capsular locus. *Eur J Clin Microbiol Infect Dis* 2012;31:233–235.
 21. Sørensen UBS, Klaas IC, Boes J, Farre M. The distribution of clones of *Streptococcus agalactiae* (group B streptococci) among herdspersons and dairy cows demonstrates lack of host specificity for some lineages. *Vet Microbiol* 2019;235:71–79.
 22. Crestani C, Seligsohn D, Forde TL, Zadoks RN. How GBS got its hump: genomic analysis of group B *Streptococcus* from camels identifies host restriction as well as mobile genetic elements shared across hosts and pathogens. *Pathogens* 2022;11:1025.
 23. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–i890.
 24. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017;13:e1005595.
 25. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, et al. The MaSuRCA genome assembler. *Bioinformatics* 2013;29:2669–2677.
 26. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, et al. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 2021;22:266.
 27. Wick RR, Holt KE. Polypolish: Short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 2022;18:e1009802.
 28. Bouras G, Judd LM, Edwards RA, Vreugde S, Stinear TP, et al. How low can you go? Short-read polishing of Oxford Nanopore bacterial genome assemblies. *Microbial Genomics* 2024;10:001254.
 29. Wick RR. Choose-your-own-adventure guide to bacterial genome assembly. Zenodo. 2021. <https://zenodo.org/records/7471199> [accessed 25 July 2024].
 30. Schwengers O, Jelonek L, Dieckmann M, Beyvers S, Blom J, et al. Bakta: rapid & standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genomics* 2021.
 31. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 2023;20:1203–1212.
 32. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;38:5315–5316.
 33. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 2018;3:124.
 34. Dieckmann MA, Beyvers S, Nkoumedjo-Fankep RC, Hanel PHG, Jelonek L, et al. EDGAR3.0: comparative genomics and phylogenomics on a scalable infrastructure. *Nucleic Acids Res* 2021;49:W185–W192.
 35. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
 36. Rothen J, Sapugahawatte DN, Li C, Lo N, Vogel G, et al. A simple, rapid typing method for *Streptococcus agalactiae* based on ribosomal subunit proteins by MALDI-TOF MS. *Sci Rep* 2020;10:8788.
 37. Harris SR. SKA: Split Kmer analysis toolkit for bacterial genomic epidemiology. *Genomics* 2018.
 38. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, et al. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob Agents Chemother* 2019;63:e00483–19.
 39. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, et al. CARD 2023: expanded curation, support for machine learning, and resistance prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2023;51:D690–D699.
 40. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25:1189–1191.
 41. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics* 2018;4:e000206.
 42. Wishart DS, Han S, Saha S, Oler E, Peters H, et al. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res* 2023;51:W443–W450.
 43. Tiruvayipati S, Tang WY, Barkham TMS, Chen SL. GBS-SBG – GBS serotyping by genome sequencing. *Microb Genom* 2021;7:000688.
 44. Inouye M, Dashnow H, Raven L-A, Schultz MB, Pope BJ, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 2014;6:90.
 45. Clausen P, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 2018;19:307.
 46. Brynildsrud O, Bohlén J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* 2016;17:238.
 47. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol* 2020;21:180.
 48. Fenske L, Jelonek L, Goesmann A, Schwengers O. BakRep – a searchable large-scale web repository for bacterial genomes, characterizations and metadata. *Microb Genomics* 2024;10.
 49. van der Linden M, Mamede R, Levina N, Helwig P, Vila-Cerqueira P, et al. Heterogeneity of penicillin-non-susceptible group B streptococci isolated from a single patient in Germany. *J Antimicrob Chemother* 2020;75:296–299.
 50. Li C, Sapugahawatte DN, Yang Y, Wong KT, Lo NWS, et al. Multidrug-resistant *Streptococcus agalactiae* strains found in human and fish with high penicillin and cefotaxime non-susceptibilities. *Microorganisms* 2020;8:1055.
 51. Nagano N, Nagano Y, Kimura K, Tamai K, Yanagisawa H, et al. Genetic heterogeneity in pbp genes among clinically isolated group B Streptococci with reduced penicillin susceptibility. *Antimicrob Agents Chemother* 2008;52:4258–4267.

52. Crestani C, Forde TL, Bell J, Lycett SJ, Oliveira LMA, et al. Genomic and functional determinants of host spectrum in Group B *Streptococcus*. *PLOS Pathog* 2024;20:e1012400.
53. Gori A, Harrison OB, Mlia E, Nishihara Y, Chan JM, et al. Pan-GWAS of *Streptococcus agalactiae* highlights lineage-specific genes associated with virulence and niche adaptation. *mBio* 2020;11:00728–20.
54. Neemuchwala A, Teatero S, Athey TBT, McGeer A, Fittipaldi N. Capsular switching and other large-scale recombination events in invasive sequence type 1 group b streptococcus. *Emerg Infect Dis* 2016;22.
55. Khan UB, Jauneikaite E, Andrews R, Chalker VJ, Spiller OB. Identifying large-scale recombination and capsular switching events in *Streptococcus agalactiae* strains causing disease in adults in the UK between 2014 and 2015. *Microbial Genomics* 2022;8:000783.
56. Pena JMS, Lannes-Costa PS, Nagao PE. Vaccines for *Streptococcus agalactiae*: current status and future perspectives. *Front Immunol* 2024;15.
57. Wong KY, Megat Mazhar Khair MH, Song AAL, Masarudin MJ, Loh JY, et al. Recombinant lactococcal-based oral vaccine for protection against *Streptococcus agalactiae* infections in tilapia (*Oreochromis niloticus*). *Fish Shellfish Immunol* 2024;149:109572.
58. Brueggemann AB, Pai R, Crook DW, Beall B. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLOS Pathog* 2007;3:e168.
59. Palmeiro JK, De Carvalho NS, Botelho ACN, Fracalanza SEL, Madeira HMF, et al. Maternal group B streptococcal immunization: capsular polysaccharide (CPS)-based vaccines and their implications on prevention. *Vaccine* 2011;29:3729–3730.
60. Caliot E, Firon A, Solgadi A, Trieu-Cuot P, Dramsi S. Lipid lysination by MprF contributes to hemolytic pigment retention in group B *Streptococcus*. *Res Microbiol* 2024;175:104231.
61. Zakerifar M, Kaboosi H, Goli HR, Rahmani Z, Peyravii Ghadikolaii F. Antibiotic resistance genes and molecular typing of *Streptococcus agalactiae* isolated from pregnant women. *BMC Pregnancy Childbirth* 2023;23:43.
62. Ernst CM, Staubitz P, Mishra NN, Yang S-J, Hornig G, et al. The bacterial defense resistance protein MprF consists of separable domains for lipid lysinylation and antimicrobial peptide repulsion. *PLOS Pathog* 2009;5:e1000660.
63. Ernst CM, Peschel A. Broad-spectrum antimicrobial peptide resistance by MprF-mediated aminoacylation and flipping of phospholipids. *Mol Microbiol* 2011;80:290–299.
64. Hu Y, Kan Y, Zhang Z, Lu Z, Li Y, et al. New mutations of penicillin-binding proteins in *Streptococcus agalactiae* isolates from cattle with decreased susceptibility to penicillin. *Microb Drug Resist* 2018;24:1236–1241.
65. Davies EV, Winstanley C, Fothergill JL, James CE. The role of temperate bacteriophages in bacterial infection. *FEMS Microbiol Lett* 2016;363:fnw015.
66. Lichvariková A, Soltys K, Szemes T, Slobodnikova L, Bukovska G, et al. Characterization of clinical and carrier *Streptococcus agalactiae* and prophage contribution to the strain variability. *Viruses* 2020;12:1323.
67. Weeks CR, Ferretti JJ. The gene for type A streptococcal exotoxin (erythrogenic toxin) is located in bacteriophage T12. *Infect Immun* 1984;46:531–536.
68. Joyce LR, Manzer HS, da C Mendonça J, Villarreal R, Nagao PE, et al. Identification of a novel cationic glycolipid in *Streptococcus agalactiae* that contributes to brain entry and meningitis. *PLoS Biol* 2022;20:e3001555.
69. Wilkinson SP, Grove A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr Issues Mol Biol* 2006;8:51–62.
70. Wang R, Li L, Huang Y, Luo F, Liang W, et al. Comparative genome analysis identifies two large deletions in the genome of highly-passaged attenuated *Streptococcus agalactiae* strain YM001 compared to the parental pathogenic strain HN016. *BMC Genomics* 2015;16.
71. Galperin MY, Moroz OV, Wilson KS, Murzin AG. House cleaning, a part of good housekeeping. *Mol Microbiol* 2006;59:5–19.
72. Harter E, Wagner EM, Zaiser A, Halecker S, Wagner M, et al. Stress survival islet 2, predominantly present in *Listeria monocytogenes* strains of sequence type 121, is involved in the alkaline and oxidative stress responses. *Appl Environ Microbiol* 2017;83:e00827–17.
73. Hein I, Klinger S, Dooms M, Flekna G, Stessl B, et al. Stress survival islet 1 (SSI-1) survey in *Listeria monocytogenes* reveals an insert common to *listeria innocua* in sequence type 121 *L. monocytogenes* strains. *Appl Environ Microbiol* 2011;77:2169–2173.
74. Jones N, Bohnsack JF, Takahashi S, Oliver KA, Chan M-S, et al. Multilocus sequence typing system for group B streptococcus. *J Clin Microbiol* 2003;41:2530–2536.

The Microbiology Society is a membership charity and not-for-profit publisher.

Your submissions to our titles support the community – ensuring that we continue to provide events, grants and professional development for microbiologists at all career stages.

Find out more and submit your article at microbiologyresearch.org