Characterization of Lancefield Group C and G Streptococci Causing Invasive Infections in Humans and Genomic Analysis of Macrolide Resistance, Portugal

Maria Inês Filipe Mendes

Thesis to obtain the Master of Science Degree in

Microbiology

Supervisor: Prof. Dr. Marcos Daniel Caetano de Borges Pinho

Co-supervisor: Prof. Dr. Ana Cristina Anjinho Madeira Viegas

Examination committee

Chairperson: Prof. Dr. Jorge Humberto Gomes Leitão

Supervisor: Prof. Dr. Marcos Daniel Caetano Borges de Pinho

Members of committee: Prof. Dr. Sofia Santos Costa

December 2024

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work presented in this thesis was performed at the Mário Ramirez Lab, Institute for Molecular Medicine, University off Lisbon (Lisbon, Portugal), during the period of September of 2023 and October of 2024, under the supervision of Prof. Dr. Marcos Daniel Caetano de Borges Pinho. The thesis was co-supervised at Instituto Superior Técnico by Prof. Dr. Ana Cristina Anjinho Madeira Viegas.

Acknowledgements

I would like to express my deepest gratitude to Prof. Dr. Marcos Pinho for all the knowledge shared, the understanding, and the time and patience. I sincerely could not have made a better decision when choosing whom to guide me in this journey.

I extend my acknowledgments to Prof. Dr. Mário Ramirez and Prof. Melo Cristino for allowing me to enroll in the amazing team of the Mário Ramirez lab, and to follow a project that I so dearly enjoyed working on. To the whole team of the Portuguese group for the study of Streptococcal Infections, my gratitude for all the help and guidance.

A special thanks to Joana, our “mother hen”, that had the most extensive patience, the best smile, the most pertinent advice, inside and outside of the lab, and who became a very dear friend. This work could not have been possible without you, and my life would have been missing a piece if I had not met you.

To Andreia and Rita, my friends, who were an essential part of these year of work, who lighten the mood and who helped me tremendously in the lab and in the writing of this work. Thank you for all the fun and knowledge that we got to share; it would not have been the same without you beside me.

To my parents and my family, to Zé Manel and to Conchas, because without them I would not be where I am today and would not have all the tools and opportunities I have had, who were always present for the ups and downs of this journey, who picked me up when I needed and who celebrated the victories.

To everyone with whom I’ve crossed paths and that have shaped me into the person I am today, thank you.

Abstract

Streptococcus dysgalactiae subsp. *equisimilis* (SDSE) is a bacterium whose pathogenesis and capacity of infection in humans has been overlooked, but several studies have shown an increasing number of infections. In 2019-2022, 183 SDSE isolates causing invasive infections in Portugal were recovered, mostly from patients over 65 years-old (68%). Most presented the Lancefield group G antigen (68%), while the remaining were group C. No resistance to penicillin, vancomycin, linezolid, trimethoprim/sulfamethoxazole and chloramphenicol was found. High antimicrobial resistance rates to erythromycin and clindamycin were detected (45% and 43%, respectively), with an increase in resistance to both these antimicrobials since 2011. Resistance rates of 28% to tetracycline and 1% to levofloxacin were registered. Around 1% of isolates presented high-level resistance to gentamycin and 3% to streptomycin. Among macrolide and lincosamide resistant isolates, 72.3%, 24.1% and 3.6% presented the $iMLS_B$, $cMLS_B$ and M phenotypes, respectively. The *erm*(A) gene was the most common (73.5%), followed by *erm*(B) (19.3%), *erm*(T) (3.6%) and *mef*(A) (3.6%) genes. Genomic analysis of SDSE invasive isolates recovered in Portugal (2011-2018) showed that macrolide and lincosamide resistance was found in different genetic backgrounds, although clonal complex 17 was associated with the high prevalence of resistance, mainly mediated by *erm*(A). The characterization of resistance genes genomic context showed that most isolates harboring the *erm*(A) gene seemed to carry ICE*Sp2907*, while the presence of the *erm*(B) gene was mainly associated to Tn*3872*. *erm*(T) presence in this population was due to plasmid *p5580,* while *mef*(A) was associated to the presence of distinct mobile genetic elements.

Keywords: *Streptococcus dysgalactiae* subsp. *equisimilis*, invasive infection, antimicrobial, macrolide and lincosamide resistance, mobile genetic element

Resumo

Streptococcus dysgalactiae subsp. *equisimilis* (SDSE) é uma bactéria cuja patogénese e capacidade de infeção em humanos tem sido desvalorizada, porém, estudos descrevem um aumento do número de infeções. Em 2019-2022, 183 estirpes de SDSE causadores de infeção invasiva em Portugal foram recebidas, maioritariamente de doentes com 65 anos ou mais (68%). A maioria pertencia ao grupo G de Lancefield (68%), enquanto as restantes ao grupo C. Nenhuma estirpe apresentou resistência a penicilina, vancomicina, linezolide, trimetoprima/sulfametoxazol ou cloranfenicol. Elevada proporção de resistência a eritromicina e clindamicina foi detetada (45% e 43%, respetivamente), com aumento da resistência a estes antimicrobianos em Portugal desde 2011. Resistências de 28% a tetraciclina e 1% a levofloxacina foram registadas. Sensivelmente 1% das estirpes apresentaram "high-level resistance" a gentamicina e 3% a estreptomicina. Das estirpes resistentes a macrólidos e lincosamidas, 72,3%, 24,1% e 3,6% apresentaram o fenótipo iMLS$_B$, cMLS$_B$ e M, respetivamente. O gene *erm*(A) predominou (73,5%), seguido do gene *erm*(B) (19,3%), *erm*(T) (3,6%) e *mef*(A) (3,6%). Análise genómica de estirpes invasivas em Portugal anteriormente recolhidas (2011-2018) demonstrou a resistência a macrólidos e lincosamidas em diferentes contextos genómicos, porém, o complexo clonal 17 foi associado à resistência, maioritariamente mediada pelo gene *erm*(A). Este gene esteve relacionado com a presença do elemento ICE*Sp2907*, enquanto a presença do gene *erm*(B) esteve relacionada com o elemento Tn*3872*. A presença do gene *erm*(T) na população esteve relacionada com o plasmídeo *p5580*, enquanto a presença do gene *mef*(A) esteve associada à presença de diversos elementos genéticos móveis.

Palavras-chave: *Streptococcus dysgalactiae* subsp. *equisimilis*, infeção invasiva, antimicrobiano, resistência a macrólidos e lincosamidas, elemento genético móvel

Table of contents

Acknowledgements i

Abstract iii

Resumo iv

List of tables vii

List of figures viii

List of abbreviations ix

1. Introduction 1

1.1. Descriptive Overview of *Streptococcus dysgalactiae* subsp. *equisimilis* 1

1.1.1. General overview 1

1.1.2. Classification of SDSE 1

1.2. Identification of SDSE 2

1.2.1. Morphology 2

1.2.2. Lancefield classification 3

1.2.3. Biochemical tests 4

1.2.4. Mass spectrometry 5

1.3. Typing 6

1.3.1. *emm*-typing 6

1.3.2. Multilocus Sequence Typing (MLST) 7

1.3.3. Whole Genome Sequencing (WGS) 7

1.4. Colonization, infection, and epidemiology 8

1.4.1. Colonization 8

1.4.2. Infection and epidemiology 9

1.5. Pathogenesis and virulence 11

1.5.1. Adhesion 12

1.5.2. The M protein and other antiphagocytic factors 12

1.5.3. Toxins and enzymes 13

1.6. Treatment and antimicrobial resistance 14

β-lactams 15

Macrolides and lincosamides 15

Tetracyclines 17

Fluoroquinolones 18

Aminoglycosides 18

2. Aims of the current work 19

3. Materials and methods 21

3.1. Isolate collection 21
3.2. Lancefield grouping 21
3.3. Antimicrobial susceptibility testing 22
3.4. DNA extraction for PCR 23
3.5. PCR for macrolide resistance gene identification 23
3.6. High throughput sequencing (HTS) 24
3.6.1. Genomic DNA extraction 24
3.6.2. Sequencing, assembly and annotation of the genome 24
3.7. Characterization of MGEs carrying macrolide and lincosamide resistance determinants 25
3.8. Statistical significance of the data 26
4. Results 27
4.1. Characterization of SDSE invasive isolates, 2019-2022 27
4.1.1. Patient demographics 27
4.1.2. Characterization of the isolates 27
4.1.3. Antimicrobial resistance 28
4.1.4. Macrolide resistance phenotypes 29
4.1.5. Macrolide resistance genotypes 30
4.2. Characterization of SDSE invasive isolates, 2011-2018 30
4.2.1. Macrolide and lincosamide resistance 32
4.2.2. Molecular typing 33
4.2.3. Characterization of MGEs carrying MLS resistance genes 36
5. Discussion 47
6. Conclusion 53

List of tables

Table 1 - *Streptococcu*s species belonging to Lancefield groups C and G 4
Table 2 - Biochemical characteristics that allow for identification of some medically relevant β-hemolytic Streptococcus species 4
Table 3 - SDSE isolates from human invasive infections (Portugal, 2019-2022) 21
Table 4 - Distribution of the macrolide resistance genes of the invasive SDSE isolates (Portugal, 2019- 2022) that express a macrolide resistance phenotype 30
Table 5 – Evolution of the distribution of the macrolide resistance genes of the invasive SDSE isolates (Portugal, 2011-2018 and 2019-2022) expressing a macrolide resistance phenotype. 32
Table 6 - Distribution of the emm-types by the CCs for all invasive SDSE isolates (Portugal, 2011-2018 34
Table 7 - Characterization of invasive SDSE isolates (2011-2018) harboring the *erm*(B) gene. 39
Table 8 - Characterization of invasive SDSE isolates (2011-2018) harboring the *erm*(T) gene.. 42
Table 9 – Characterization of invasive SDSE isolates (2011- 2018) harboring the *mef*(A) gene 43
Table 10 - Distance matrix of *mef* alleles 43

List of figures

Figure 1 – SDSE morphology 3
Figure 2 - Schematic representation of the S. pyogenes M protein 13
Figure 3 - Macrolide, lincosamide and streptogramin B resistance phenotypes. 15
Figure 4 - Location of SDSE MGE insertion regions relative to the GGS_124 reference genome 16
Figure 5 - Distribution per year of SDSE isolates causing invasive infections in Portugal in 2019-2022. 27
Figure 6 - Antimicrobial resistance rates of invasive SDSE isolates recovered in Portugal from 2019 to 2022 28
Figure 7 - Distribution of the percentage of resistance to antimicrobials from 2019 to 2022 29
Figure 8 - Distribution of the percentage of macrolide resistance phenotypes of the invasive SDSE isolates recovered in Portugal from 2019 to 2022 29
Figure 9 - Distribution per year of SDSE isolates causing invasive infections in Portugal in 2011-2022 31
Figure 10 - Distribution of the percentage of resistance to antimicrobials from 2011 to 2022 ... 31
Figure 11 - Macrolide resistance gene distribution and macrolide susceptibility/resistance phenotype proportion of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 33
Figure 12 - Distribution of the *emm*-types and macrolide susceptibility/resistance phenotype proportion of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 33
Figure 13 - Distribution of the MLSB resistance/susceptibility by the CCs defined by MLST for all invasive SDSE isolates recovered in Portugal from 2011 to 2018 35
Figure 14 - Clonal complex (CC) distribution and MLS_B resistance genes proportion for the invasive SDSE isolates recovered in Portugal from 2011 to 2018. 36
Figure 15 - Distribution of the CC of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 harboring the *erm*(A) gene. 36
Figure 16 – **(A)** Detected complete elements containing the *erm*(A) gene **(B)** Representative of the most commonly found element 38
Figure 17 – Distribution of the CC of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 harboring the *erm*(B) gene. 39
Figure 18 – Detected complete elements containing the *erm*(B) gene 40
Figure 19 - SDSE plasmid *p5580*. 42
Figure 20 - Neighbor-joining tree including the *mef*(A) harboring invasive SDSE isolates recovered in Portugal from 2011 to 2018. 44
Figure 21 – Detected complete elements containing the *mef*(A) gene 45

List of abbreviations

API – Analytical profile index

CC – Clonal complex

CDC – Center for Disease Control and Prevention

CLSI – Clinical and Laboratory Standards Institute

EUCAST – European Committee on Antimicrobial Susceptibility Testing

FBP – Fibronectin binding protein

FCT – Fibronectin and collagen binding protein and T antigen encoding *loci*

GCGS – Group C and G streptococci

HGT – Horizontal gene transfer

HTS – High throughput sequencing

ICE – Integrative and conjugative element

IE – Infective endocarditis

IgG – Immunoglobulin G

IME – Integrative and mobilizable element

MALDI-TOF MS – Matrix-assisted laser desorption/ionization - time of flight mass spectrometry

MIC – Minimum inhibitory concentration

MLS_B – Macrolide, lincosamide and streptogramin B

MLST – Multilocus sequence typing

PCR – Polymerase chain reaction

PJI – Periprosthetic joint infections

SDSD – *Streptococcus dysgalactiae* subsp. *dysgalactiae*

SDSE – *Streptococcus dysgalactiae* subsp. *equisimilis*

SLO – Streptolysin O

SLS – Streptolysin S

SPE -– Streptococcal pyrogenic exotoxins

SSTIs – Skin and soft tissue infections

ST – Sequence type

STSS – Streptococcal toxic shock syndrome

TSA – Tryptone soy agar

TSB – Tryptone soy broth

1. Introduction

1.1. Descriptive Overview of *Streptococcus dysgalactiae* subsp. *equisimilis*

1.1.1. General overview

Streptococcus dysgalactiae subsp. *equisimilis* (SDSE) is a β-hemolytic streptococci that has been commonly overlooked as a pathogen, but most recently, its importance as an animal and human infection agent has surfaced [1–3]. SDSE belongs to a group of *Streptococcus* species known as group C and G streptococci (GCGS) regarding the presence of Lancefield antigens from these groups in their cell wall. As other species here included, SDSE was more commonly viewed as a colonizer and its capacity to cause infection was overlooked. The species included in GCGS have, in the past, been acknowledged as a group and not individually, and, as such, the isolated importance of each species was not adequately addressed. However, it is now understood that SDSE is the GCGS species most commonly linked to human infections.

The disease profile and the virulence mechanisms of this bacterium are very similar to those of *Streptococcus pyogenes*, an outstanding human pathogen that can cause a wide range of diseases and that is responsible for a significant number of deaths each year [4]. As such, the ever-growing understanding of the similarities between the two species highlights the need for a deeper comprehension of the SDSE and its infection mechanisms, as well as of its antimicrobial resistance characteristics and epidemiological portrait.

1.1.2. Classification of SDSE

Although in 1980 the species now known as *Streptococcus dysgalactiae* was for some unknown reason removed from the Approved List of Bacterial Names [5], in 1983, Garvie *et al*. published an article detailing the microbiological characteristics of this species and attributing it a reference strain [6]. Years later, in 1986, Vandamme *et al*. proposed the distinction into two subspecies, one with animal isolates, designated *S. dysgalactiae* subsp. *dysgalactiae*, and the other with human isolates, known as SDSE. The first subspecies accommodated large colony-forming Lancefield groups C and L streptococci isolates with variable hemolysis and that infected animals, whilst the latter included streptococci isolates from Lancefield groups C and G, with large β-hemolytic colonies and that infected humans [7].

However, in 1998, Vieira *et al*. performed some DNA-DNA reassociation experiments and multilocus enzyme electrophoresis (MEE) works. Although agreeing with the separation into two subspecies, Vieira *et al.* proposed the following distinction: *S. dysgalactiae* subsp. *dysgalactiae* (SDSD)

strains were α-hemolytic or γ-hemolytic (non-hemolytic), with Lancefield group C antigen and cause infection in animals, specially bovines; alongside, all SDSE strains were β-hemolytic, presenting Lancefield group C, G or L antigens and can cause infection in humans and animals [3]. In 2001, it was proposed that SDSE and SDSD should be considered two separate species, but this motion was denied by the International Committee on Systematic Bacteriology due to the lack of taxonomic evidence to support this claim, even if the subspecies present different clinical and epidemiological basis. As such, the classification and distinction of subspecies by Vieira *et al*. is the one now most often considered when characterizing the isolates [8].

While some genomic inaccuracies regarding the classification method by Vieira *et al*. were presented, the inconsistences were attributed to horizontal gene transfer (HGT) events that appear to have happened from other β-hemolytic streptococci species, especially from *S. pyogenes* to SDSE [4,9,10]. Indeed, there is evidence of transfer of virulence genes, such as the *emm* gene encoding for the M protein, a gene known to encode a classical virulence gene of *S. pyogenes* [11,12]. Other studies [13,14] have also utilized multilocus sequence typing (MLST) schemes that were originally developed for *S. pyogenes* to confirm the close relationship between SDSE and *S. pyogenes* and also found examples of closely related or identical alleles amongst these two species [15,16].

1.2. Identification of SDSE

Routinely, clinical laboratories base the identification of SDSE isolates on phenotypic characteristics, such as the type of hemolysis, colony size, and the Lancefield group, in conjugation with biochemical systems like analytical profile index (API) galleries or the more recently introduced matrix-assisted laser desorption/ionization - time of flight mass spectrometry (MALDI-TOF MS) technique [17,18]. Genotypic tests, such as polymerase chain reaction (PCR) based assays, were developed, allowing SDSE distinction from other β-hemolytic species, although these are not used on a routine basis.

1.2.1. Morphology

Similarly to other species of the *Streptococcus* genus, SDSE is a non-motile and non-spore forming Gram-positive bacterium composed of cocci/ovoid units, usually in chains [16]. When grown in blood agar (Tryptone Soy Agar, TSA, supplemented with blood), the colonies are large (>0.5 mm), round with a white/greyish colour and present a clear halo around them, indicating a complete lysis of the erythrocytes (β-hemolysis) [19] (**Figure 1**). The size of the colonies can help visually distinguish SDSE from other β-hemolytic streptococci that present small colonies, such as the ones present in the *anginosus*/*milleri* group, which are phylogenetically distinct [19,20]. SDSE is facultative anaerobe and catalase negative

bacterium with complex growth conditions, and, as such, it is incubated in media complemented with blood or serum at an optimal temperature of around 35-37°C[21,22].

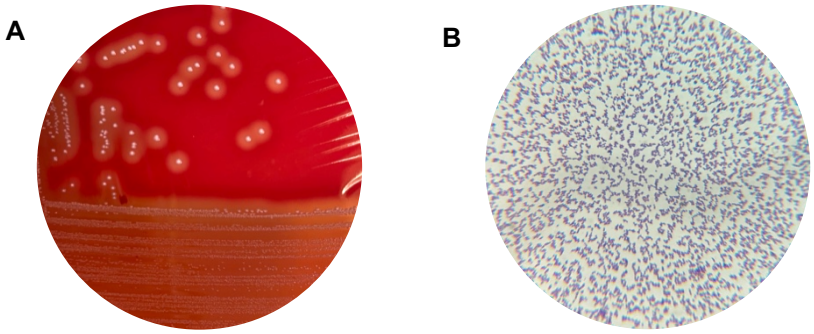


Figure 1 – SDSE morphology. (A) SDSE colonies grown after 24h of incubation at 35°C in Tryptone Soy Agar (TSA) supplemented with 5% (v/v) sheep blood. (B) Gram-staining of a SDSE strain at 1000x zoom.

The β-hemolysis phenotype occurs due to the production of an exotoxin by the cells known as streptolysin S (SLS). This hemolytic toxin is oxygen stable and is encoded by the *sagA* gene. It is also resistant to oxidation and sensitive to trypan blue, which affects its hemolytic activity. SLS not only lysis red blood cells, but is also capable of damaging the membrane of other cell types and cell organelles [23,24]. It is regulated by both the *covRS* and *fasCAX* two-component regulatory systems [25,26].

1.2.2. Lancefield classification

In 1933, Rebecca C. Lancefield, an American bacteriologist, proposed a method of differentiation and classification of β-hemolytic streptococci based on a precipitation reaction between the antigens of the samples and antibodies of sera [27]. This technique allowed for the distinction of groups based on antigenic differences of the polysaccharide, known as carbohydrate C, present in their cell wall [24]. This identification can be performed in the laboratory using test kits commonly available.

Although 20 recognized Lancefield groups exist (A-H and K-V), groups A, B, C, G and D are the most medically relevant [27,28]. As for SDSE strains, these may harbor, most commonly, group C or group G carbohydrates, therefore being included in the group known as group C and G streptococci (GCGS), which includes several species [19]. Less frequently, SDSE may also present group A or group L antigens. Although some of the other GCGS species have been occasionally isolated from human specimens, SDSE is by far the most commonly isolated in human infections [29]. *S. canis* (group G) and *S. equi* subsp. *zooepidemicus* (group C) are well-recognized zoonotic agents that are infrequently isolated from humans.

Table 1 - *Streptococcus* species belonging to Lancefield groups C and G and their respective hosts. [Adapted from 24]

Species	Lancefield group	Host
S. dysgalactiae		
subsp. *equisimilis*	A, C, G, L	Animal, human
subsp. *dysgalactiae*	C	Animal
S. canis	G	Animal, human
S. equi		
subsp. *zooepidemicus*	C	Animal, human
subsp. *equi*	C	Animal
subsp. *ruminatorum*	C	Animal, human
S. phocae	C, F, G, none	Animal

1.2.3. Biochemical tests

Biochemical tests (**Table 2**) are commonly employed for identification to the species level alongside with Lancefield serotyping and hemolysis observation. These tests allow for the distinction of the species based on a set of phenotypic characteristics [30]. There are several kits available for lab use, such as the API® 20 Strep test kit (bioMérieux), an identification system for Streptococcaceae and related organisms, and manual tests for characteristics like catalase, using hydrogen peroxide, and oxidase, with an oxidase test strip [17]. Automated systems for bacterial identification are nowadays commonly used in hospital laboratories, such as the Phoenix system (BD Diagnostics) or the VITEK® 2 system (bioMérieux).

Table 2 - Biochemical characteristics that allow for identification of some medically relevant β-hemolytic *Streptococcus* species. [Adapted from 30]

Species	Lancefield group	Bac	PYR	Cam	VP	Hip	Arg	Esc	Str	Sbl	Tre	Rib	αGal	βGal	βGlu
S. pyogenes	A	+	+	-	-	-	+	V	-	-	NA	-	NA	NA	NA
S. dysgalactiae															
subsp. *equisimilis*	A, C, G, L	-	-	-	-	-	+	V	-	V	+	+	NA	NA	NA
subsp. *dysgalactiae*	C	-	-	-	-	-	+	+	-	-	+	+	-	-	+
S. canis	G	-	-	-	-	-	+	V	+	V	V	NA	NA	NA	NA
S. equi															
subsp. zooepidemicus	C	-	-	+	-	-	+	+	-	-	V	NA	+	+	-
S. anginosus group	A, C, F, G	-	-	-	+	-	+	+	-	-	+	NA	NA	NA	NA

Abbreviations: Bac, bacitracin susceptibility; PYR, pyrrolidonylarylamidase; Cam, CAMP reaction; VP, Voges-Proskauer reaction; Hip, hydrolysis of hippurate; Arg, deamination of arginine; Esc, hydrolysis of esculin; Str, hydrolysis of starch; Sbl, production of acid in sorbitol broth; Tre, production of acid in trehalose broth; Rib, production of acid in ribose broth; αGal, α-galactosidase test; βGal, β-galactosidase test, βGlu, β-glucuronidase test. Results: +, positive reaction > 95% of strains; -, negative reaction >95% of strains; v, variable: 6 to 94% positive; NA, not applicable. Although *S. dysgalactiae* subsp. *dysgalactiae* is not β-hemolytic, it is included for taxonomic reasons.

1.2.4. Mass spectrometry

Different bacterial cultures present distinctive particle and molecule profiles that allow for identification of isolates. MALDI-TOF MS is a technique that enables microbial identification by ionizing chemical compounds present in the bacterial samples into charged molecules, and then measuring their mass to charge ratio by their time of flight [18]. This generates a molecular profile specific to each sample that can be compared with reference spectra available in a database. However, correct species identification depends on good culture conditions and sample preparation, as well as being related to curation of the reference spectra available in the database [31]. Furthermore, there seems to be a limited capacity for discriminating *S. dysgalactiae* subspecies from other streptococci, such as *S. canis,* due to genetic similarities, and so the distinction of SDSE from other species using mass spectrometry may sometimes be troublesome [32].

To solve this question, Nybakken *et al*. [31] demonstrated the importance of tailoring a database in the identification of *S. dysgalactiae*. When identifying SDSE by MALDI-TOF MS, the manufacturer spectra database allowed for correct identification of only 62% of *S. dysgalactiae* isolates. After the curating of the database, the proportion of successful identifications of the isolates as *S. dysgalactiae* increased from 62% to 94%.

1.2.5. Amplification methods

SDSE can also be detected by amplification of specific sequences by polymerase chain reaction (PCR). This technique is mainly used due to how quickly it can be performed and to its cost-effectiveness. However, it is species specific and, as such, it should be used when the presence of this species is suspected. Different target sequences can be used for identification of SDSE, although a fragment of the streptokinase precursor gene is the sequence most commonly searched for when performing a PCR with the aim of distinction between SDSE strains [33].

The 16S rRNA gene and 23S rRNA gene sequences can also be used for identification of streptococci, as they are adequately conserved in each species. The 16S rRNA gene codes for a structural RNA included in the 30S ribosomal subunit, being present in all bacteria, thus being a universal target for bacterial identification, and its function has remained constant throughout time. It is also large enough to arbor statistically relevant information that allows for species recognition [34]. Similarly, the 23S rRNA is a structural RNA that is present in the 50S ribosomal subunit. However, when in comparison with the 16S rRNA gene, the 23S rRNA gene's longer sequence arbors more characteristic sequences and unique insertions and deletions that may allow for a better phylogenetic resolution [35].

1.3. Typing

Different techniques for bacterial typing have been developed in order to differentiate between related bacterial isolates. Strain typing is key in epidemiological surveillance and investigation, as it provides enlightenment into the dynamics of bacterial population and is useful for the characterization and understanding of the epidemiological correlations in an outbreak setting [36].

Throughout time, several phenotypic methods have been designed and used to characterize SDSE strains. While M or T protein serotyping [37], typing of phages and bacteriocins [38,39], and biochemical testing [30] are methods not currently performed, Lancefield group classification is still routinely used in laboratory for SDSE strain typing [27]. Furthermore, techniques such as *emm*-typing and multilocus sequence typing (MLST) have been developed and genome-based methods for SDSE typing have been explored [40,41]. These methods allow for the analyzes of the sequence of one or more genes, respectively, with *emm*-typing being the one most commonly performed for the epidemiological characterization of SDSE strains. Pulse-field gel electrophoresis (PFGE) was, in time, considered the gold standard of genotypic typing methods mainly due to its discriminatory potential [42]. However, most recently, techniques such as MLST have become more commonly used.

1.3.1. *emm*-typing

The M protein present in SDSE is an important virulence factor and can help understand the epidemiology and distribution of the species. In 1928, Rebecca Lancefield developed a system for the typing of differently structured M proteins which involved the production of M-specific typing antisera [40]. This system was used mainly for characterization of the M protein of *S. pyogenes*, however, due to high non-typeability and limited supply of antisera, a molecular method for the typing of the *emm* gene, encoding for the M protein, has been developed [40,43].

Similarly to *S. pyogenes*, the M protein is present in all SDSE strains. It presents a conserved region anchored to the cell wall, a central domain and hypervariable N-terminal. The genetic portion that codifies for this N-terminal region is the one that gets amplified by PCR and sequenced, allowing for *emm*-type determination [43]. The Centers for Disease Control and Prevention (CDC) provides protocols and databases (https://www.cdc.gov/streplab/groupa-strep/index.html) that include *emm* sequences from *S. pyogenes* and SDSE, as well as the rules followed by CDC when attributing a new *emm*-type and subtype [20].

More than 125 *emm*-types have been described for SDSE, with their distribution changing according to the geographical location, as well as some *emm*-types being more prevalent in invasive and other in non-invasive disease [14,44]. Prior research indicates that, from 2011 to 2017, in Portugal, 25 *emm* types were detected among SDSE invasive isolates, with *emm*-type *stG62647* representing more than a quarter of all isolates collected, being the most commonly found *emm*-type [45].

1.3.2. Multilocus Sequence Typing (MLST)

Multilocus sequence typing (MLST) is a technique that allows for the characterization of bacterial isolates in a standardized and reproducible way. It was developed in 1998 as researchers found the methods at the time used presented flaws, such as a poor reproducibility between laboratories and the ineffectiveness at assessing the relationship between isolates [41]. This approach uses housekeeping genes, codifying for proteins with essential functions, where the nucleotide sequence for each of the *loci* is compared to a database with described allele sequences. From here, an allele number is attributed to each *loci*, forming the allelic profile of the strain, to which a sequence type (ST) number is attributed [41,46]. Grouping algorithms, such as eBURST[46] or goeBURST[47], analyze the allelic profile differences between the different STs and group them into distinct clonal complexes (CCs), groups of closely related genotypes. The CCs are named after a founder ST, defined by the ST that differs in only one *locus* from the largest number of STs inside of that same CC [46,47]. With the increase in population, variants with changes in only one of the MLST *loci*, known as single-*locus* variants (SLVs), will arise, which will diversify further and produce double-*locus* variants and triple-*locus* variants, and so forth, and gathered into a CC [46].

The current MLST scheme for SDSE was established in 2009 and uses seven housekeeping genes, codifying for proteins with essential functions: *gki* (glucose kinase), *gtr* (glutamine transporter protein), *murI* (glutamate racemase), *mutS* (DNA mismatch repair protein), *recP* (transketolase), *xpt* (xanthine phosphoribosyl), and *yqiZ*, also referred to as *atoB* (acetyl-CoA acetyltransferase) [15,48].

Research found that, in the period between 2011 and 2017, the SDSE isolate causing invasive infection in Portugal belonged to a total of 76 distinct STs that were grouped into 16 CCs and 8 singletons, suggesting that several different SDSE lineages are responsible for invasive disease [45].

1.3.3. Whole Genome Sequencing (WGS)

In 1976, Fiers *et al.* [49] presented the first complete genome sequence of a non-living organisms, the bacteriophage MS2. After that, researchers went on to decipher the complete genome of some eukaryote and prokaryote organisms using whole genome sequencing (WGS) techniques that allow for the description of the complete genomic sequence of an organism. WGS has allowed for the expansion of some techniques, including MLST, that commonly uses 7 to 8 housekeeping genes for typing, developing typing methods such as core genome multilocus sequence typing (cgMLST) [50]. This way, basing the technique on a much larger group of *loci* will result in a greater typing precision and in a supply of important clinical and epidemiological information, while providing more resolution regarding genomic information, such as antimicrobial resistance, serotyping, and pathogenicity and virulence [36].

Research has shown that SDSE complete genome is comprised of a circular chromosome, with a length of around 2.1 Mbp and a median G+C content of 39.4% (https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/823/). In 2011, Shimomura *et al.* [51] presented the first complete genome sequence of a group G SDSE isolate retrieved from a patient presenting streptococcal toxic shock syndrome, which is commonly referred to as SDSE GGS_124. In 2013, Watanabe *et al.* [52] demonstrated a 90% genome similarity between SDSE strains, whilst illustrating the role of prophages in genomic diversity, as little to no homology was found in prophage regions.

1.4. Colonization, infection, and epidemiology

SDSE can be present in various bodily sites, such as the skin or the pharynx, without causing harm to the human, behaving as a colonizer. When the host's immune defenses become weaker, the bacteria may seize the opportunity and cause infection [53,54]. These sites of colonization and infection serve as reservoir for the human strains of SDSE and to bacterial transmission amongst human hosts [20,24].

1.4.1. Colonization

SDSE is part of the microbiota of the upper respiratory tract, with pharyngeal asymptomatic carriage being more common in adolescents and young adults [53], and it is also present in the skin, the gastrointestinal and in the female urogenital tract [54]. However, the colonization rates might deviate in accordance with different factors.

In fact, Agerhäll *et al*., in 2021, detected pharyngeal colonization by SDSE in 7.8% of the Swedish adolescent and young adult individuals included in this study [53]. McDonald *et al*. [55] also described different throat carriage rates for GCGS in accordance with geographical distribution. Curiously, a study conducted in Brazil found an inverse association between colonization with SDSE and sore throat caused by *S. pyogenes* in children ages three to 15 in slum and non-slum settings. It also proposed that, in line with the results obtained, children in slum settings may have developed immunity against symptomatic infections more promptly due to early exposure to causative agents [56]. SDSE skin colonization is also common, with research demonstrating that anal colonization [57] and the presence of β-hemolytic bacterial on foot skin [58] can function as a reservoir for this bacterium and lead to the recurrence of skin infections like cellulitis and erysipelas. Additionally, a study from Finland, conducted by Jaalama *et al*., also found the presence of GCGS in 2.9% of the rectovaginal samples of pregnant women [54].

1.4.2. Infection and epidemiology

The colonization of human hosts by SDSE is common, with the bacteria's ability of inhabiting the human organism asymptomatically. However, the number of reports of disease caused by GCGS continuously escalates. As such, this bacterium is considered an opportunistic pathogen, as it causes disease in circumstances where the host is susceptible to infection, including periods of weaker immune response or of disruption of the normal microbiota [59]. Higher rates of infection by SDSE are more commonly registered amongst the elderly since, with age, the immune system tends to weaken. Not only that, but older patients more often present some underlying comorbid conditions, such as diabetes mellitus and cardiovascular disease, as well as poor health habits that further deteriorate the capacity of the immune system to respond to an invading microorganism [19].

Non-invasive infections

When the infection occurs in usually non-sterile sites of the host's organism, it is considered to be a non-invasive infection. Different types of non-invasive SDSE infections have been reported, with tonsillopharyngitis and superficial skin and soft tissue infections (SSTIs) being some of the most common non-invasive diseases [19].

Skin and soft tissue infections (SSTIs) are the most common manifestations of SDSE infections, which may present as cellulitis, erysipelas, impetigo, pyoderma, or more superficial wounds [19,60]. These infections may be a gateway for invasive disease as they sometimes allow for the bacteria to enter the bloodstream and reach other tissues. Pre-existing wounds associated to underlying conditions, such as diabetes mellitus, or venous and lymphatic compromise, may be negatively affected by the infection [19].

Tonsillopharyngitis is described as the inflammation of anatomic structures of the upper respiratory tract, leading to a sore and scratchy throat, with discomfort when swallowing [61]. The symptoms due to SDSE infection are very similar to those of tonsillopharyngitis caused by *S. pyogenes*. SDSE is usually linked to outbreaks of tonsillopharyngitis and less commonly related to sporadic cases in young adults, while its responsibility in endemic or sporadic cases in children is less understood [20]. As previously mentioned, SDSE can be a colonizer of the pharynx. To understand the capacity of infection, Turner *et al*. [62] performed throat cultures from college students who presented complaints of sore throat and the results were compared with those of healthy college students that presented no symptoms. It was found that the isolation of group C streptococci presented significant differences, being isolated from 26% of the complaining students, in comparison with 11% of the healthy individuals.

Invasive infections

An invasive infection caused by SDSE is characterized by the isolation of this bacterium from a site that is commonly sterile or from a non-sterile site in combination with signs of necrotizing fasciitis or streptococcal toxic shock syndrome (STSS) [63]. Most of the patients presenting this type of infection by

SDSE are elderly people that possess underlying conditions, or that, in some other way, present a debilitated immune system [64].

Bacteremia is a commonly acquired SDSE invasive infection, being characterized by the isolation of this bacterium from the blood. Given these cases, the bacteria can tour the bloodstream and ultimately cause disease in diverse host tissues. This invasive infection is commonly secondary to SSTIs, and its recurrence is more predominant in patients with a history of skin erosion, cellulitis, or chronic eczema, as well as diabetes and heart disease [65]. Research has shown that bacteremia by SDSE can lead to case fatalities of 2 to 18% [66–68]. In 2009, a population-based study regarding bacteremia caused by β-hemolytic streptococci in Finland [69] reported 314 cases of this invasive infection. From the total number of infections, the most frequent etiological agent was SDSE, being responsible for around 45% (n=140) of cases, whilst the remaining infections were related to group A, group B and other group C β-hemolytic streptococci. Diabetes was common amongst patients with SDSE bacteremia, as well as episodes of cellulitis prior the current infection. The 30-day mortality for patients with bacteremia caused by SDSE was around 15%. Furthermore, an observational study in Taiwan [70] from the year of 1998 to 2004 identified 92 patients presenting group G streptococcal bacteremia, of which 86 cases were confirmed to be caused by SDSE isolates. The majority of the cases were secondary bacteremia due to cellulitis, and the mortality rate was low, of 3%.

Infective endocarditis (IE), meaning the presence of bacteria at the endocardium, is another invasive disease that can be caused by SDSE. It develops after the bacteria enter the bloodstream (bacteremia) and attach to the cells of this tissue of the heart. Although the report of cases of IE caused by β-hemolytic streptococci is uncommon, the disease is characterized by severe manifestations [67,71]. A study by Oppegaard *et al*. [67], in 2017, demonstrated that the median age for cases of endocarditis caused by SDSE in Norway, in the period of 1999 to 2013, was older than patients with the same disease caused by *S. pyogenes* (63 versus 51 years-old). Diabetes mellitus and malignancy were associated with SDSE and the mortality at 30 days was of around 22%. A Canadian study, by Lother *et al*. [71], also demonstrated that, from the patients suffering from bacteremia by SDSE involved in the study, 6% presented endocarditis.

Although not customary, the number of cases of joint and bone infection caused by SDSE seems to be rising, which may be a result of the usually weakened immune system of the hosts [72]. Some cases of osteomyelitis by SDSE have been described, with infection in sites such as the cervical [73] or the lumbar spine [74], or general reports of vertebral osteomyelitis and spondylodiscitis [72,75]. From the reports, patients presenting the types of infection above mentioned usually experience fever, back pain and tenderness, as well as general malaise and weakness. Cases of septic osteoarthritis, i.e. the inflammation of the joints secondary to a microbial infection, have been described in the literature. Amongst them, Mesa-del-Castillo *et al*. [76], in 2016, reported a case of a three-year-old toddler presenting an acute and severe osteoarticular infection of the left hip. The toddler showed extensive periodontal disease, which, by the deterioration of the gums, may allow the bacteria to enter the bloodstream. Some case studies have also reported

periprosthetic joint infections (PJI), meaning the infection of a bone structure in close relation to an implant. In 2020, Erden *et al*. [77] and Pham *et al*. [78] separately reported cases of PJI caused by SDSE, with both patients presenting swelling and redness of the skin surrounding the infection.

Invasive SSTIs by SDSE are characterized by infection of the skin or soft tissue with the additional isolation of bacteria from a normally sterile body site, such as blood samples or surgical specimens [60]. These types of infections can lead to severe necrotizing skin and soft tissue diseases such as necrotizing fasciitis, affecting the muscular fascia, necrotizing myositis, when the infection occurs in the muscle, or in localized sites of the body, being Fournier's gangrene an example [19,79]. Injection drug users, burn patients in need of skin grafts and elderly patients with underlying immunosuppressive conditions are the most likely to develop these types of infections [20]. Bruun *et al*. [80] characterized the necrotizing soft tissue infections by *S. pyogenes* and SDSE reported in Norway from the years of 2000 to 2009. Out of the 70 reported cases, 61 were caused by *S. pyogenes*, whilst the remaining nine were associated to SDSE. The in-hospital case fatalities were of 11% and 30% for *S. pyogenes* and SDSE, respectively.

Streptococcal toxic shock syndrome (STSS) is a life-threatening condition that may, in rare occasions, be caused by SDSE, although *S. pyogenes* is the most frequent etiologic agent. It is related to the combination of the effect of toxins and other enzymes produced by the bacteria and the host's immune response and is characterized by organism-wide symptoms, such as fever, hypotension, skin rash and organ damage and failure [81]. In 2017, a study by Ikebe *et al*. [82] contemplated all reported STSS cases caused by SDSE in Japan from the year of 2014 to 2016. In total, 173 cases were described, and the mortality rate was 31.2%. Other studies also provide data supporting the claim that the number of cases of STSS by SDSE in Japan has been increasing [83,84]. In 2013, Hagiya *et al*. [85] reported the death of an 88-year-old man due to a SDSE infection that led to STSS. It was proposed that the bacterium might have been introduced into the patient's organism 30 hours prior via an intramuscular injection, demonstrating how rapid the progression of this disease can be.

1.5. Pathogenesis and virulence

For all microorganisms, evolution leads for the acquisition and preservation of the most useful genomic traits, and consequent phenotypical characteristics, that prompt species survival and perseverance. Such happens for virulence and pathogenicity factors, which allow the bacteria to easily infect, survive and to cause damage to the host's cells. As mentioned previously, the infection patterns and mechanisms of SDSE are very similar to those of *S. pyogenes*, which are believed to have been acquired both by horizontal genes transfer (HGT) events or simply due to common ancestry between the species [10]. In fact, Shimomura *et al*., in 2011 [51], sequenced the first complete genome of a group G SDSE strain and compared the sequence against those of 11 other streptococcal species, including *S. pyogenes*. The SDSE strain had the highest sequence similarity with *S. pyogenes*, of around 72%. These virulence factors

can control diverse aspects, such as adhesion, antiphagocytic activity and toxin and enzyme production, amongst other features.

1.5.1. Adhesion

The capacity of adhesion is a crucial attribute of bacteria for the development of infection, and therefore, microorganisms have evolved specific mechanisms that allow for the adhesion to extracellular components of the host's cells [86]. SDSE mainly utilizes fibronectin as an attachment target present in the soluble form in the extracellular matrix of host cells, a high-molecular-weight glycoprotein that is responsible for substrate adhesion in eukaryotic cells, as well as other functions related to the extracellular network [87]. This way, the fibronectin-binding proteins (FBPs) of SDSE are capable of attaching to specific target cells, like the ones of the upper respiratory tract of the human host, enabling colonization of this site and the development of an infection.

Two FBPs have been identified in SDSE, FnbA and FnbB, and these proteins bind fibronectin by its N-terminal region [88]. GfbA is another FBP, which binds to fibronectin through a repeat region and an upstream region that contains the LAGESGET amino acid motif [24]. This protein shows a significant amino acid identity with *S. pyogenes* FBPs and is able to induce adhesion phenotypes in nonadherent strains once the *gbfA* gene is introduced [89].

Some FBPs, such as T antigen genes and *pili* encoding genes, are chromosomally located in fibronectin and collagen binding protein and T antigen encoding *loci* (FCT) in *S. pyogenes* [90]. Research has shown the presence of regions with high similarity to these FCTs regions in SDSE [91,92]. SDSE can also adhere to other extracellular matrix compounds, such as fibrinogen, laminin, plasminogen and vitronectin. It was demonstrated that the M protein of SDSE can also play a role in adhesion as it can bind collagen [24]. Additionally, SDSE possesses a laminin binding protein, Lmb, encoded by the *lmb* gene. This gene was first found in *Streptococcus agalactiae*, but a virtually identical copy was found in SDSE. The adhesion to this glycoprotein can be important in bacterial colonization and invasion into the bloodstream [20,93,94].

1.5.2. The M protein and other antiphagocytic factors

The expression of the M protein is valuable for streptococcal immune evasion. It was first described in 1927 by Rebecca Lancefield as being important for *S. pyogenes* virulence, primarily due to its antiphagocytic properties [40]. One of the mechanisms by which this protein can inhibit opsonization is by functioning as an immunoglobulin G (IgG) Fc receptor, meaning it can bind the crystallizable fragment (Fc) of IgG proteins [95,96]. *S. pyogenes* and SDSE M proteins have similar biological, immunochemical, and

structural features, being encoded by the same *locus* in both species – the *emm* gene. This fibrillar coiled-coil protein is surface-exposed and can allow typing of both SDSE and *S. pyogenes* into various *emm*-types due to its variability [97]. Its N-terminal region presents high heterogenicity in its amino acid sequence, whilst the C-terminal region, anchored to the cell wall, is more conserved along the different *emm*-types [97–99]. Repeats with distinct length and composition are integrated throughout the protein, with the conserved portion presenting C and D repeats, while the hypervariable portion of the protein presents A and B repeat regions. A signal peptide is then located at the protein's N-terminus [97,100,101] (**Figure 2**).

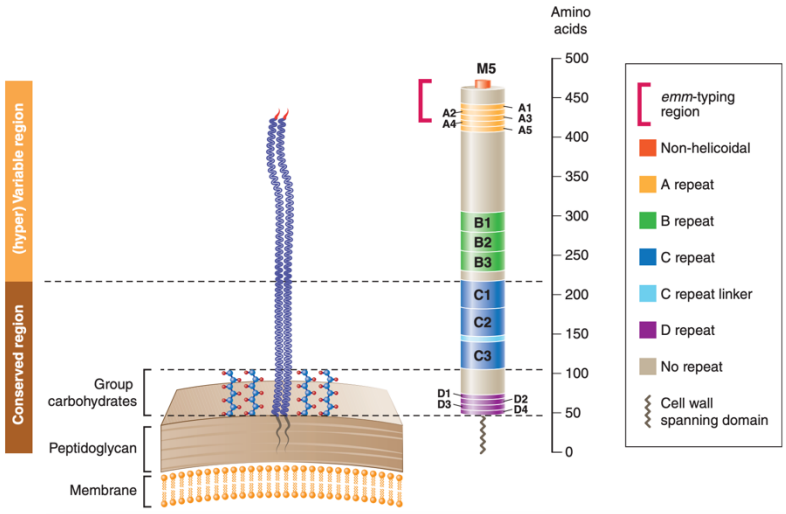


Figure 2 - Schematic representation of the *S. pyogenes* M protein. [Adapted from [97]

One study has shown that the protein resulting from a specific SDSE *emm*-type can have important antiphagocytic action by binding fibrinogen [102]. With functional similarities to the M protein, SDSE also possesses protein G, an IgG Fc receptor, that exclusively binds this protein with high affinity [103,104].

Additionally, a C5a peptidase has been identified in SDSE and has the capacity to cleave complement factors circulating in the blood, proteins that enhance phagocytosis [24].

1.5.3. Toxins and enzymes

The capacity of proliferation of the bacterial cells in the host tissues is essential for the dissemination of the infection. For this, SDSE isolates use an array of streptococcal toxins and enzymes. Streptokinase is an enzyme that can form complexes with plasminogen, where the plasminogen binding site is exposed, which catalyzes the conversion of other plasminogen molecules into plasmin [105]. This molecule is rather important for tissue invasion since plasmin is essential in the breakdown of tissue barriers

[24,106]. Streptokinase is encoded by the *ska* gene, which presents a high level of variation, leading to different level of plasminogen activation and differences in tissue tropism [106].

SDSE produces not only SLS, as mentioned above, as well as Streptolysin O (SLO). SLO is an example of a thiol-activated cytolysin, encoded by the *slo* gene, and it plays a role in bacterial virulence. This oxygen-labile toxin is able to damage the cytoplasmic membrane of several host cell types, such as erythrocytes, leading to cell lysis [23]. Not only that, but it is also capable of depleting the host cells of their energy storage [24,107].

Streptococcal pyrogenic exotoxins (SPE) are well established virulence factors of *S. pyogenes*, being classified as superantigens as they can lead to an aggravated proliferation of immune T cells and release of inflammatory cytokines, promoting inflammation and leading to STSS [24,108]. Given that SDSE also causes STSS, it has been hypothesized that SDSE strains also produces SPEs. The *spegg* or *speG^{dys}* gene is the only repeatedly identified gene encoding for a SPE in multiple SDSE isolates, being orthologous to the gene *speG* of *S. pyogenes* [109,110]. In fact, comparing *S. pyogenes* and SDSE genomic profiles has provided evidence that no HGT events lead to the acquisition of this gene but rather it being the direct descendant of a common ancestor [91]. Still, *sepgg*'s role in virulence is yet not fully understood, with evidence deeming no association of the clinical presentation of the disease and the presence of this gene [111]. Furthermore, some studies failed to detect any mitogenic activity in *spegg* positive SDSE isolates [109].

1.6. Treatment and antimicrobial resistance

When treating SDSE infections, penicillin and other β-lactam antimicrobials, which act by inhibiting peptidoglycan synthesis, are considered the drugs of choice. Such occurs because, similarly to other β-hemolytic streptococci, there seems to be virtually no resistance of SDSE to this class of antimicrobials [19].

With the aim of avoiding poor response to antimicrobial treatment during SDSE infections, the use of other antimicrobials alongside a β-lactam treatment has been advised [20]. The addition of aminoglycosides may benefit treatment, as these drugs belong to a class of antimicrobials with a different effect on the cell (inhibition of protein synthesis) and were shown to act synergistically with β-lactams. The use of clindamycin, a lincosamide, in combination with a penicillin treatment as also shown to be beneficial for invasive *S. pyogenes* infections. Since the characteristics of both *S. pyogenes* and SDSE infections are similar, treatment with clindamycin in combination with a β-lactam may be valuable [112]. In instances where the patient presents an allergy to penicillin or other β-lactam, a macrolide or lincosamide treatment may also be adequate.

β-lactams

Although being the number one line of treatment against SDSE infections, a single study described the occurrence of resistance to penicillin. Fuursted *et al*. [113] investigated four cases of SDSE penicillin resistance in three patients with bacteremia, in Denmark. The isolates all presented some amino acid changes in their penicillin-binding proteins (PBPs), conferring resistance to this drug. It was found that the four isolates belong to a single clone and that there was a possible epidemiological link between the patients. No other studies in the literature detected SDSE penicillin resistant strains. However, some cases of tolerance to the use of *β*-lactams have been reported in the past. A study conducted by Biedenbach *et al*., in 2006 [114], showed that some of the group GCGS isolates under consideration exhibited a higher minimal inhibitory concentration (MIC) value for penicillin, in both Europe and North America. The mechanisms responsible for this observation were not further discussed.

Macrolides and lincosamides

Although macrolides and lincosamides may be advised in the treatment of non-invasive and invasive SDSE infections, respectively, the rates of non-susceptibility to these antimicrobials might be substantial, highlighting the importance of antimicrobial susceptibility testing. Macrolides and lincosamides, as well as streptogramin B antibiotics, have the same cell target, aiming for the 50S ribosomal subunit, where they attach to, inhibiting protein synthesis [115]. However, the cell can overcome this mechanism, and resistance may express itself in three distinct phenotypes, which can be distinguished in the laboratory, as represented in **Figure 3**: the M phenotype (**A**), where the bacteria exhibits a resistance to 14- and 15-membered macrolides; the $cMLS_B$ phenotype (**B**), in which the bacteria has a constitutive resistance to macrolides, lincosamides and streptogramin B; and the $iMLS_B$ phenotype (**C**), where the bacteria is resistant to macrolides and shows inducible resistance to lincosamides and streptogramin B in the presence of a macrolide in antimicrobial susceptibility testing [116,117].

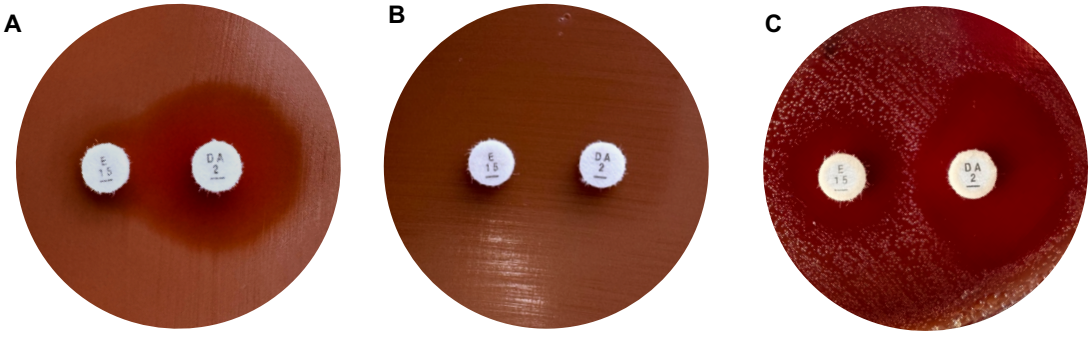


Figure 3 - Macrolide, lincosamide and streptogramin B resistance phenotypes revealed by disk diffusion testing. (A) M phenotype. (B) $cMLS_B$ phenotype. (C) $iMLS_B$ phenotype.

The resistance to these antimicrobials is mediated by two distinct mechanisms encoded by different gene groups. The M phenotype results from *mef* genes encoding for an efflux pump that expels the drug out of the bacterial cell before it can get to its target, described as an antimicrobial target protection mechanism. SDSE isolates can usually present the *mef(*A) or the *mef*(E) gene, which are found in tandem with the *msr*(D) gene [118,119]. The *mef*(A) is commonly present in the phage φ1207.3, mostly found in *S. pyogenes*, although Tn*1207.1* has also been associated with the spread of this gene, and φm46.1 has also been linked to the presence of the mef(A) gene in *S. pyogenes* [120]. In parallel, the *mef*(E) gene is carried by the macrolide efflux genetic assembly (MEGA) structure, that also carries a homolog of the *msr*(D) gene [121–123].

On the other hand, MLS_B phenotypes result from *erm* genes, which encode a methylase that has the capacity of altering the ribosomal site in the 50S subunit to which these groups of antimicrobials attach, referred as a mechanism of antimicrobial target alteration. This way, the binding capacity decreases, not allowing for a proper attachment and action of the drug [124]. From this group of genes, *erm*(TR) (a subclass of *erm*(A)) or *erm*(B) are the most commonly found in SDSE strains, while *erm*(T) is less frequent [124,125].

These genes are believed to be present in the genome of SDSE due to events of HGT from other streptococcal species. Xie *et al.* [126] has identified putative insertion regions in SDSE, some shared with *S. pyogenes*, where mobile elements carrying antimicrobial resistance genes may be located, as demonstrated in **Figure 4**.

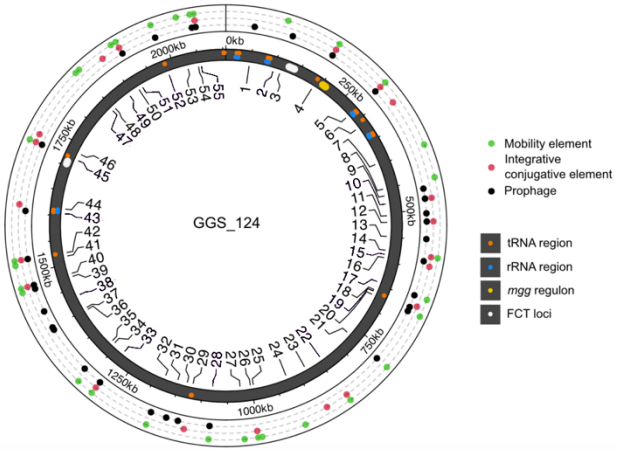


Figure 4 - Location of SDSE MGE insertion regions relative to the GGS_124 reference genome (GenBank sequence number: NC_012891.1). Insertion regions are labeled from 1 through 55 [Adapted from126].

Several MGEs have been found to carry the *erm*(A) gene, in special in *S. pyogenes*, with IME*Sp2907* being the principal carrier integrative and mobilizable element (IME) that can, itself, be inserted in more complex integrative and conjugative elements (ICEs) [116,122,127]. This way, the presence of the *erm*(A) gene may be accompanied by other antimicrobial resistance genes also inserted in the ICE. Examples of that are the ICE*Sp2905*, that carries the *tet*(O) gene, or the ICE*SpHKU165* or ICE*SagTR7*

carrying the *tet*(M) gene, both tetracycline resistance genes [127,128]. Additionally, the *erm*(A)-element has been described in ICE*Sp1108*-like structures, also reported in *S. suis* and *S. agalactiae*, and that commonly carry the *tet*(M) gene [122,127,129].

The transposons of the Tn*916* family seem to be the elements most commonly associated with the carriage of the *erm*(B) gene [121,123]. The most frequent elements of this family are the Tn*6002*, from the insertion of the *erm*(B) into the Tn*916*, and Tn*6003*, with the presence of two copies of the *erm*(B) gene and other antimicrobial resistance determinants [116]. The Tn*3872* element and the ICE*Sp1070HUB* structure, besides harboring the *erm*(B) gene, also include other genes that lead to a multidrug-resistance phenotype [116,122]. Furthermore, the Tn*2010* element is behind dual-macrolide resistance events, as it carries both the *erm*(B) and the *mef*(E) gene [121].

As for the *erm*(T) gene, it is present in non-self-transmissible plasmids that benefit from ICE*Sde3396*-like structures [130]. These ICEs allow for the transfer of the *erm*(T)-carrying plasmid in streptococci, as they present the machinery necessary for conjugation, although the co-transfer of the ICE structure itself is not mandatory [125].

Throughout time, the evolution of the resistance to antimicrobials, in special to macrolides, has been under attention. A review by Berbel *et al*. [116] compiling various studies from different geographical locations indicated that the proportion of macrolide resistant SDSE strains in the USA and Europe has been increasing, reaching values of around 20 to 30% in the last decade. In Asia, the numbers range from around 18 to 24% in Japan and Central Taiwan, but reaching 40 to 72% in Korea, Southern Taiwan, and China. In Africa and the middle east, although studies reporting macrolide resistance are scarce, the numbers range between 8 and 28%. MLS_B phenotypes prevail in SDSE isolates, with *erm*(A) and *erm*(B) being the most commonly found genes encoding macrolide resistance.

Tetracyclines

SDSE isolates have also shown high levels of tetracycline resistance, which acts by blocking protein synthesis as it reversibly bind to the ribosomal 30S subunit, not being the most suitable line of treatment [131,132]. The resistance to this antimicrobial is mediated by three main mechanisms: active efflux pumps, enzyme inactivation, and ribosomal protection, with the latter being the most important within the *Streptoccocus* genus and the *tet*(M) the most frequent gene [133]. Other genes of the *tet* family may also be found in the SDSE genome, such as *tet*(O), *tet*(S) or *tet*(L), although not as commonly [134–136]. Research has shown that the presence of these genes in the genome of SDSE is mainly due to the horizontal transfer of IMEs and other mobile genetic elements (MGE) [137,138], with new elements carrying tetracycline resistance genes recently described [133].

Fluoroquinolones

Fluoroquinolones, such as norfloxacin and levofloxacin, are another class of antimicrobials suitable for treatment of SDSE infections, although some higher rates of resistance have been described in Portugal [139]. These antimicrobials work by inhibiting the bacterial DNA gyrase and DNA topoisomerase IV, which are topoisomerases that work by unraveling supercoils in the DNA and allowing DNA synthesis to occur. The proteins are composed of two subunits, GyrA and GyrB in gyrase, and ParC and ParE in topoisomerase IV [140,141]. Resistance to fluoroquinolones commonly arises due to mutations in the so called quinolone resistance-determining regions (QRDRs) of *gyrA* and *parC* genes, that are thought to occur spontaneously, while mutations in *gyrB* and *parE* genes do not seem to be as important [139,140].

Aminoglycosides

Aminoglycosides, including gentamycin and streptomycin, are yet another commonly used antimicrobial group of which the agents act by binding the 30S ribosomal subunit, although some aminoglycosides seem to bind the 50S subunit [142]. However, similarly to enterococci, streptococci have an intrinsic resistance to low levels of this group of antimicrobials, as their capacity of trespassing the cell wall is reduced. As such, these antibiotics can be used more profitably in combination with cell-wall-active agents, such as β-lactams and glycopeptides [143]. Some high levels of aminoglycoside resistance have also been detected and attributed to MGEs containing genes that encode for aminoglycoside-modification enzymes, which will lead to failure in the inhibition of the 30S ribosomal subunit [143].

2. Aims of the current work

Although disregarded as a pathogen in the past, the rising number of SDSE invasive infections worldwide has enlighten the importance of understanding the mechanism of action and epidemiological traits of this bacteria, as well as the significance in the detection and monitorization of the infections. Not only that, but SDSE antimicrobial resistance must be discussed as to acknowledge the best course of action when facing SDSE invasive infections, and to further understand which mechanisms are behind this trait.

As such, this work aims to:

- Phenotypically and genotypically characterize the SDSE isolates causing invasive infection in Portugal in the period of 2019 to 2022;
- Analyze the epidemiological traits of the isolates;
- Determine the antimicrobial susceptibility of the SDSE isolates to antimicrobials of different classes;
- Describe the macrolide and lincosamide resistance patterns through the determination of the resistance phenotypes and the genes responsible for resistance;
- Compare the antimicrobial resistance rates with results from years prior;
- Investigate the genomic regions in which the macrolide and lincosamide resistance genes are inserted and compare with mobile genetic elements already known to harbor these resistance genes;
- Identify the main macrolide and lincosamide resistant SDSE lineages and investigate the role of clonality in the dissemination of resistance.

The current work aimed at characterizing Lancefield group C and G streptococci (GCGS) causing invasive infections in Portugal during the period of 2019 to 2022 as part of an ongoing surveillance study started in 2011.

Between 2019 and 2022, the invasive isolates identified as *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) were characterized by Lancefield grouping and antimicrobial susceptibility testing, showing high rates of resistance to macrolides and lincosamides. Resistance genes were searched by PCR in isolates resistant to these antimicrobials. DNA extraction, purification and quality control were performed for the the isolates recovered during this period and sent to whole-genome sequencing (WGS), with the purpose of studying the dynamics of macrolide and lincosamide resistance in this population. WGS was performed by an external service and the results were not available by the end of this thesis, preventing further characterization of the isolates recovered between 2019 and 2022.

Therefore, WGS data for invasive SDSE isolates recovered in Portugal between 2011 and 2018, obtained from previous research projects in the host laboratory, was used for further analysis. The

molecular typing information and the presence of mobile genetic elements (MGEs) associated with the spread of macrolide and lincosamide resistance genes were explored. This part of the work provides a framework for further characterization of the SDSE macrolide and lincosamide resistant population once the WGS data for isolates recovered in the 2019 to 2022 period is available.

3. Materials and methods

3.1. Isolate collection

SDSE isolates causing human invasive infections (i.e. isolated from normally sterile sites) were collected in Portugal between January of 2019 and December of 2022. These isolates were collected in 23 laboratories in Portugal (**Supplementary table 1**) that were part of the Portuguese Group for the Study of Streptococcal Infections and submitted to Instituto de Microbiologia at Instituto de Medicina Molecular (Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal). Isolates were identified at the species level by the submitting laboratory. The number of isolates and sample types are identified in **Table 3**. In total, 183 isolates were recovered.

Table 3 - SDSE isolates from human invasive infections in the years of 2019 to 2022 in Portugal (n=183).

Product	Number of isolates	Percentage of isolates
Blood	168	92.8%
Synovial fluid	11	6.1%
Peritoneal fluid	2	1.1%
Bone biopsy	1	0.5%
Pleural fluid	1	0.5%

The collected isolates were preserved at -80°C in tryptone soy broth (TSB) (Oxoid, Basingstoke, UK) supplemented with 15% (v/v) glycerol (Sigma-Aldrich, St. Louis, MO, USA). To perform the characterization of the isolates, these were cultured in tryptone soy agar (TSA) (Oxoid, Basingstoke, UK) supplemented with 5% (v/v) defibrinated sheep blood (Probiológica, Lisboa, Portugal) at 35°C in ambient atmosphere from 24 hours. The colonies were then observed to confirm that characteristics such as color, size and hemolysis were coherent with an SDSE identification.

3.2. Lancefield grouping

Lancefield group identification was performed using the Streptococcal Grouping Kit using Latex Agglutination (Oxoid, Basingstoke, UK) and following the manufacturers use instructions. All strains were tested with group C and G sera.

3.3. Antimicrobial susceptibility testing

Antimicrobial susceptibility testing was performed by disc diffusion using the Kirby-Bauer method according to the most recent European Committee on Antimicrobial Susceptibility testing (EUCAST) or Clinical and Laboratory Standards Institute (CLSI) guidelines (EUCAST, 2023 [144]; CLSI, 2023 [145]), using *Streptococcus pneumoniae* ATCC 49619 as quality control strain as preconized by the guidelines. Incubation was conducted at 35°C in an atmosphere with 5% CO_2 on Mueller-Hinton agar (Oxoid, Basingstoke, UK): to follow EUCAST guidelines, the media was supplemented with 5% defibrinated horse blood (Probiológica, Lisbon, Portugal) and 0.001% hydrated β-NAD at 20 mg/mL; to follow CLSI guidelines, the media was supplemented with 5% defibrinated sheep blood (Probiológica, Lisbon, Portugal). Some colonies were suspended in tubes with 0.85% NaCl to a turbidity of approximately 0.5 on the McFarland scale and inoculated on the agar plates.

For EUCAST media, the disks (Oxoid, Basingstone, UK) used to perform the tests were penicillin (1U), norfloxacin (10 µg), levofloxacin (5 µg), vancomycin (5 µg), linezolid (10 µg), trimethoprim/sulfamethoxazole (25 µg), erythromycin (15 µg), clindamycin (2 µg), tetracycline (30 µg), gentamycin (30 µg) and streptomycin (300 µg); gentamycin and streptomycin were used to detect high-level resistance to aminoglycosides following the EUCAST guideline values for enterococcus, given that no breakpoints are available for streptococci: isolates identified as resistant (R) to these antimicrobials in fact present high-level resistance to this group. In accordance with the EUCAST guidelines, the norfloxacin disk diffusion test can be used to screen for fluoroquinolone resistance. In cases of norfloxacin susceptibility, the isolates can be reported as “susceptible increased exposure” (I) to levofloxacin. In case of norfloxacin resistance, the disk diffusion test should be performed for levofloxacin, as isolates resistant to norfloxacin can be categorized as “resistant” (R) or “susceptible increased exposure” (I) to levofloxacin. Erythromycin and clindamycin disks were placed in proximity to detect MLS_B resistance phenotypes also according with the guidelines to distinguish between the $cMLS_B$ and the $iMLS_B$ phenotype. Isolates expressing the $iMLS_B$ phenotype, although being susceptible to clindamycin when this antimicrobial is tested alone, should be considered clinically resistant to lincosamides according to EUCAST criteria and should not be used in treatment.

Since there are no EUCAST guidelines to test the susceptibility to chloramphenicol in group C and G streptococci, testing was performed in CLSI media using chloramphenicol (30 µg) disks (Oxoid, Basingstone, UK). To test for reduced susceptibility to β-lactams, oxacillin (1 µg), ceftizoxime (30 µg) and ceftibuten (30 µg) disks (Oxoid, Basingstone, UK) were used according to the screening protocol described by Kimura *et al.* [146]. Briefly, since CLSI guidelines do not determine the cutoff values for susceptibility to these antimicrobials, provisional “reduced susceptibility” values were set for the three antibiotics using as a guide the values of susceptibility of the strains with the smallest-diameter values: in this case, <17 mm for oxacillin, <29 mm for ceftizoxime, and <20 mm for ceftibuten. This protocol was described for Lancefield group B streptococci but here applied for SDSE.

3.4. DNA extraction for PCR

DNA extraction was performed for macrolide and lincosamide resistant isolates in order to identify the macrolide and lincosamide resistance genes present in their genomes, while using Chelex® 100 Molecular Biology Grade Resin (Bio-Rad Laboratories, Hercules, CA, USA) following the manufacturer's instructions. In sum, a solution was prepared using Chelex® 100 Molecular Biology Grade Resin at 500 mg/mL and 0.2 mg/mL of proteinase K, and then distributed in Eppendorfs with 202 µL of solution. With a sterile loop, colonies were suspended in each Eppendorf and incubated at 56ºC for one hour. The solution was then resuspended, incubated for 10 minutes at 90ºC and again vortexed. Lastly, it was centrifuged at 12000 rpm for 3 minutes and preserved at -20°C until needed for PCR execution.

3.5. PCR for macrolide resistance gene identification

Isolates presenting resistance to erythromycin underwent PCR.

In short, the PCR multiplex mix for identification of *erm*(A) (sub-class *erm*(TR)), *erm*(B), and *mef*(A) genes consisted of 1x Green GoTaq Flexi Buffer (Promega, Madison, WI, USA), 2.5 mM $MgCl_2$ (Promega, Madison, WI, USA), 0.2 mM dNTP Mix (Quantabio, Beverly, MA, USA), 0.2 pmol/µL of each primer [147] – erm(TR)1 - CTTGTGGAAATGAGTCAACGG, erm(TR)2 - TTGTTCATTGGATAATTTATC, 0.4 pmol/µL; erm(B)2.3 - AGTAAYGGTACTTAAATTGTTTAC, erm(B)1.2 - GAAAARGTACTCAACCAAATA, 0.25 pmol/µL; mef(A)/E1 - GGCAAGCAGTATCATTAATCAC, mef(A)/E2 - GACTGCAAAGACTGACTATAG, 0.05 pmol/µL (Invitrogen, Carlsbad, CA, USA), and 0.02 U/µL G2 Flex DNA polymerase (Promega, Madison, WI, USA). In each PCR reaction tube, 1 µL of DNA was added to 24 µL of the PCR mix. The mix for identification [adapted from [148]] of the *ermT* gene consisted of 1x Green GoTaq Flexi Buffer (Promega, Madison, WI, USA), 2 mM $MgCl_2$ (Promega, Madison, WI, USA), 0.2 mM dNTPs Mix (Quantabio, Beverly, MA, USA), 0.2 pmol/µL of each primer [125] – ermT(up), 5′-AGGATCTA TTTCAATGGCGG-3′, ermT(down), 5′-TCCAAAAACCAATTTTATCAAGC-3′ (Invitrogen, Carlsbad, CA, USA), and 0.01 U/µL GoTaq G2 Flex DNA polymerase (Promega, Madison, WI, USA). In each PCR reaction tube, 0.4 µL of DNA were added to 19.6 µL of the PCR mix. DNA extracted from SDSE isolates in the laboratory's collection for which whole-genome sequence was available were used as positive controls: *erm*(A) – isolate SH16980; *erm*(B) – isolate SH16918; *mef*(A) – isolate SH18147; *erm*(T) – isolate SH18902. The negative control used was distilled water.

The PCR program for both PCRs included a 5-minute initial phase at 94°C, 40 cycles of 30 seconds at 94°C, 30 seconds at 51°C and 1 minute at 72°C, followed by pause stage at 16°C.

After that, a gel electrophoresis was performed using a 1% (w/v) agarose gel prepared with electrophoresis grade agarose (Bio-Rad, Hercules, California, USA) in 0.5X TBE buffer (Bio-Rad, Hercules,

California, USA). The ladder 1KB plus (Invitrogen, Carlsbad, California, USA) was used as a molecular weight marker and the separation was performed at 120V for about 60 minutes.

3.6. High throughput sequencing (HTS)

3.6.1. Genomic DNA extraction

All SDSE invasive isolates were characterized by HTS. Genome extraction for HTS was done using the Invitrogen PureLink® Genomic DNA extraction kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). Slight modifications were applied to the manufacturer's instructions. A bacterial suspension was obtained by inoculation of one colony in 5 ml of Todd-Hewitt broth (Becton Dickinson, New Jersey, USA) and incubated overnight at 35°C. A volume of 1 ml of the bacterial suspension was centrifuged at 15060 rpm for 5 minutes and the supernatant was discarded. The lysozyme digestion buffer was prepared, with a ratio of 20 mg of fresh lysozyme for 1000 µL of buffer, with the addition of 75 U of mutanolysin (Sigma-Aldrich, St. Louis, Missouri, USA) to the mixture. After adding proteinase K, 400 µg of RNase was added and incubated for 1 minute at room temperature. To assess DNA purity, a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) was used and the 260/280 nm and 260/230 nm absorbance ratios were measured. DNA integrity was assessed by gel electrophoresis using a 1% (w/v) agarose gel prepared with electrophoresis grade agarose (Bio-Rad, Hercules, California, USA) in 0.5X TBE buffer (Bio-Rad, Hercules, California, USA). The ladder 1KB plus (Invitrogen, Carlsbad, California, USA) was used as a molecular weight marker.

The nucleic acid concentration was measured by fluorometry with an Invitrogen QubitTM dsDNA HS assay kit (Thermo Fisher Scientific Inc., Waltham, MA, USA) following the manufacturer's instructions. Lastly, the DNA concentration was adjusted to 10 ng/µL by dilution in Tris-HCL 10 mM pH8 (Sigma-Aldrich, St. Louis, Missouri, EUA).

3.6.2. Sequencing, assembly and annotation of the genome

The complete genome sequencing was performed at the Gene Express Unit at Instituto Gulbenkian de Ciência (Oeiras, Portugal). High throughput libraries were prepared with the Nextera XT DNA Library Prep paired-end kit and the Index v2 kit (Illumina©, San Diego, California, USA) and sequenced using the Illumina NextSeq 2000 (Illumina©) system that uses the NextSeq 500/550 Mid-Output kit (version 2) (300 cycles). The genome assembly was performed by the bioinformatics team at Instituto de Microbiologia of Faculdade de Medicina de Lisboa. The evaluation and control of the quality of the paired-end reads, as well as the *de novo* assembly, were obtained using the INNUca pipeline v3.1 (https://github.com/B-UMMI/INNUca). This pipeline was run using the Docker image "ummidock/innuca:3.1" (https://hub.docker.com/r/ummidock/innuca/), using sequences of adapters Nextera XT for removal of adapters, and an expected genome size of 2.1 Mb. The quality of the reads was evaluated using the FastQC

tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and then processed with Trimmomatic [149]. The *de novo* assembly of the reads processed in draft genomes was obtained with SPAdes [150] and polished with Pilon[151]. The polished draft genomes were annotated using the Prokka v1.12 pipeline [152] using the Docker image "ummidock/prokka:1.12" (https://hub.docker.com/r/ummidock/prokka/). The tools SignalP v4.1 [153], used to identify the signal peptides in the coding sequences (CDS), and RNAmmer v1.2 [154], used to identify ribosomal RNA, were externally provided to the Docker image. Prokka was run using a genus database created using complete SDSE genomes and regarding the following parameters: --*addgenes* --*usegenus* --*rfam* --*rnammer* --*gram pos* --*increment* 10 --*mincontiglen* 1 -- *gcode* 11 -- *kingdom Bacteria* --genus *Streptococcus* --*species dysgalactiae*.

3.6.3. Identification of antimicrobial resistance genes and extraction of typing data from SDSE genomes

The genetic relationship between the SDSE invasive isolates was determined based on the available MLST scheme for *S. dysgalactiae* (https://pubmlst.org/organisms/streptococcus-dysgalactiae/). To obtain the allelic profile (*gki, gtr, murI, mutS, recP, xpt* and *atoB* genes) and the corresponding ST from the draft genome of each isolate, the MLST (https://github.com/tseemann/mlst) present in the INNUca pipeline was used. The new alleles or profiles identified were submitted to the PubMLST database for attribution of a number. To define the CCs, the goeBURST algorithm implemented in PHYLOViZ was used, including all allelic profiles of the *S. dysgalactiae* strains available in the database (https://pubmlst.org/organisms/streptococcus-dysgalactiae/) in the analysis. The CCs were defined at SLV level. The Comprehensive Antibiotic Resistance Database (CARD) (https://card.mcmaster.ca/) for identification of the resistance determinants present in the genomes. The *emm*-type was determined using e*mm*Typer v0.2.0 (https://github.com/MDU-PHL/emmtyper/releases/tag/v0.2.0).

3.7. Characterization of MGEs carrying macrolide and lincosamide resistance determinants

To understand the distribution of macrolide resistance genes, FASTA files containing the sequenced genome of the macrolide resistant isolates causing invasive disease in Portugal in the period of 2011 to 2018 were uploaded to Geneious Prime® 2024.0.2. Alongside it, representative sequences of macrolide resistance genes *erm*(A) (GenBank accession number: EU348756), *erm*(B) (GenBank accession number: AP026914), *erm*(T) (GenBank accession number: M64090), different *mef(*A) subclasses: *mef*(A): GenBank accession number SPU70055 [155]; *mef*(B): GenBank accession number DQ445269 [156]; *mef*(E): GenBank accession number: SPU83667 [157]; *mef(*G): GenBank accession number: AM168138 [158]; *mef*(I): GenBank accession number: AJ971089 [159]; *mef*(O): GenBank accession number: DQ016305 [160]), and

lnu(D) (Genbank accession number: ABR14060 [161]) were retrieved from CARD (https://card.mcmaster.ca) and uploaded to Geneious Prime® 2024.0.2. Databases including all isolates presenting each of the resistance genes were created in accordance with the results from the prior genomic analysis. The basic local alignment search tool (BLAST) was used to detect each of the resistance genes in the genomes and to identify the contig in which the gene was detected. Each contig harboring macrolide resistance genes was compared by BLAST to the GGS_124 reference genome (GenBank reference sequence: NC_012891 [51]) to identify genomic conserved regions, SDSE insertion regions as defined by Xie *et al.* [126], and to establish the boundaries of putative mobile genetic elements. The full sequence of the identified putative MGEs was extracted from each contig. The sequences of previously described MGEs containing the macrolide resistance genes in streptococci and known to have been previously described in SDSE, including *erm*(A)-IMESp2907 (GenBank accession number: FR691055) [122], ICE*Sp2905* (GenBank accession number: FR691055) [127]; *erm*(B)-Tn*3872* (GenBank accession number: AM411624), Tn*6002* (GenBank accession number: AY898750), Tn*6003* (GenBank accession number: AM410044) [116]; *mef*(A)-φ1207.3 [122] (GenBank accession number: AY657002), were used for comparison purposes. A BLAST was performed to search for the presence of these elements against each contig containing the respective macrolide resistance gene. The identification of putative prophages, IMEs and insertion sites was done using PHASTEST and ICEfinder 1.0 [162]. Easyfig 2.2.2 [163] was used to align and compared the contigs and the elements present in the contigs. Nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to search for related elements in the NCBI databases.

3.8. Statistical significance of the data

The determine associations between variables, Fisher's exact test was used. The statistical significance of the obtained results was calculated using the Cochran-Armitage test for trend with R Studio, with the aim of understanding if there were significant changes over time in the percentage of antimicrobial resistance. The p-values for both tests were corrected using the false discovery rate (FDR) in RStudio®.The results with p-values lower than 0.05 were deemed statistically significant. The Simpson's Index of Diversity (SID) and corresponding 95% confidence intervals (95%CI) were calculated for characterizing population diversity as inferred by applying a given typing method [164] . It represents the probability of any two strains randomly selected from a population belonging to two different groups, and it varies from 0 to 1, with population diversity increasing as SID approaches 1. The online tool available at http://www.comparingpartitions.info/ was used.

4. Results

4.1. Characterization of SDSE invasive isolates, 2019-2022

4.1.1. Patient demographics

The total number of SDSE invasive isolates received in 2019-2022 was 183. **Figure 5** shows the distribution of SDSE invasive isolates from the years 2019 to 2022. Overall, despite some fluctuations between years, the number remained around 40 to 50 cases per year and there seemed to be no substantial tendency for the number of invasive infections by SDSE to increase between 2019 and 2022.

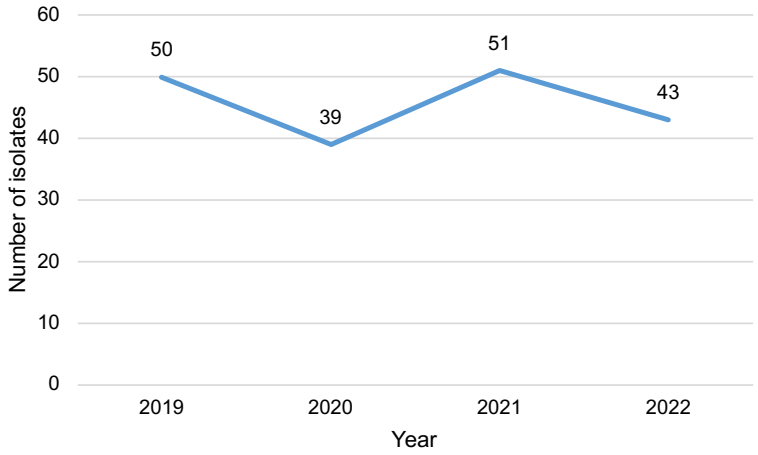


Figure 5 - Distribution per year of SDSE isolates causing invasive infections in Portugal in 2019-2022 (n=183).

Regarding the distribution of the patients with SDSE invasive infections in Portugal in 2019-2022 by sex, around 43% (n=78) of patients were female, whilst the remaining 57% (n=105) were male. Regarding the patients' age, no infections were observed in the pediatric patients (<18 years old), around 15% (n=27) of infected individuals were 18 to 49 years old, around 19% (n=35) were 50 to 64 years old, whilst the remaining 68% (n=121) were of age 65 or older.

4.1.2. Characterization of the isolates

After 24h incubation, all isolates presented the typical morphology of SDSE, with large white/greyish colonies (>0.5 mm) and clear halos in the media surrounding them, demonstrating β-hemolysis. Lancefield group identification characterized 32% (n=59) of the isolates as Lancefield group C isolates, whilst the remaining 68% (n=124) as Lancefield group G isolates. No Lancefield group A or L SDSE isolates were detected during the study period.

4.1.3. Antimicrobial resistance

All isolates were susceptible to penicillin, vancomycin, linezolid, trimethoprim/sulfamethoxazole and chloramphenicol. The percentage of susceptibility/nonsusceptibility for all other antimicrobials is represented in **Figure 6**.

The highest resistance rate was observed towards erythromycin (45.4%, n=83), followed by clindamycin (43.7%, n=80). The proportion of clindamycin resistant isolates include all isolates considered clinically resistant do this antimicrobial, englobing both $cMLS_B$ and $iMLS_B$ phenotypes. Resistance to tetracycline was 27.9% (n=51), while for other antimicrobial agents, resistance rates were below 5%. Only four isolates were resistant to norfloxacin (2.2%), two of which were also resistant to levofloxacin (indicated as R in Figure 6). All the other isolates were classified as "susceptible increased exposure" (I) to levofloxacin according to EUCAST guidelines. High-level resistance to aminoglycosides was also low, with rates of 3.3.% (n=6) to streptomycin and 0.5% (n=1) to gentamycin.

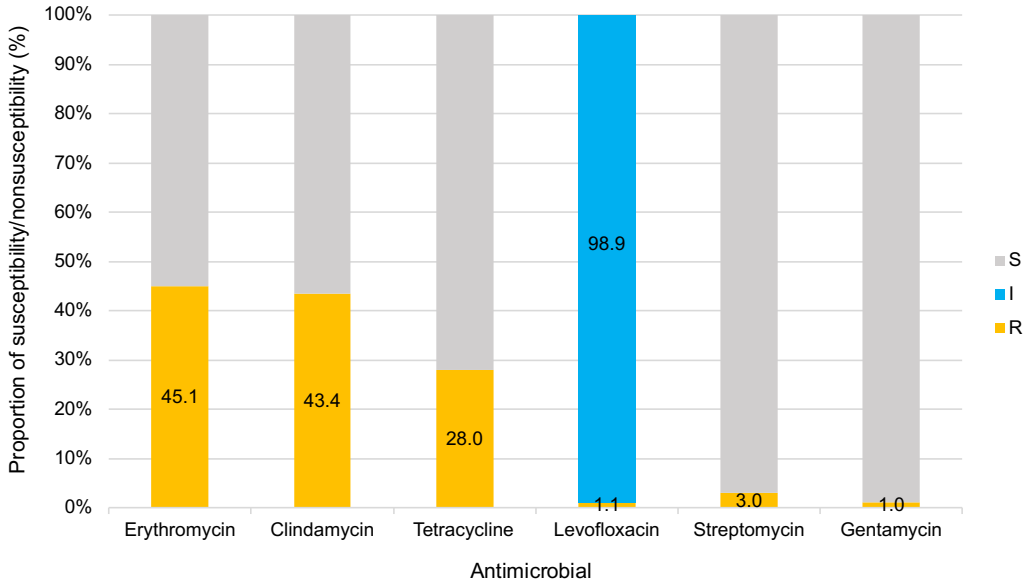


Figure 6 - Antimicrobial resistance rates of invasive SDSE isolates recovered in Portugal from 2019 to 2022.
Abbreviations: R, resistant; I, "susceptible increased exposure"; S, susceptible.

When screening isolates for reduced susceptibility to β-lactams, a single isolate presented reduced susceptibility to both ceftizoxime and ceftibuten, as the measured inhibition halos to these antimicrobials were reduced by 1 mm and equivalent to breakpoints, respectively.

Overall, 111 (69.7%) isolates were resistant to at least one antimicrobial agent and 22 (12.0%) were resistant to both macrolides and tetracycline.

The temporal evolution of resistance rates to the different antimicrobial agents tested is represented in **Figure 7**. After performing the Cochran–Armitage test for trend, no antimicrobial resistance fluctuation was deemed statistically significant for this time period. Erythromycin/clindamycin presented the highest proportion of resistant strains, followed by tetracycline, with values above 50% and 20% by the end of the study, respectively.

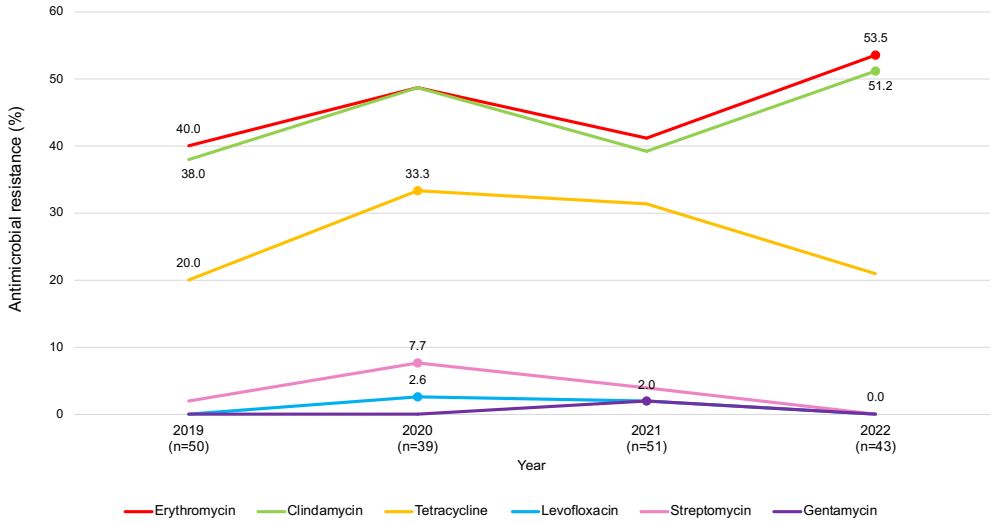


Figure 7 - Distribution of the percentage of resistance to antimicrobials from 2019 to 2022 (n=183), with identification of the years of highest and lowest resistance to each antimicrobial.

4.1.4. Macrolide resistance phenotypes

Overall, 60 and 20 isolates presented the iMLS$_B$ and cMLS$_B$ phenotypes, respectively. The M phenotype was expressed by a small number of isolates (n=3). The evolution of the macrolide resistance phenotypes in the years of 2019 to 2022 is represented in **Figure 8**. Over the study period, no differences in phenotype prevalence were detected, with the iMLS$_B$ phenotype prevailing in all years of the study.

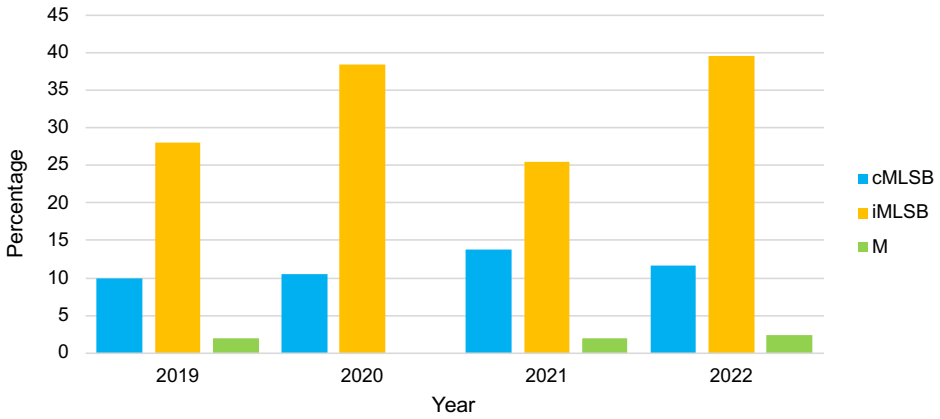


Figure 8 - Distribution of the percentage of macrolide resistance phenotypes of the invasive SDSE isolates recovered in Portugal from 2019 to 2022 (n=183).

4.1.5. Macrolide resistance genotypes

For all the isolates expressing macrolide and lincosamide resistance (n=83), PCR was performed in order to determine the gene encoding for the resistance. The results are presented in **Table 4**.

Table 4 - Distribution of the macrolide resistance genes of the invasive SDSE isolates recovered in Portugal from 2019 to 2022 that express a macrolide resistance phenotype (n=83).

	***erm* (A)**	***erm* (B)**	***erm* (T)**	***mef* (A)**	Total
$cMLS_B$	4 (4.8%)	16 (19.3%)			20 (24.1%)
$iMLS_B$	57 (68.7%)		3 (3.6%)		59 (72.3%)
M				3 (3.6%)	3 (3.6%)
Total	61 (73.5%)	16 (19.3%)	3 (3.6%)	3 (3.6%)	83 (100%)

Overall, 73.5% (n=61) of isolates with resistance to macrolides presented the *erm*(A) gene, of which 93.4% (n=57) expressed the $iMLS_B$ phenotype, while the remaining 6.6% (n=4) exhibited the $cMLS_B$ phenotype. The *erm*(B*)* gene was present in 19.3% (n=16) of the isolates, and only $cMLS_B$ expressing isolates possessed this gene. Additionally, three isolates exhibiting the $iMLS_B$ phenotype were shown to carry the *erm*(T) gene. As expected, the remaining isolates expressing the M phenotype possessed the *mef*(A) gene.

The genomic sequencing results for SDSE isolates causing invasive infections in Portugal from 2019 to 2022 were not available in time for assessment of the genotypic characteristics of the SDSE macrolide and resistance population during this period. Therefore, WGS data from SDSE invasive isolates recovered in prior years (2011 to 2018), partly presented in prior work by Castro [45] and additional data available in the host laboratory, were used to further explore the dynamics of macrolide and lincosamide resistance in SDSE invasive isolates recovered in Portugal and the origin of the resistance determinants.

4.2. Characterization of SDSE invasive isolates, 2011-2018

The total number of SDSE invasive isolates received in 2011-2022 was 562 (n=379 in 2011-2018 and n=183 in 2019-2022). **Figure 10** shows the distribution of SDSE invasive isolates from the years 2011 to 2022. Some oscillations in the number of SDSE invasive isolates collected are visible, with a peak in 2018 followed by a slight drop afterwards. Overall, no tendency for the number of invasive infections by SDSE to increase from between 2011 and 2022 could be detected.

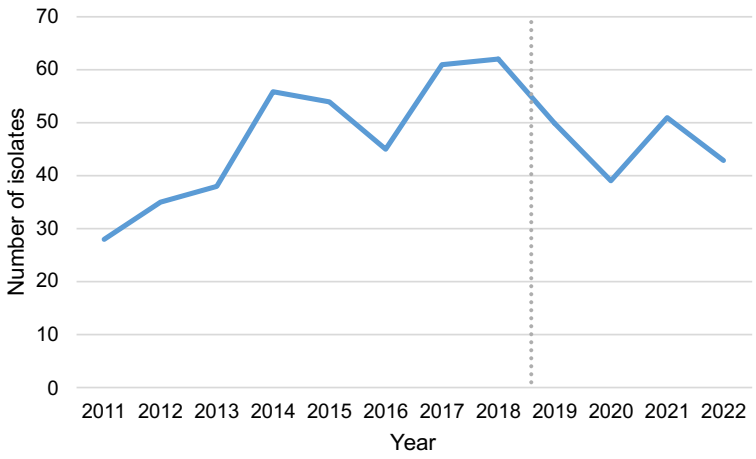


Figure 9 - Distribution per year of SDSE isolates causing invasive infections in Portugal in 2011-2022 (n=562).

To further understand how the antimicrobial resistance rates are progressing, the whole results from 2011 to 2022 were evaluated and are represented in **Figure 10**. After performing the Cochran–Armitage test for trend, it is possible to conclude that there was a tendency for erythromycin and clindamycin resistance to rise in the years 2011 to 2022 (p-value < 0.05, still significant after FDR correction). Additionally, resistance to levofloxacin decreased in the years of 2011 to 2022 (p-value < 0.05, still significant after FDR correction). Tetracycline is also the antimicrobial presenting the third highest proportion of resistant strains. No other antimicrobial resistance suffered significant differences in the years of 2011 to 2022.

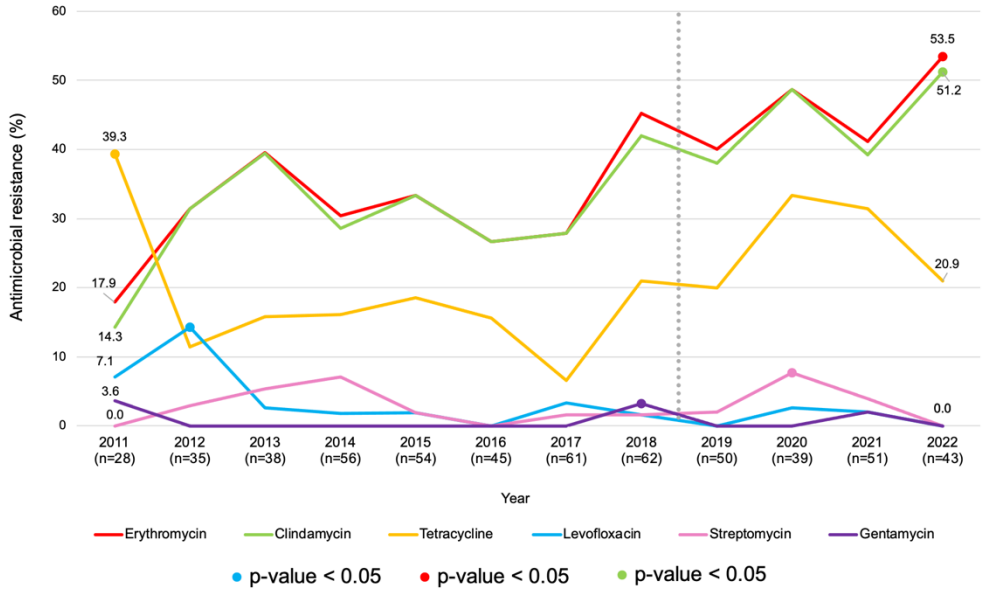


Figure 10 - Distribution of the percentage of resistance to antimicrobials from 2011 to 2022 (n=562), with identification of the years of highest resistance to each antimicrobial. Only the number of isolates resistant to erythromycin, clindamycin, and levofloxacin showed a tendency to increase and decrease, respectively, in the years of 2011 to 2022 (p-value < 0.05).

4.2.1. Macrolide and lincosamide resistance

Between years 2011 to 2018, 32.7% (n=124) of the 379 SDSE isolates causing invasive disease in Portugal presented macrolide resistance. From these, 73.4% (n=91) presented the iMLS$_B$ phenotype, 22.6% (n=28) presented the cMLS$_B$ phenotype, while the remaining 4.0% (n=5) presented the M phenotype. As for the distribution by macrolide resistance genes, most isolates (80.7%, n=101) presented the *erm*(A) gene, followed by the *erm*(B) (11.3%, n=15), *mef*(A) and *erm*(T) genes (2.4%, n=3 each). Despite an increase in the overall macrolide and lincosamide resistance rate when comparing the two periods (32.7% erythromycin resistance in 2011-2018 to 45.4% in 2019-2022, [p-value<0.05, still significant after FDR correction]), no differences were observed in terms of MLS resistance phenotypes or genotypes when comparing both periods (**Table 5**).

Table 5 – Evolution of the distribution of the macrolide resistance genes of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 and from 2019 to 2022 that express a macrolide resistance phenotype.

	2011-2018						2019-2022				
	***erm* (A)**	***erm* (B)**	***erm* (T)**	***mef* (A)**	**None**	Total	***erm* (A)**	***erm* (B)**	***erm* (T)**	***mef* (A)**	Total
cMLS$_B$	15 (12.1%)	13 (10.5%)[a]		1 (0.8%)[a]		28 (22.6%)	4 (4.8%)	16 (19.3%)			20 (24.1%)
iMLS$_B$	86 (69.4%)	2 (1.6%)	3 (2.4%)			91 (76.4%)	57 (68.7%)		3 (3.6%)		59 (72.3%)
M				3 (2.4%)	2 (1.6%)	5 (4.0%)				3 (3.6%)	3 (3.6%)
Total	101 (81.5%)	15 (12.1%)[a]	3 (2.4%)	4 (3.2%)[a]	2 (1.6%)	124 (100%)	61 (73.5%)	16 (19.3%)	3 (3.6%)	3 (3.6%)	83 (100%)

[a] This value includes the isolate presenting both the *erm*(B) and *mef*(A) resistance genes.

Analysis of the genomic sequences of resistant isolates further allowed to detect the *lsa*(C) gene, which confers resistance to lincosamides and streptogramins A [165] in one of the *erm*(A) carrying isolates. Another isolate presented two copies of the *mef*(A)/*msr*(D) genes, together with a copy of the *erm*(B) gene, and additionally, a copy of the *lnu*(D) gene, which confers resistance to the lincosamide lincomycin [161]. Furthermore, WGS confirmed that *mef(*A) was found in tandem with the *msr*(D) gene in all isolates carrying these genes. Finally, two isolates expressing the M phenotype, which after testing for the minimum inhibitory concentration were confirmed to be resistant to erythromycin, did not seem to harbor any macrolide resistance genes, as none could be detected in the genomic sequence. **Figure 11** resumes the findings of resistance genotypes and phenotypes based on the genomic data.

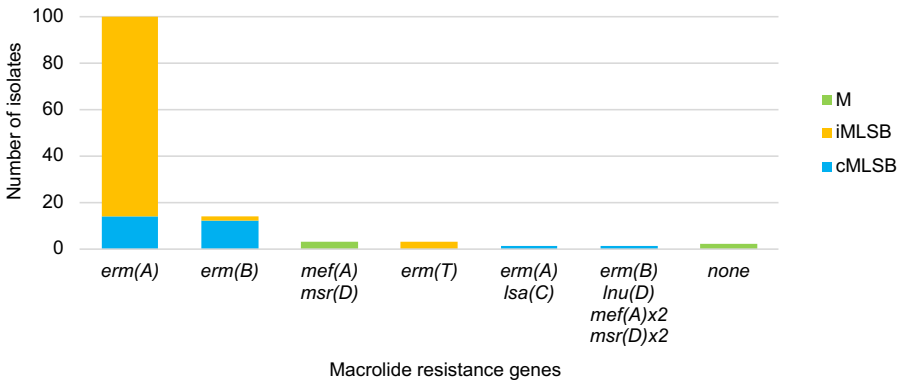


Figure 11 - Macrolide resistance gene distribution and macrolide susceptibility/resistance phenotype proportion of the invasive SDSE isolates recovered in Portugal from 2011 to 2018.

4.2.2. Molecular typing

4.2.2.1. *emm*-type distribution

All 379 isolates collected were separated into 28 distinct *emm*-types and two non-typeable isolates (SID=0.893; 95%CI: 0.876-0.910). *emm*-type *stG62647* (24.5%, n=93) clearly stand out as nearly one quarter of the isolates belonged to this *emm*-type, followed by *stC74a* (10.8%, n=41) and *stG6* (9.5%, n=36). The diversity of *emm* types observed among the 124 isolates with resistance to macrolide and lincosamides (SID=0.886; 95%CI: (0.854-0.918)), which belonged to 21 *emm* types, was similar to the one of the entire SDSE invasive population.

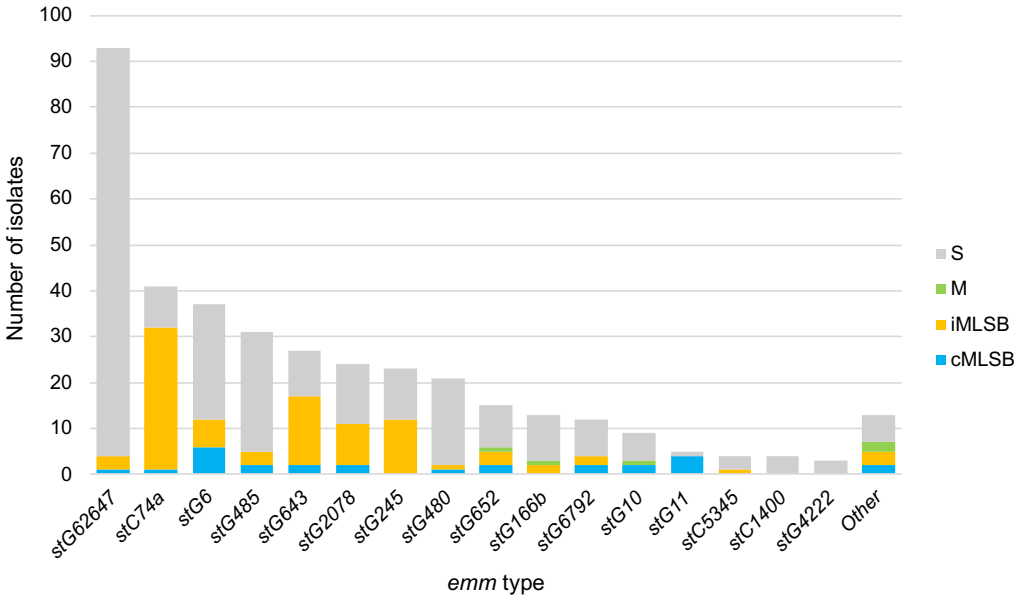


Figure 12 - Distribution of the *emm*-types and macrolide susceptibility/resistance phenotype proportion of the invasive SDSE isolates recovered in Portugal from 2011 to 2018. "Other" includes *emm*-types with only one or two isolates.

In **Figure 12**, it is possible to observe the distribution of the $cMLS_B$ (blue), $iMLS_B$ (orange), and M (green) macrolide resistance phenotypes, as well as the proportion of macrolide susceptible (grey) isolates in each *emm*-type. From the *emm*-types shown, at the exception of two (*stC1400* and *stG4222*), each distinct *emm*-type group presented diversity in their macrolide susceptibility/resistance phenotype.

The Fisher's exact test was used to evaluate the relationship between the resistance to erythromycin and clindamycin and the different *emm*-types. *emm*-type *stC74a* was associated with resistance to both erythromycin and clindamycin (p-value <0.001, still significant after FDR correction), as well as *stG643* (p-value=0.001 for erythromycin and p-value <0.001 for clindamycin, still significant after FDR correction).

4.2.2.2. MLST CC distribution

All 379 isolates collected were separated into 81 distinct STs that were grouped into 11 CCs and 10 singletons (SID=0.786; 95%CI: 0.762-0.811). Around 74.4% (n=282) of the isolates were included in the three most common CCs, CC17 (35.4%, n=134), CC20 (22.7%, n=85), and CC29 (16.6%, n=63). While most CCs included isolates of distinct *emm*-types, as was the case of the most frequent CC, CC17, CC20 stands out as it included almost exclusively isolates of the most common *emm*-type, *stG62647*. **Table 6** also shows the distribution of MLS resistant isolates by MLST CCs and the *emm*-type distribution. The 124 macrolide and lincosamide isolates were distributed over 12 CCs (SID=0.619; 95%CI: (0.526-0.711)), showing slight less diversity than the entire SDSE population.

Table 6 - Distribution of the *emm*-types by the CCs defined by MLST for all invasive SDSE isolates recovered in Portugal from 2011 to 2018. "Other" includes *emm*-types and CCs with only one or two isolates. The results shown differentiate between the number of MLS resistant/susceptible isolates (124 Resistant/255 Susceptible).

		Clonal Complex (CC)										
		CC17	CC20	CC29	CC15	CC269	CC8	CC128	CC34	CC44	Other	Total
emm type	***stG62647***	3/8	1/81									4/89 (93)
	stC74a	26/0		6/8							0/1	32/9 (41)
	stG6	1/1	0/1	5/22	4/0						2/1	12/25 (37)
	stG485	2/9		3/10				0/6			0/1	5/26 (31)
	stG643	15/7					2/3					17/10 (27)
	stG2078	11/12				0/1						11/13 (24)
	stG245	12/5		0/2	0/3	0/1						12/11 (23)
	stG480	2/12			0/1		0/5			0/1		2/19 (21)
	stG652			0/1	2/7	2/1	1/0				1/0	6/9 (15)
	stG166b	0/3		0/2	2/2	0/3					1/0	3/10 (13)
	stG6792					4/7					0/1	4/8 (12)
	stG10	1/0			2/6							3/6 (9)
	stG11						4/1					4/1 (5)
	stc5345	1/1			0/1						0/1	1/3 (4)
	stC1400				0/2				0/1	0/1		0/4 (4)
	stG4222	0/1		0/2								0/3 (3)
	Other	0/1	0/2	1/1	1/0	1/1			3/0	0/1	2/3	8/9 (17)
	Total	74/60 (134)	1/84 (85)	15/48 (63)	11/22 (33)	7/14 (21)	7/9 (16)	0/6 (6)	3/1 (4)	0/3 (3)	6/8 (14)	379

Figure 13 further highlights the proportion of macrolide and lincosamide resistance and susceptibility in each of the CCs. Most CCs present both susceptible and resistant isolates.

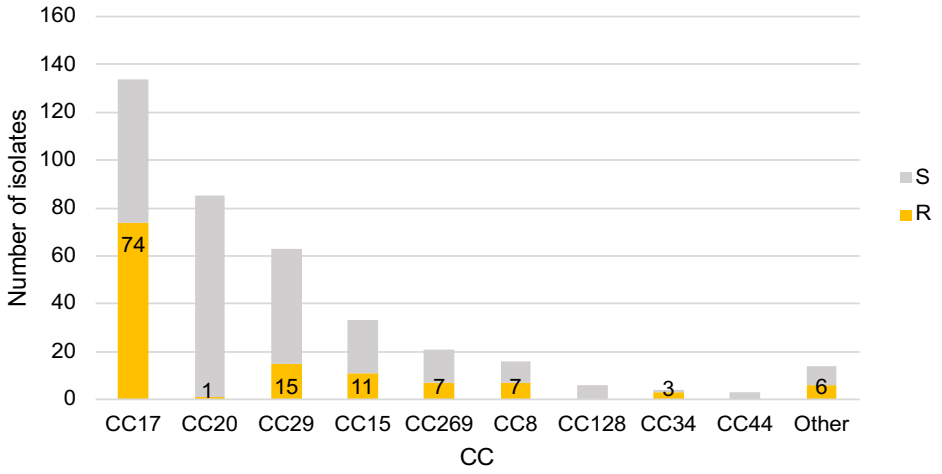


Figure 13 - Distribution of the MLSB resistance/susceptibility by the CCs defined by MLST for all invasive SDSE isolates recovered in Portugal from 2011 to 2018. "Other" includes CCs with only one or two isolates. **Abbreviations**: R, resistant; S, susceptible.

The Fisher's exact test was used to evaluate the association between the resistance to erythromycin and clindamycin and the different CCs. CC17, including almost 60% (74/124) of all resistant isolates, was the only CC associated with resistance to both erythromycin and clindamycin (p-value<0.001, also significant after FDR correction). On the other hand, the second most common CC, CC20, included a single resistant isolate.

In **Figure 14**, it is possible to observe the distribution of the *erm*(A) (green), *erm*(B) (orange), *erm*(B)+*mef*(A) (red), *erm*(T) (purple) and *mef*(A) (blue) macrolide and lincosamide resistance genes for each CC that include resistant isolates. Most *erm*(A) isolates were included in CC17 although other CCs also included a noticeable proportion of these isolates. CC8 included the highest proportion of *erm*(B) carrying isolates, but isolates with this gene were found in several other CCs, most also including *erm(*A) isolates. *erm*(T) was exclusively present in CC15 isolates and *mef*(A) in isolates of three distinct CCs.

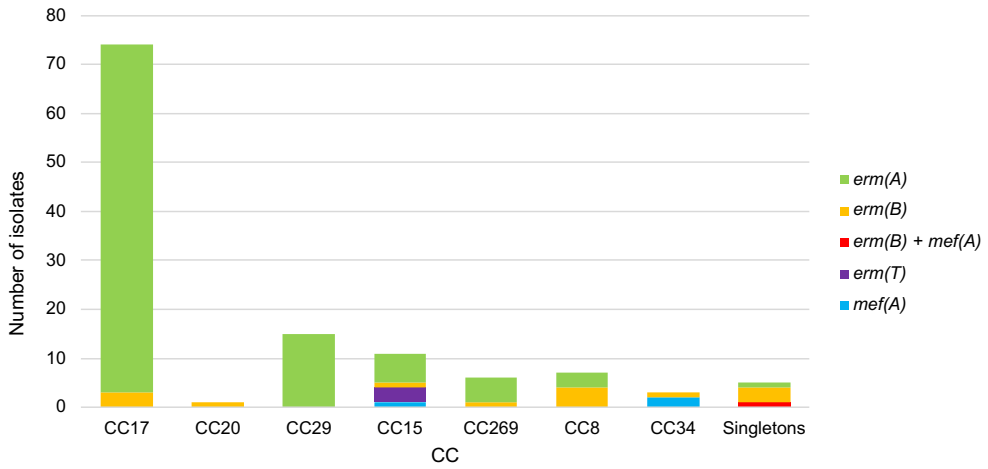


Figure 14 - Clonal complex (CC) distribution and MLS_B resistance genes proportion for the invasive SDSE isolates recovered in Portugal from 2011 to 2018.

4.2.3. Characterization of MGEs carrying MLS resistance genes

4.2.3.1 *erm*(A) gene

The *erm*(A) gene was found in 101 isolates recovered from 18 different Portuguese hospitals. The *erm*(A) isolates belonged to six different CCs (SID=0.482; 95%CI: (0.371-0.592)) and 16 distinct *emm*-types (SID=0.845; 95%CI: (0.800-0.890)). The majority (70.3%) of *erm*(A) carrying isolates belonged to CC17, although this CC included 10 distinct *emm*-types.

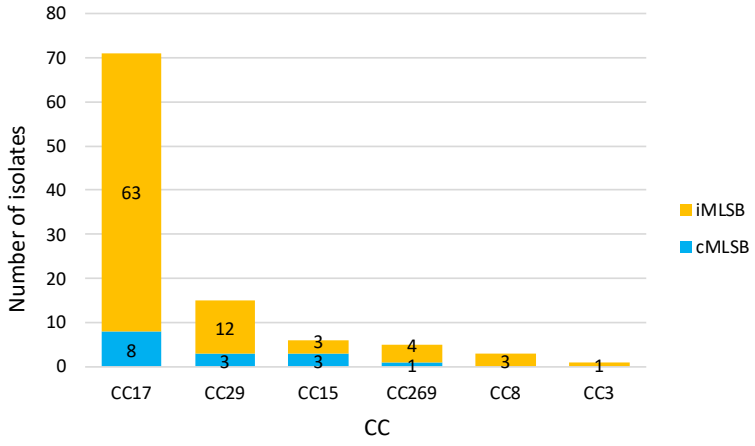


Figure 15 - Distribution of the CC of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 harboring the *erm*(A) gene.

The *erm*(A) sequences present at the 101 isolates shared >99.5% sequence similarity. BLASTn searches of *erm*(A) containing contigs against the GGS_124 genome (GenBank accession number: NC_012891) were used to identify regions conserved in GGS_124 and to define the boundaries of putative MGEs in each contig. This approach allowed to identify the putative elements to their full extension, ranging from 56.4Kb to 119.4Kb, in only four isolates (**Figure 16**), but showed that most *erm*(A) containing elements were integrated in the bacterial chromosome near the *rlmD* gene (GGS_124 locus tag SDEG_RS06630, encoding a 23S rRNA-methyltransferase). The *rlmD* gene corresponds to the 3' extremity of SDSE insertion region 33 recently identified by Xie *et al.*[92] (**Figure 16 (A).1** and **(A).2**). Thus, 85 isolates had *erm*(A) containing elements present in insertion region 33, two isolates in insertion region 5 and two other in insertion region 42 (**Figure 16**). For 12 isolates the insertion region could not be determined because no BLAST hit with the GGS_124 genome was obtained, suggesting that all the contig corresponded to only a portion of the element carrying *erm*(A).

Figure 16 (B) shows the structure of the most commonly detected *erm*(A) element. This 75Kb element started with a C-5 cytosine-specific DNA methylase CDS at the 5' start of the contig and finished with a hypothetical protein CDS immediately upstream of the *rlmD* gene. As in the majority of the isolates, the 5' end of the element could not be determined with certainty because, in most cases, it corresponded to the beginning of the contig.

Alignment between the most frequent *erm*(A) containing element (**Figure 16 (B)**) with the full 65.4 Kb ICE*Sp2905* element (GenBank accession number: FR691055) from *S. pyogenes* and previously detected in SDSE, showed only 67.3% sequence identity. On the other hand, an alignment with the 12.6Kb region common to both ICE*Sp2905* and ICE*Sp2907* showed 90.4% sequence identity. From all isolates containing the *erm*(A) gene, 59 (58.4%) presented an element similar to the one presented in **Figure 16 (B)**, which was shown to have high similarity to ICE*Sp2907*. Of these isolates, 56 (94.9%) belonged to CC17, from which 24 (42.9%) were *emm*-type *stC74a*. The remaining isolates (n=42) presented elements that diverged significantly from ICE*Sp2907* and that were very diverse in sequence amongst each other. This way, a more detailed analysis of these elements should be performed in the future.

Although 9 *erm*(A) isolates were also resistant to tetracycline, in only two isolates (both belonging to CC8) was a tetracycline resistance gene (*tet*(W)) detected within the *erm*(A) containing element identified. The *ant*(6)-Ia gene, conferring resistance to aminoglycosides, was also found in five putative elements containing the *erm*(A) gene, with a distance of about 6-6.5 Kb between each of the genes.

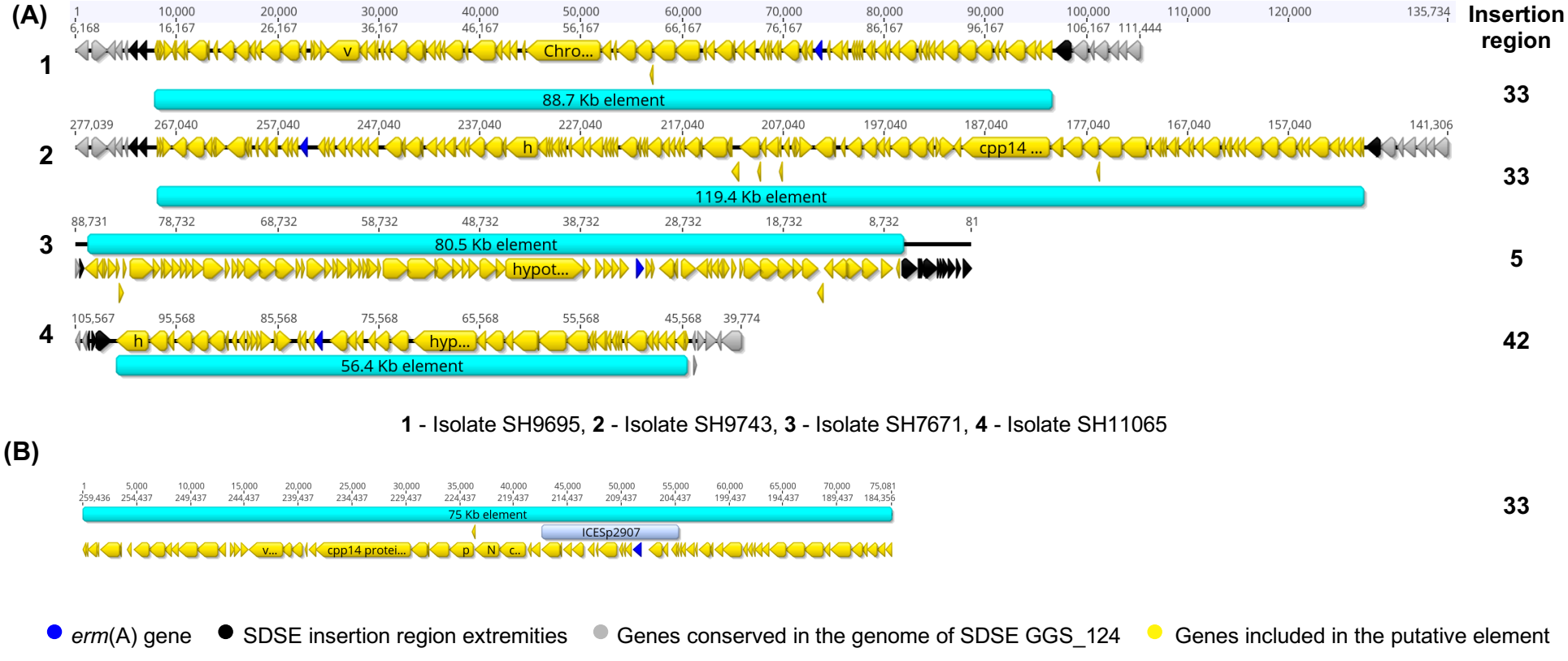


Figure 16 – **(A)** Detected complete elements containing the *erm*(A) gene and the corresponding SDSE GGS_124 insertion regions: 1 – full element located in insertion region 33; 2 – full element located in insertion region 33; 3 – full element located in insertion region 5; 4 – full element located in insertion region 42. **(B)** Representative of the most commonly found element among all contigs containing the *erm*(A) gene, here represented in the genome of isolate SH10053 and located in insertion region 33.

4.2.3.2. *erm*(B) gene

The *erm*(B) gene was found in 15 isolates recovered from ten different Portuguese hospitals. As for the hospitals in which SDSE was collected more than once throughout the years of study, there seems to be no correlation amongst the isolates, as none but two isolates were collected in the same hospital in the same year or in consecutive years. The isolates belonged to nine different CCs and 11 distinct *emm*-types.

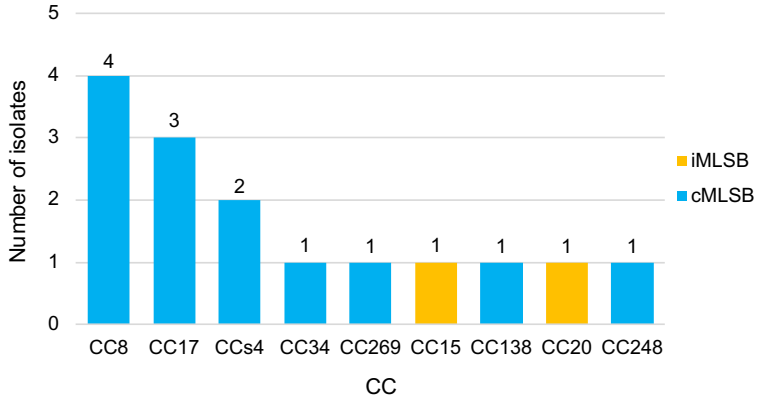
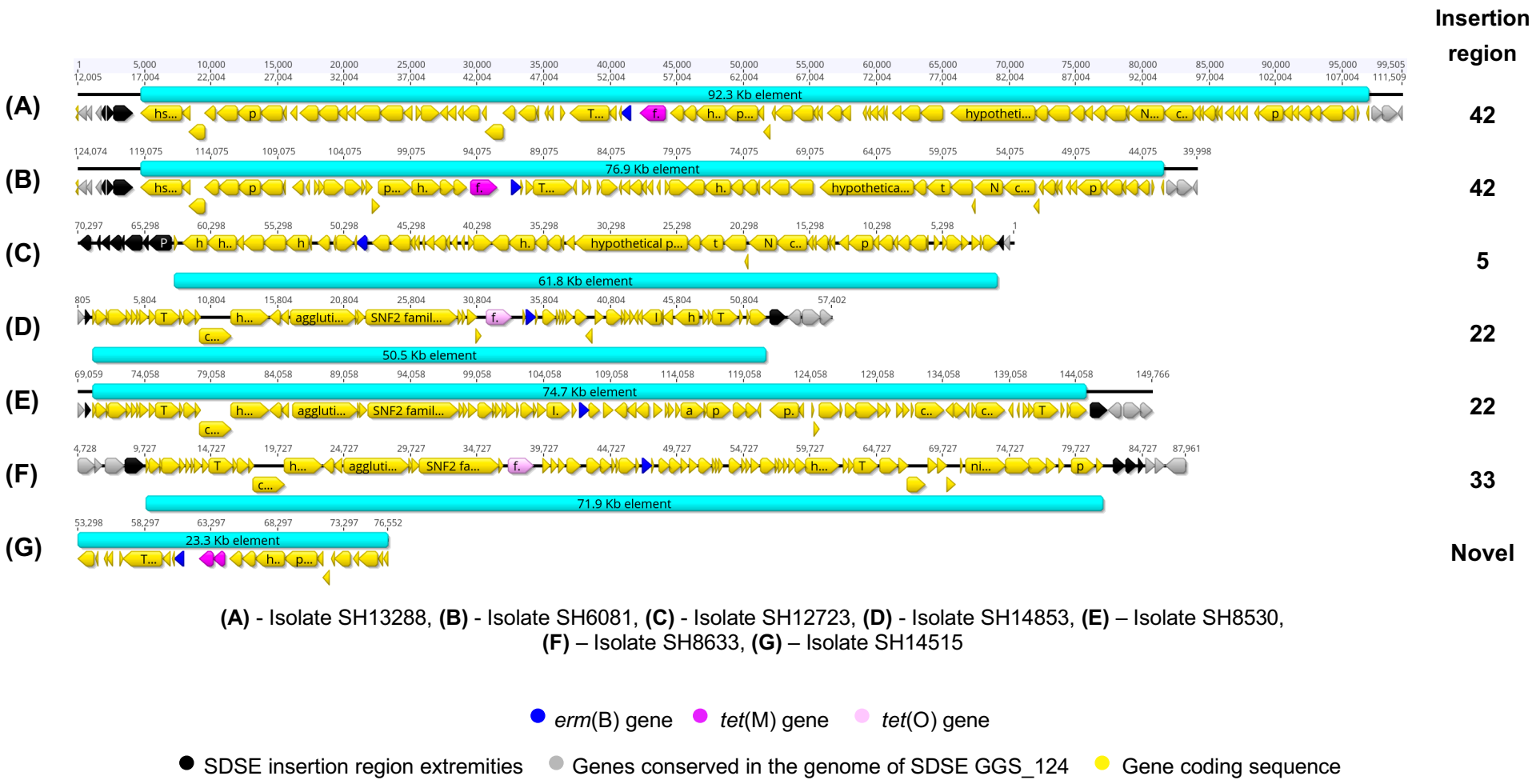


Figure 17 – Distribution of the CC of the invasive SDSE isolates recovered in Portugal from 2011 to 2018 harboring the *erm*(B) gene.

Table 7 - Characterization of invasive SDSE isolates (2011-2018) harboring the *erm*(B) gene.

Isolate	Year	Lancefield group	Product	Age	Sex	MLS phenotype	MLS genotype	*emm* type	MLST ST	CC
SH6081	2011	G	Blood	88	M	cMLSB	*erm*(B)	stG11	8	8
SH6409	2011	G	Blood	62	F	cMLSB	*erm*(B)	stG652	138	Singleton
SH7689	2012	G	Blood	77	F	cMLSB	*erm*(B)	stG2078	17	17
SH8223	2013	G	Blood	74	M	cMLSB	*erm*(B)	stG11	8	8
SH8530	2014	C	Blood	74	F	iMLSB	*erm*(B)	stG62647	20	20
SH8633	2014	L	Blood	51	F	cMLSB	*erm*(B) *lnu*(D) *mef*(A)**x2** *msr*(D)**x2**	stL1376	249	Singleton
SH9723	2014	G	Synovial fluid	67	M	cMLSB	*erm*(B)	stG11	8	8
SH11911	2016	C	Blood	34	F	cMLSB	*erm*(B)	stG6	499	s4
SH12320	2016	G	Blood	61	M	cMLSB	*erm*(B)	stG485	17	17
SH12723	2016	G	Blood	56	F	cMLSB	*erm*(B)	stG2078	17	17
SH13288	2016	G	Blood	77	M	cMLSB	*erm*(B)	stG643	8	8
SH14293	2017	C	Blood	35	F	cMLSB	*erm*(B)	stG6	499	s4
SH14515	2017	G	Blood	73	M	cMLSB	*erm*(B)	stG840	269	269
SH14853	2018	C	Blood	80	M	cMLSB	*erm*(B)	stGM220	34	34
SH15607	2018	G	Blood	73	F	iMLSB	*erm*(B)	stG166b	15	15

The *erm*(B) sequences of the 15 isolates were 99.3% identical. The full element (with conserved regions of the genome present at both ends) was identified in 8 of the 15 *erm*(B) isolates and ranged from 23.3 Kb to 92.3 Kb. The elements were found inserted in insertion regions 33 (n=4), region 42 (n=4, all CC8), region 22 (n=3), region 5 (n=2) and in one region not previously described (**Supplementary table 2**). In the last case, it was inserted between a bacteriocin immunity protein gene (locus tag SDEG_RS08455 in GGS_124 genome) and transketolase (*tkt*) gene (locus tag SDEG_RS08460 in GGS_124 genome). The insertion region could not be determined for one isolate because it was incomplete at both ends.



(A) - Isolate SH13288, **(B)** - Isolate SH6081, **(C)** - Isolate SH12723, **(D)** - Isolate SH14853, **(E)** – Isolate SH8530, **(F)** – Isolate SH8633, **(G)** – Isolate SH14515

● *erm*(B) gene ● *tet*(M) gene ● *tet*(O) gene

● SDSE insertion region extremities ● Genes conserved in the genome of SDSE GGS_124 ● Gene coding sequence

Figure 18 – Detected complete elements containing the *erm*(B) gene and the corresponding SDSE GGS_124 insertion regions: **(A)** – element from SH13288 located in insertion region 42; **(B)** - element from SH6081 located in insertion region 42; **(C)** – element from SH12723 located in insertion region 5; **(D)** – element from SH14853 located in insertion region 22; **(E)** – element from SH8530 located in insertion region 22; **(F)** – element from SH8633 located in insertion region 33; **(G)** – element from SH14515 located in a novel insertion region.

The element described in **Figure 18 (A**) was found in only one isolate. A BLAST search allowed to identify the presence of Tn*3872* (GenBank accession number: OP715845 [166]) with a similarity of around 96.2%. This ~23 Kb long element was located roughly between bases 28.5 Kb and 50 Kb. Simultaneously, this MGE was also found in the four isolates presenting the element described in **Figure 18 (B)**, although, in this case, the similarity between Tn*3872* and the corresponding region was >99%. Additionally, the element represented in **Figure 18 (G**) corresponds to the totally of Tn*3872* with an identity of 97.1%, in which the *tet*(M) gene was inserted. All isolates belonged to CC8, with the exception of one of the isolates presenting the element represented in **Figure 18 (B)**, and the isolate presenting the element described in **Figure 18 (G)**.

Two isolates presented the element described in **Figure 18 (C)**. This element showed a similarity of 91.1% to ICE*San*95*_hsdM* (GenBank accession number: PP359593). Although the extremities of the element found in the contig were highly different from the extremities of the MGE, the sequence (~50 Kb) between them presented a similarity of >99% to ICE*San95_hsdM*.

Figure 18 (D) describes an element found exclusively in one isolate. A BLAST search found the element to have an identity of 94.2% to a section of the genome of a strain of *Streptococcus suis* (GenBank accession number: KY400494). However, this similar region did not include the *erm*(B) gene. The element described in **Figure 18 (E)**, shared by three isolates, does not seem to correspond to any previously reported MGE carrying the *erm*(B) gene. Only a small section (~5.4 Kb) of the element, which included the resistance determinant, showed an identity of >99% to a portion of the genome of an *Enterococcus gallinarum* strain (GenBank accession number: CP050816 [167]), but this element also shared high similarity with sections of the genomes of other species of the *Enterococcus* genus. As for the element present in **Figure 18 (F)**, found in three isolates, a BLASTn search indicates similarity of 98% with a strain of *Streptococcus agalactiae* (GenBank accession number: CP118043), but the HIT corresponds only to the *erm*(B) gene also found in the genome of this *S. agalactiae* strain. One of the isolates presenting this element also possessed two copies of the mef(A) gene in a different contig.

4.2.3.3 *erm*(T) gene

The *erm*(T) gene was found in the genomes of three SDSE, all belonging to CC15. Two of the isolates were *emm*-type *stG652* while the other was non-typeable (NT). When comparing the laboratory and geographical location of collection, two isolates were retrieved in the same hospital but three years apart (2015 and 2018), whilst the other isolate was collected in another location in 2013 (**Table 8**). As such, no epidemiological relation between the isolates was shown.

Table 8 - Characterization of invasive SDSE isolates (2011-2018) harboring the *erm*(T) gene. **Abbreviations**: NT – non-typeable.

Isolate	Year	Lancefield group	Product	Age	Sex	MLS phenotype	MLS genotype	*emm* type	MLST ST	CC
SH7955	2013	G	Blood	47	M	iMLSB	*erm*(T)	*stG652*	15	15
SH10515	2015	G	Blood	67	F	iMLSB	*erm*(T)	*NT*	15	15
SH15900	2018	G	Blood	86	F	iMLSB	*erm*(T)	*stG652*	15	15

The gene present in the three isolates was 100% identical and was present in small contigs (5044bp) that were also identical in terms of sequence.

When performing a BLASTn search on NCBI using the sequence of one of the contigs presenting the *erm*(T) gene, a 100% query cover was obtained with SDSE plasmid *p5580* complete sequence (GenBank accession number: HE862394) for which was more than 99% sequence identity. This sequence was 4950 bp long (**Figure 19**).

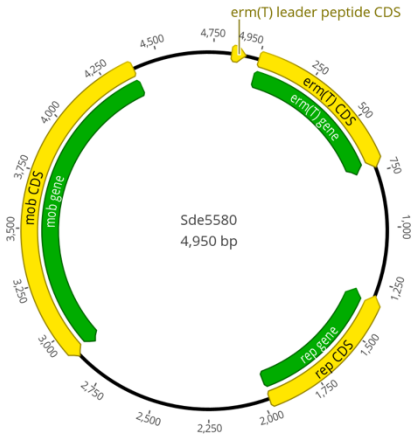


Figure 19 - SDSE plasmid *p5580* (GenBank accession number: HE862394).

4.2.3.4 *mef*(A) gene

The *mef*(A) gene was found in the genomes of four SDSE invasive isolates recovered from different patients in different hospitals. The isolates belonged to three CCs and four different emm-types. The *mef* gene was always found in tandem with the *msr*(D) gene. *mef*(A) sequences were extracted from the genome of these isolates (three M phenotype and one $cMLS_B$ phenotype). The isolate expressing the $cMLS_B$ phenotype (SH8633) possessed two copies of the *mef*(A) (below denoted as SH8633_A and SH8633_B) located in the same contig nearly 10kb apart, as well as a copy of the *erm*(B) gene in a different contig. The overall characteristics of the isolates are summarized in **Table 9**.

Table 9 – Characterization of invasive SDSE isolates (2011 to 2018) harboring the *mef*(A) gene.

Isolate	Year	Lancefield group	Product	Age	Sex	MLS phenotype	MLS genotype	*emm* type	MLST ST	CC
SH6079	2011	G	Blood	83	F	M	*mef*(A) *msr*(D)	*stG10*	15	15
SH8633	2014	L	Blood	51	F	cMLSB	*erm*(B) *lnu*(D) *mef*(A)**x2** *msr*(D)**x2**	*stL1376*	249	Singleton
SH16552	2018	C	Blood	92	F	M	*mef*(A) *msr*(D)	*stC7934*	35	34
SH17123	2018	C	Blood	86	M	M	*mef*(A) *msr*(D)	*stCK401*	35	34

The *mef*(A) gene can be divided into various subclasses, including *mef*(A), *mef*(B), *mef*(E), *mef*(G), *mef*(I) and *mef*(O) according to differences in sequence [168,169]. Geneious® was used to align sequences and build a neighbor-joining tree with the Tamura-Nei model, as shown in **Table 10** and **Figure 20.**

Four distinct alleles were found: the sequence of SH6079 showed high similarity to *mef*(A) sub-class; SH8633_A showed high similarity to *mef*(B) sub-class; SH8633_B showed high similarity to *mef*(E) sub-class; SH16552 and SH17123 had identical sequences which showed high similarity to sub-classes *mef*(I) (95%) and *mef*(E) (94%).

Table 10 - Distance matrix of *mef* alleles. Green highlights the highest identities between the *mef* genes found in each invasive SDSE isolate recovered in Portugal from 2011 to 2018 and the described sub-classes. Lighter green indicates a high identity to a sub-class, although not the highest for that isolate.

	mef(A)	*mef*(B)	*mef*(E)	*mef*(G)	*mef*(I)	*mef*(O)
SH6079	99.43	89.41	90.39	88.18	90.48	88.34
SH8633_A	89.33	100.00	91.46	89.41	92.04	88.83
SH8633_B	89.90	91.54	99.67	89.82	94.25	89.49
SH16552	89.49	92.36	93.76	92.36	94.99	91.54
SH17123	89.49	92.36	93.76	92.36	94.99	91.54

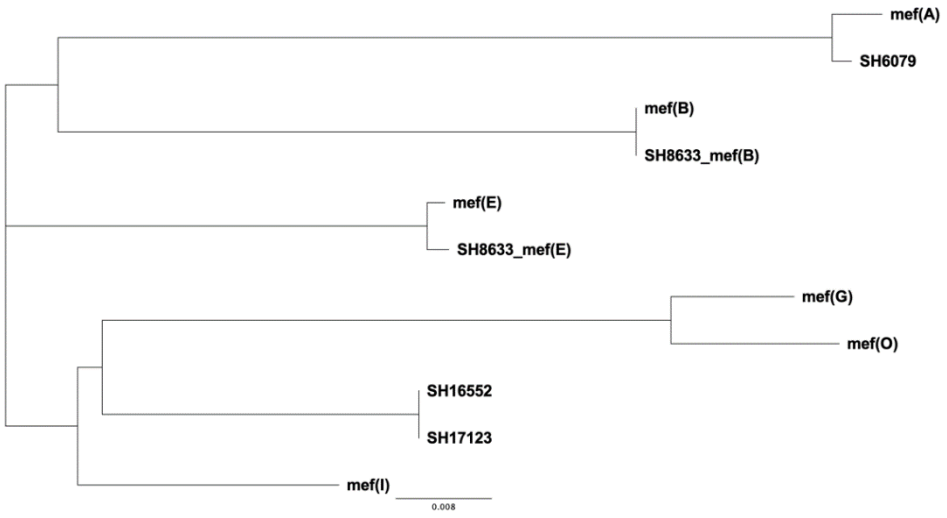


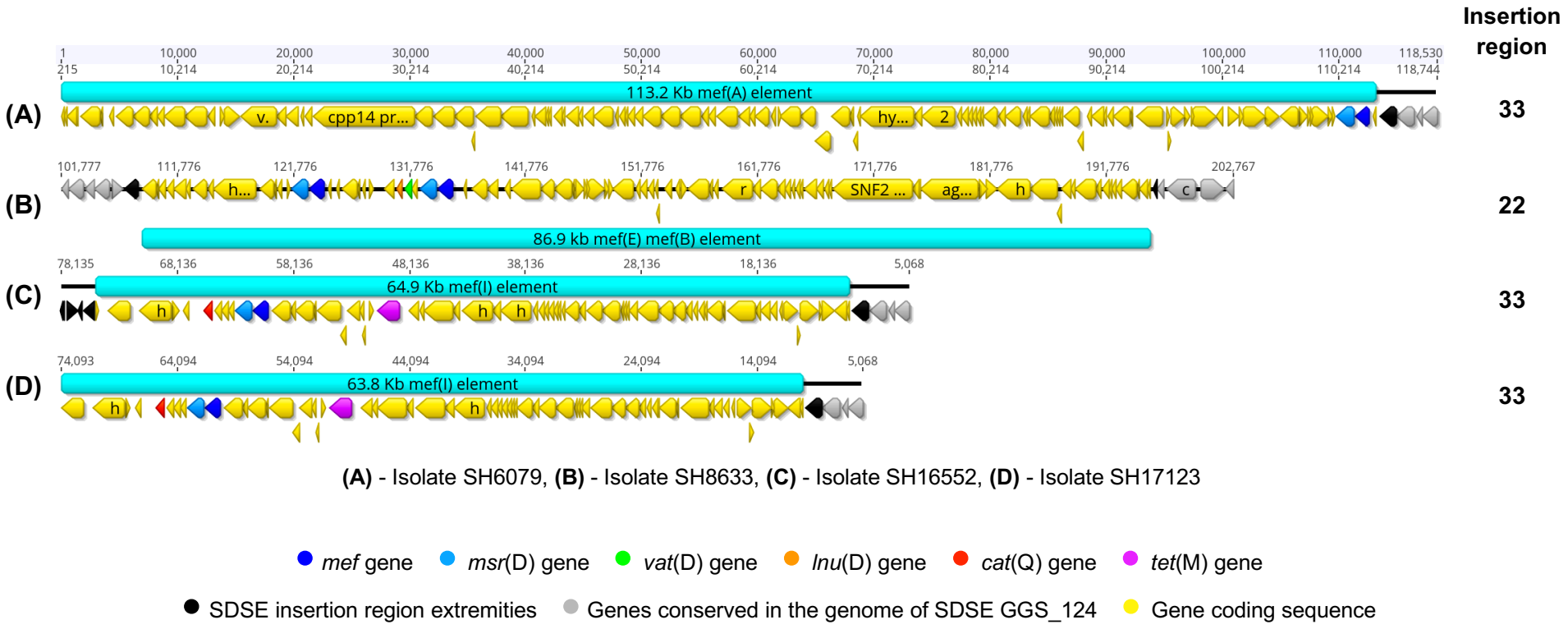
Figure 20 - Neighbor-joining tree including the *mef*(A) harboring invasive SDSE isolates recovered in Portugal from 2011 to 2018 and the *mef*(A) gene sub-classes.

The elements carrying *mef*(A) ranged from 113.2 Kb to 63.8 Kb and were inserted in insertion region 33 (n=3) or 22 (n=1) (**Supplementary table 3**).

SH6079 isolate carrying sub-class *mef*(A)

The *mef*(A)/*msr*(D) genes were present in a contig where a extremity of insertion region 33 of SDSE GGS_124 was identified, and downstream, a region conserved in SDSE GGS_124 was found (**Figure 21 (A)**). The other extremity of that same insertion region was identified in a different contig, followed by a region also conserved in SDSE GGS_124. Therefore, the presence of further genes in the element could not be confirmed, as it is not possible to exclude their presence in another contig.

PHASTEST identified a region in this contig as an intact phage and ICEfinder identified a putative ICE. The results of a BLAST search for these regions in NCBI (https://www.ncbi.nlm.nih.gov/) found a nearly identical phage in a *Streptococcus anginosus* isolate (GenBank accession number: CP118053) and the most similar ICE region was found in isolate *Streptococcus dysgalactiae* strain FDAARGOS_654. This isolate was also tetracycline resistant but the *tet*(M) gene was located in a distinct contig.



(A) - Isolate SH6079, **(B)** - Isolate SH8633, **(C)** - Isolate SH16552, **(D)** - Isolate SH17123

Figure 21 – Detected complete elements containing the *mef*(A) gene and the corresponding SDSE GGS_124 insertion regions: **(A)** – element from SH6079 located in insertion region 33; **(B)** - element from SH8633 located in insertion region 22; **(C)** – element from SH16552 located in insertion region 33; **(D)** – element from SH17123 located in insertion region 33.

SH8633 isolate carrying sub-class *mef*(B) and *mef*(E)

Isolate SH8633 was a group L isolate expressing the $cMLS_B$ phenotype, and additionally resistant to tetracycline and streptomycin (**Figure 21 (B)**). It possessed two copies of *mef* located in the same contig: one identical to *mef*(B) (upstream) and other nearly identical to *mef*(E) (downstream). An element containing both *mef* genes could be identified and, when in comparison to the genome of SDSE GGS_124, was found to be located in insertion region 22 of this genome. Between the two *mef* genes, the *vat*(D) gene, encoding for a streptogramin A acetyltransferase that induces resistance to streptogramin A was found, followed by the *lnu*(D) gene, encoding for a lincosamide nucleotidyltransferase, which confers resitance to lincosamides.

A BLASTn search NCBI (https://www.ncbi.nlm.nih.gov/) for the element found to harbor both genes, found higher similarity 93.3% to ICE*Sg5302* (GenBank accession number: MZ503511), but not to the region including the *mef* genes.

SH16552 and SH17123 isolates carrying sub-class *mef*(I)

The *mef* allele found in two related group C isolates, SH16552 (**Figure 21 (C)**) and SH17123 (**Figure 21 (D)**), had 95% sequence identity to the *mef*(I) allele described by Cochetti *et al*., in 2005 [159]. Comparison between the isolates' sequences and the genome of SDSE GGS_124 identified nearly identical mobile elements (~65 Kb) located in insertion region 33. A BLAST search of this element revealed a sequence identity of >95% to an *Anaerotignum* spp. isolate (GenBank accession number: CP097072), which did not encompass the region containing the resistance gene. However, for the small portion (~3300 bp) of the element encompassing the *mef*(I) and *msr*(D) genes, a sequence identity of >99% to a *Vagococcus lutrae* (GenBank accession number: CP133078) isolate was found for both isolates.

This element containing the *mef*(I) gene in the two isolates also presented both the *cat*(Q) (downstream of *mef*(I)) and *tet*(M) (upstream of *mef*(I)) genes, conferring resistance to tetracycline and chloramphenicol, respectively. A BLASTn search of the region including both the *cat*(Q) and *mef*(I) genes found similarity with isolates of various streptococcal species, but not to a specific mobile genetic element. Similarly, a BLASTn search of the region containing both *tet*(M) and *mef*(I) genes showed hit results with isolates of various bacterial genera but no specific mobile element.

5. Discussion

The rise of human invasive infections by SDSE described in various studies worldwide in recent years [170,171] has brought attention to the importance of this bacterium as a pathogen to the human host. In this study, characterization of the SDSE strains causing invasive disease in Portugal in the years 2019 to 2022 was performed in the context of an ongoing surveillance of SDSE invasive infections in the country. In total, 183 isolates were collected country-wide during this period, the majority being recovered from blood samples. Phenotypic characterization of the isolates was based on hemolysis observation and Lancefield group typing. Antimicrobial susceptibility testing was performed and the resistance patterns throughout this time period were evaluated, as well as compared with the resistance distribution of the years prior to the current study.

The number of SDSE invasive isolates received by the laboratory oscillated throughout the study period and no tendency for the number of cases to increase could be observed. SDSE invasive infections were more common amongst patients 65 years old or older (68%), with a median of 70 years old, with mostly male patients (57%). These results are in accordance with most studies that present the older age-groups as the most prone to these types of infections. In Europe, research has shown that invasive events caused by GCGS seem to be more prevalent amongst the older age groups when in comparison to other β-hemolytic streptococci [68]. Studies in Japan and Taiwan have also demonstrated the elderly to be more prone to invasive SDSE infections [70,172]. Interestingly, a population-based surveillance study conducted in the United States found that a substantial fraction (60%) of the patients were 64 years old or younger; however, around 42% and 30% of patients presented diabetes mellitus and obesity, respectively, highlighting the contribution of these comorbidities in the development of invasive SDSE diseases [173].

Most isolates presented Lancefield group G antigens (68%), followed by group C (32%), and with no occurrences of Lancefield groups L or A antigens. These results are in agreement with the ones reported from research in Portugal [174], as well as worldwide [55,172,173,175,176], that indicate that the SDSE isolates recovered from human disease predominantly present the Lancefield group G antigen.

Antimicrobial resistance

All 183 isolates in study were susceptible to penicillin. Given some concerns of reduced susceptibility to penicillin by beta-hemolytic streptococci, including the recent, yet rare, description of an SDSE strain resistant to this antibiotic [113], a screening protocol for detection of reduced susceptibility to beta-lactams was applied in the current study. A single isolate with reduced halos to ceftizoxime and ceftibuten was detected. While further characterization is needed to confirm reduced susceptibility in this isolate, the observations of the current study further support the claim that SDSE presents virtually no resistance to β-lactams. Simultaneously, all isolates were also susceptible to vancomycin, linezolid,

trimethoprim/sulfamethoxazole and chloramphenicol, antimicrobial agents that, despite not being first-line treatment options for SDSE invasive infections, retain activity against this bacterium.

Throughout the study period, 45.4% of all isolates were resistant to erythromycin. Additionally, 43.7% of the isolates were clinically resistant to clindamycin, meaning that, although the results of antimicrobial susceptibility testing may deem the isolate susceptible to this antimicrobial, if it presents the $iMLS_B$ resistance phenotype, the treatment will not include lincosamides. In 2022, the levels of resistance to erythromycin and clindamycin peaked, with values above 50% detected for the first time in the country, meaning that, in the last year of the study period, over half of the collected SDSE causing invasive disease in Portugal were resistant to both these antimicrobials. A review by Berbel *et al*. [116] has shown that, in the last decade, the proportion of macrolide and clindamycin resistant SDSE strains in the USA and Europe have reached values of 20 to 30%; numbers between 18 and 24% in Japan and Central Taiwan and 40 to 72% in Korea, Southern Taiwan, and China have been registered; in Africa and the Middle East, the numbers range between 8 and 28%. Additionally, studies only focusing in isolates causing invasive infections have reported macrolide and lincosamide resistance rates of 10% [84] and 31.5%[131] in Japan, and 21.4% in Hungary[171]. The results of this study seem to indicate that the macrolide resistance proportions of SDSE invasive isolates in Portugal is rising and above the average when in comparison with the values in Europe of the past years.

The distribution of the macrolide resistance phenotypes suffered no changes in occurrence throughout the study period. Overall, $iMLS_B$ was the most common of the phenotypes on all years, with 72.3% of macrolide resistant isolates displaying this characteristic, followed by 24.1% of isolates presenting the $cMLS_B$ phenotype, and 3.6% of the isolates exhibiting the M phenotype. These results are in accordance with research done both in Austria [177] and Argentina [178] that shows the $iMLS_B$ phenotype as the most common in SDSE. In contrast, studies in France [63] and Japan [14] have presented the $cMLS_B$ phenotype as the most common, and additionally, research in Brazil [179] found the M phenotype to be the most common amongst the SDSE isolates in study. These differences may be attributable to variations in the prevalence of specific macrolide and lincosamide resistant SDSE lineages across different countries associated with distinct MLS resistance phenotypes and resistance genes.

From all macrolide resistant isolates, the most part presented the *erm*(TR), a subclass of the *erm*(A) gene, present in most $iMLS_B$ phenotype expressing isolates, followed by the *erm*(B) gene, only found in isolates exhibiting the $cMLS_B$ phenotype. Many studies do not present results for the genes responsible for each of the phenotypes, or, in most cases, the results show the *erm*(A) and *erm*(B) encoding for both the $iMLS_B$ and the $cMLS_B$ [119,178,179]. However, some studies seem to demonstrate the presence of the *erm*(A) gene in association with the $iMLS_B$ phenotype, and the carriage of the *erm*(B) gene with the $cMLS_B$ phenotype. In fact, Wajima *et al*. found that, in a collection of SDSE isolates causing invasive disease in Japan, all isolates presenting the *erm*(A) gene expressed the $iMLS_B$ phenotype, and all harboring the *erm*(B) gene expressed the $cMLS_B$ phenotype, which corroborates the results of the present study. Three

isolates expressing the $iMLS_B$ phenotype presented the *erm*(T) gene, demonstrating that SDSE isolates can harbor the *erm*(T) gene, as previous studies have seldom shown [180], although *erm*(A) and *erm*(B) genes are most frequently found [121,125,180,181]. As expected, the remaining isolates expressing the M phenotype all presented the *mef*(A) gene.

A substantial rate of resistance to tetracycline was found, with around one-third (27.9%) of SDSE invasive isolates being resistant to this agent. This observation is in line with previous studies that have shown that tetracycline resistance is frequent in SDSE. In France, SDSE resistance to tetracycline was around 35% from 2006 to 2010, similarly to the values of tetracycline resistance in the last years in Portugal [63]. Research in Israel [44] has shown that, from all invasive SDSE isolates in study, 47.4% were resistant to tetracycline, while in Austria [177], tetracycline resistance rates of 62% were detected for SDSE isolates.

Throughout the period of 2019 to 2022, the proportion of resistance to some of the antimicrobials suffered some fluctuations. Overall, no tendency for the increase or decrease of resistance to any of the antimicrobials was detected. Only when examining the evolution of antimicrobial resistance from 2011 to 2022 is it possible to conclude that, in fact, there was a tendency for the resistance to erythromycin and clindamycin to increase amongst SDSE isolates causing invasive infections in Portugal. In contrast, there was a tendency for the number of levofloxacin resistant isolates to decrease, with the high levels of resistance to this antimicrobial detected in SDSE in the first years of the survey [139] no longer occurring.

The observed changes in resistance rates to antimicrobials, particularly the notorious increase in resistance to macrolides and lincosamides, prompted the molecular typing characterization of the SDSE invasive population, aiming at clarifying if this increase was attributable to specific SDSE resistant lineages and characterizing the MGEs contributing to the dispersion of resistance genes. Given that WGS data was only available to SDSE invasive isolates recovered up to 2018, all further studies were conducted with these isolates.

Genotypic characterization of macrolide and lincosamide resistance

Overall, 32.5% of the SDSE invasive isolates recovered between 2011 and 2018 were resistant to macrolides and lincosamides, a value expectably lower than the 45.4% detected in the period 2019-2022 in the context of rising erythromycin and clindamycin resistance. Nevertheless, no differences in the prevalence of phenotypes or genotypes of macrolide and lincosamide resistance were observed when comparing the two periods of study: the majority of resistant isolates recovered up to 2018 presented the $iMLS_B$ phenotype and only a minority the M phenotype; Concerning the resistance genes, most isolates harbored the *erm*(A) gene, followed by *erm*(B), *erm*(T) and *mef*(A). These observations suggest that increase in resistance to these antimicrobials was not accompanied by major changes in the underlying mechanisms.

MLST and *emm*-typing information extracted from WGS data was used to characterize SDSE invasive isolates. All isolates were separated into 81 MLST ST, 11 CCs and 28 *emm*-types. Overall, CC17 was the most common clonal complex. Various studies worldwide have similarly demonstrated CC17 to be one of the most prevalent CCs, such as in Norway [182], Canada [2] and Japan [1,131]. CC17 isolates in our study demonstrated high heterogeneity as it included 13 distinct *emm*-types, the most abundant being *stC74a*, which has also been shown to be amongst the most prevalent SDSE *emm*-types worldwide [183]. In Japan, *stG6792* is the main *emm*-type associated with CC17 [14], suggesting that different lineages within CCs may predominate in different geographic locations. Additionally, CC20 was the second most common CC, with 96.5% of its isolates belonging to *emm*-type *stG62647*, demonstrating a higher *emm*-type homogeneity. These results are in accordance to other studies that showed the emergence of a CC20-*stG62647* SDSE lineage that become one of the most prevalent in European countries in the last years [182,183].

In the present study, CC17 was associated with resistance to erythromycin and clindamycin. In parallel, *emm*-type *stC74a*, the most common *emm*-type in CC17, was the one also associated with resistance to both antimicrobials. As stated above, various studies show *stC74a* to be amongst the most prevalent *emm*-types in Europe [177,182,183], however, they do not indicate an association between *stC74a* and macrolide and lincosamide resistance. A study by Shinohara *et al*. [131] demonstrated CC17 as the most common in isolates causing bacteremia in Japan, as 34.2% of isolates belonged to this CC. However, CC25 was the one associated with macrolide and lincosamide resistance.

The *erm*(A) gene was the most common in invasive SDSE isolates from Portugal, being detected in 81.5% of all macrolide and lincosamide resistant isolates. Most (70.3%) of these isolates were included in CC17, while the remaining isolates were found sparsely in five other CCs. The genomic analysis of the contigs in which this gene was found allowed the identification of several elements most of them inserted in the same genomic location in the SDSE genome. The central part of the element most commonly found amongst *erm*(A) carrying isolates showed a similarity of 90.4% to ICE*Sp2907*. This ICE is inserted in a larger integrative and conjugative element, ICE*Sp2905*. However, only the region corresponding to both ICE*Sp2905* and ICE*Sp2907* (which contains the *erm(A) gene*) seemed to be identical to the isolates sequence. Various studies indicate ICE*Sp2907* as one of the typically found *erm*(A) carrying MGEs in different species [127]. A study by Oppegaard *et al.* [184] found that the presence of the *erm*(A) gene was related to ICE*Sp2907* in all SDSE isolates in study harboring this resistance determinant, found in 14 of the 19 isolates.

The *erm*(B) gene was found in 15 isolates, which seemed to have a heterogeneous CC distribution, as it was found in nine different CCs. These include the isolate containing two copies of the *mef*(A) gene. One element, found to be present in only one contig (Figure 19 (A)), showed a similarity of 96.2% to Tn*3872* (GenBank accession number: OP715845 [166]). Likewise, the element described in Figure 19 (B), found in four contigs, present an even higher similarity to Tn*3872*, of >99%, and, additionally, the element in Figure

19 (G) presented a similarity of 97.1% to Tn*3872*. This element was first found in *S. agalactiae* and resulted from the insertion of the Tn*917* transposon containing the *erm*(B) gene into Tn*916* [185]. Five of the six isolates containing this element belonged to CC8. The element shown in Figure 19 (C) was found in two isolates, and it presented a high similarity to ICE*San95_hsdM* [186]. This element was first found in *Streptococcus anginosus* and its cross-species transfer from *S. anginosus* to other streptococci was demonstrated. Both isolates belonged to CC17 and presented the same *emm*-type. All other elements (Figure 19 (D), (E), (F)) did not seem to correspond to any described element containing the *erm*(B) gene. Shinohara *et al.* [131] demonstrated a link between the *erm*(B) gene to the presence of Tn916-like ICEs, but, in contrast with the present study, the *erm*(B) harboring isolates belonged mainly to CC25. Overall, despite present in a limited number of isolates, the diversity of *erm*(B) containing elements highlights the ability of several SDSE lineages to acquire these resistance determinants.

Three isolates harbor the *erm*(T) gene and expressed the iMLS$_B$ phenotype, which corroborates the idea that, although it may be found in SDSE, it is uncommon in this species [121,125,180,181]. On all isolates, the *erm*(T) gene was located in a small contig simlar to the plasmid *p5580* previously described in SDSE. In the work by Palmieri *et al*. [125], an SDSE isolate (Sde5580) with macrolide resistance was investigated for the presence of a macrolide resistance gene. It was found that the isolate harbored the erm(T) gene and that the sequence encoding for it was present in an *erm*(T)-carrying plasmid, designated p5580. Researchers learned that this plasmid was, in fact, almost identical to a non-self-transmissible *erm*(T)-carrying plasmid carried by *S. pyogenes* and *S. agalactiae*, and that the highest BLASTn score (>99% identity) was for the pGB2001 [187] plasmid of *S. agalactiae*. This isolate also presented an integrative and conjugative element (ICE*Sde3396*-like) from the ICESa2603 family that presents its own machinery for conjugation. Since p5580 is a non-self-transmissible plasmid, the trans mobilization activity of ICE*Sde3396*-like seems to be essential for the transfer and conjugation of the plasmid to other isolates, allowing for the transfer of the plasmid while the transfer of the ICE*Sde3396*-like itself is not necessary [125].

Four isolates presented the *mef*(A) macrolide resistance gene, and all expressed the M phenotype. Allelic variants of the *mef*(A) gene found in streptococcal species have been separated in subclasses showing sequence divergence at the nucleotide level of 88 to 94% [168] (reviewed in [169]). Previous studies have shown that *mef*(A) subclasses are associated with distinct genetic elements, with some of these being shared between *S. pyogenes* and *S. pneumoniae* [168]. Literature has shown that subclasses *mef*(A) and *mef*(E) are the most common *mef*(A)-class genes present in these two species, however, their prevalence might change [160,188]. Moreover, not many studies report and differentiate these *mef*(A) gene subclasses in SDSE isolates, and commonly group them as *mef*(A/E) [179]. Some reports that do distinguish them do not show a staggering difference in the proportion of SDSE isolates that harbor the *mef*(A) and the ones that carry the *mef*(E) gene subclass [82,189]. Despite the small number of isolates carrying *mef* genes in our study, sequence analysis demonstrated that invasive SDSE isolates carried *mef* genes from different subclasses, *mef*(A), *mef*(E), *mef*(B) and *mef*(I). Additionally, one of the isolates presented two copies of the *mef*(A)

gene, as well as a copy of the *erm*(B) gene. This same isolates belonged to Lancefield group L, which is less commonly associated with human infections and more commonly to animal SDSE infections [30]. There seems to be no published research documenting an SDSE isolate carrying both the *mef*(B) and *mef*(E) genes in close proximity such as the one this work reports.

The molecular characterization of SDSE macrolide and lincosamide resistant isolates in the present study was instrumental to show that, although clonal expansion of CC17 might have contributed to the high rates of resistance to these antimicrobials seen in the last years of the study, resistance is present in many different unrelated genetic backgrounds, as genes conferring resistance were found in isolates of many distinct CCs and *emm*-types. Moreover, exploration of the genomic context in which these genes are present showed the presence of diverse MGEs.

6. Conclusion

SDSE is an emergent pathogen with the capacity of causing a diverse range of invasive and non-invasive infections, with research supporting the increase of the number of cases of SDSE invasive infections throughout the world.

This study allowed to gather an insight on how SDSE invasive infections have evolved in the most recent years in Portugal. Between the years of 2019 and 2022, no tendency for SDSE invasive infections increasing in Portugal is evident, with only slight fluctuations in the number of cases being registered from year to year. Nonetheless, antimicrobial susceptibility testing determined rather high macrolide, lincosamide, and tetracycline resistance rates, which had a tendency to increase from 2011 to 2022. The lack of knowledge on the mechanisms of action and virulence of SDSE, paired with the higher rates of antimicrobial resistance registered, highlight the importance of a continuous monitorization of these infections.

This work also demonstrated the possibility of acquisition of antimicrobial resistance genes, possibly from events of HGT, as elements containing the different macrolide and lincosamide resistance determinants originally detected in other species were here found in SDSE. This aspect further intensifies the need to monitor resistance in SDSE, as well as other species.

6.1. Future perspectives

Although there seemed to be no tendency for the increase of the number of human invasive SDSE infections in Portugal in the past decade, the proportion of macrolide and lincosamide resistance suffered a notable increase. This way, several aspects could be explored in future work:

- Study the clonal distribution of the invasive SDSE isolates recovered in Portugal from 2019 to 2022 once the sequencing results are available;
- Study the antimicrobial resistance determinants distribution of the invasive SDSE isolates recovered in Portugal from 2019 to 2022 once the sequencing results are available;
- Utilize techniques using data collected from WGS, such as cgMLST, to further analyze the clonality of the isolates;
- Further investigate the presence of mobile genetic elements in the genomes of SDSE isolates, their origin, and their mobilization mechanisms;
- Investigate the presence of virulence genes important for the spread of infections caused by SDSE;
- Continue the investigation of the changes in antimicrobial resistance, in special to macrolides and lincosamides and to tetracycline;
- Maintain a continuous epidemiological surveillance of the invasive SDSE infections occurring in Portugal to be able to contain and treat these infections.

7. References

1. Fujita, T., Horiuchi, A., Ogawa, M., Yoshida, H., Hirose, Y., Nagano, N., & Takahashi, T. (2017). Genetic diversity in *Streptococcus dysgalactiae* subspecies *equisimilis* isolates from patients with invasive and noninvasive infections in a Japanese university hospital (2014-2015). *Japanese Journal of Infectious Diseases*, *70*(1), 100–104.
2. Lother, S. A., Demczuk, W., Martin, I., Mulvey, M., Dufault, B., Lagacé-Wiens, P., & Keynan, Y. (2017). Clonal clusters and virulence factors of group C and G *Streptococcus* causing severe infections, Manitoba, Canada, 2012–2014. *Emerging Infectious Diseases*, *23*(7), 1079–1088.
3. Vieira, V. V., Teixeira, L. M., Zahner, V., Momen, H., Facklam, R. R., Steigerwalt, A. G., Brenner, D. J., & Castro, A. C. D. (1998). Genetic relationships among the different phenotypes of *Streptococcus dysgalactiae* strains. *International Journal of Systematic Bacteriology*, *48*(4), 1231–1243.
4. McNeilly, C. L., & McMillan, D. J. (2014). Horizontal gene transfer and recombination in *Streptococcus dysgalactiae* subspecies *equisimilis*. *Frontiers in Microbiology*, 676(*5)*.
5. Skerman, V. B. D., McGowan, V., & Sneath, P. H. A. (Eds.). (1989). *Approved lists of bacterial names (amended)*. ASM Press.
6. Garvie, E. I., Farrow, J. A. E., & Bramley, A. J. (1983). *Streptococcus dysgalactiae* (Diernhofer) nom. Rev. *International Journal of Systematic and Evolutionary Microbiology*, *33*(2), 404–405.
7. Vandamme, P., Pot, B., Falsen, E., Kersters, K., & Devriese, L. A. (1996). Taxonomic study of Lancefield streptococcal groups C, G, and L (*Streptococcus dysgalactiae*) and proposal of *S. dysgalactiae* subspecies *equisimilis* subspecies nov. *International Journal of Systematic Bacteriology*, *46*(3), 774–781.
8. Kloos, W., Hardie, J., & Whiley, R. (2001). International Committee on Systematic Bacteriology. Subcommittee on the taxonomy of staphylococci and streptococci. *International Journal of Systematic and Evolutionary Microbiology*, *51*(2), 717–718.
9. Ciszewski, M., Zegarski, K., & Szewczyk, E. M. (2016). *Streptococcus dysgalactiae* subspecies *equisimilis* isolated from infections in dogs and humans: Are current subspecies identification criteria accurate? *Current Microbiology*, *73*(5), 684–688.

10. Jensen, A., & Kilian, M. (2012). Delineation of *Streptococcus dysgalactiae*, its subspecies, and its clinical and phylogenetic relationship to *Streptococcus pyogenes*. *Journal of Clinical Microbiology*, *50*(1), 113–126.

11. Simpson, W. J., Robbins, J. C., & Cleary, P. P. (1987). Evidence for group A-related M protein genes in human but not animal-associated group G streptococcal pathogens. *Microbial Pathogenesis*, *3*(5), 339–350.

12. Sriprakash, K. S., & Hartas, J. (1996). Lateral genetic transfers between group A and G streptococci for M-like genes are ongoing. *Microbial Pathogenesis*, *20*(5), 275–285.

13. McMillan, D. J., Vu, T., Bramhachari, P. V., Kaul, S. Y., Bouvet, A., Shaila, M. S., Karmarkar, M. G., & Sriprakash, K. S. (2010). Molecular markers for discriminating *Streptococcus pyogenes* and *S. dysgalactiae* subspecies *equisimilis*. *European Journal of Clinical Microbiology & Infectious Diseases*, *29*(5), 585–589.

14. Wajima, T., Morozumi, M., Hanada, S., Sunaoshi, K., Chiba, N., Iwata, S., & Ubukata, K. (2016). Molecular characterization of invasive *Streptococcus dysgalactiae* subspecies *equisimilis*, Japan. *Emerging Infectious Diseases*, *22*(2), 247–254.

15. Ahmad, Y., Gertz, R. E., Li, Z., Sakota, V., Broyles, L. N., Van Beneden, C., Facklam, R., Shewmaker, P. L., Reingold, A., Farley, M. M., & Beall, B. W. (2009). Genetic relationships deduced from *emm* and multilocus sequence typing of invasive *Streptococcus dysgalactiae* subspecies *equisimilis* and *S. canis* recovered from isolates collected in the United States. *Journal of Clinical Microbiology*, *47*(7), 2046–2054.

16. Patterson, M. J. (1996). Streptococcus. In S. Baron (Ed.), *Medical Microbiology* (4th ed.). University of Texas Medical Branch at Galveston.

17. Holding, A. J., & Collee, J. G. (1971). Chapter I: routine biochemical tests. In J. R. Norris & D. W. Ribbons (Eds.), *Methods in Microbiology* (Vol. 6, pp. 1–32). Academic Press.

18. Singhal, N., Kumar, M., Kanaujia, P. K., & Virdi, J. S. (2015). MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*, *6*, 791.

19. Baracco, G. J. (2019). Infections caused by group C and G *Streptococcus* (*Streptococcus dysgalactiae* subspecies *equisimilis* and others): Epidemiological and clinical aspects. *Microbiology Spectrum*, *7*(2), 7.2.32.

20. Brandt, C. M., & Spellerberg, B. (2009). Human infections due to *Streptococcus dysgalactiae* subspecies *equisimilis*. *Clinical Infectious Diseases*, *49*(5), 766–772.

21. Haslam, D. B. (2018). *122: Groups C and G streptococci*, 736.

22. Farrow, J. A. E., & Collins, M. D. (1984). Taxonomic studies on streptococci of serological groups C, G and L and possibly related taxa. *Systematic and Applied Microbiology*, *5*(4), 483–493.

23. Sierig, G., Cywes, C., Wessels, M. R., & Ashbaugh, C. D. (2003). Cytotoxic effects of streptolysin O and streptolysin S enhance the virulence of poorly encapsulated group A streptococci. *Infection and Immunity*, *71*(1), 446–455.

24. Turner, C. E., Bubba, L., & Efstratiou, A. (2019). Pathogenicity factors in group C and G streptococci. *Microbiology Spectrum*, *7*(3), 7.3.4.

25. Nizet, V. (2002). Streptococcal beta-hemolysins: Genetics and role in disease pathogenesis. *Trends in Microbiology*, *10*(12), 575–580.

26. Steiner, K., & Malke, H. (2002). Dual control of streptokinase and streptolysin S production by the *covRS* and *fasCAX* two-component regulators in *Streptococcus dysgalactiae* subspecies *equisimilis*. *Infection and Immunity*, *70*(7), 3627–3636.

27. Lancefield, R. C. (1933). A serological differentiation of human and other groups of hemolyitc streptococci. *The Journal of Experimental Medicine*, *57*(4), 571–595.

28. *Table: Lancefield Classification**. (2023). MSD Manual Professional Edition. https://www.msdmanuals.com/professional/multimedia/table/lancefield-classification

29. Pinho, M. D. (2014). *Lancefield group C and G streptococci in human infection: Molecular typing, virulence and antimicrobial susceptibility* [doctoralThesis].

30. Facklam, R. (2002). What happened to the streptococci: Overview of taxonomic and nomenclature changes. *Clinical Microbiology Reviews*, *15*(4), 613–630.

31. Nybakken, E. J., Oppegaard, O., Gilhuus, M., Jensen, C. S., & Mylvaganam, H. (2021). Identification of *Streptococcus dysgalactiae* using matrix-assisted laser desorption/ionization-time of flight mass

spectrometry; refining the database for improved identification. *Diagnostic Microbiology and Infectious Disease*, *99*(1), 115207.

32. Tsuyuki, Y., Kurita, G., Murata, Y., Goto, M., & Takahashi, T. (2017). Identification of group G streptococcal isolates from companion animals in Japan and their antimicrobial resistance patterns. *Japanese Journal of Infectious Diseases*, *70*(4), 394–398.
33. Preziuso, S., Pinho, M. D., Attili, A. R., Melo-Cristino, J., Acke, E., Midwinter, A. C., Cuteri, V., & Ramirez, M. (2014). PCR based differentiation between *Streptococcus dysgalactiae* subspecies *equisimilis* strains isolated from humans and horses. *Comparative Immunology, Microbiology and Infectious Diseases*, *37*(3), 169–172.
34. Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Molecular Diagnosis*, *6*(4), 313–321.
35. Pei, A., Nossa, C. W., Chokshi, P., Blaser, M. J., Yang, L., Rosmarin, D. M., & Pei, Z. (2009). Diversity of 23S rRNA genes within individual prokaryotic genomes. *PLOS ONE*, *4*(5), e5437.
36. Simar, S. R., Hanson, B. M., & Arias, C. A. (2021). Techniques in bacterial strain typing: Past, present, and future. *Current Opinion in Infectious Diseases*, *34*(4), 339–345.
37. Lancefield, R. C. (1943). Studies on the antigenic composition of group A hemolytic streptococci: I. Effects of proteolytic enxymes on streptococcal cells. *The Journal of Experimental Medicine*, *78*(6), 465–476.
38. Brandis, H. (1978). Chapter X: Bacteriocins of streptococci and bacteriocin typing. In T. Bergan & J. R. Norris (Eds.), *Methods in Microbiology* (Vol. 12, pp. 221–239). Academic Press.
39. Efstratiou, A. (1987). *The pyogenic streptococci of Lancefield group C and group G as pathogens in man* [Doctoral]. City of London Polytechnic.
40. Lancefield, R. C. (1928). The aantigenic complex of *Streptococcus haemolyticus*. *The Journal of Experimental Medicine*, *47*(1), 91–103.
41. Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., & Spratt, B. G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms.

Proceedings of the National Academy of Sciences of the United States of America, *95*(6), 3140–3145.

42. Schwartz, D. C., & Cantor, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, *37*(1), 67–75.

43. Facklam, R., Beall, B., Efstratiou, A., Fischetti, V., Johnson, D., Kaplan, E., Kriz, P., Lovgren, M., Martin, D., Schwartz, B., Totolian, A., Bessen, D., Hollingshead, S., Rubin, F., Scott, J., & Tyrrell, G. (1999). Emm typing and validation of provisional M types for group A streptococci. *Emerging Infectious Diseases*, *5*(2), 247–253.

44. Halperin, T., Levine, H., Korenman, Z., Burstein, S., Amber, R., Sela, T., & Valinsky, L. (2016). Molecular characterization and antibiotic resistance of group G streptococci in Israel: Comparison of invasive, non-invasive and carriage isolates. *European Journal of Clinical Microbiology & Infectious Diseases*, *35*(10), 1649–1654.

45. Castro, A. C. S. R. de. (2020). *Caracterização genómica de* Streptococcus dysgalactiae *subspecies* equisimilis *responsável por infeção invasiva no Homem* [masterThesis]. https://repositorio.ul.pt/handle/10451/47419

46. Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., & Spratt, B. G. (2004). eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *Journal of Bacteriology*, *186*(5), 1518–1530.

47. Francisco, A. P., Bugalho, M., Ramirez, M., & Carriço, J. A. (2009). Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics*, *10*(1), 152.

48. McMillan, D. J., Bessen, D. E., Pinho, M., Ford, C., Hall, G. S., Melo-Cristino, J., & Ramirez, M. (2010). Population genetics of *Streptococcus dysgalactiae* subspecies *equisimilis* reveals widely dispersed clones and extensive recombination. *PLOS ONE*, *5*(7), e11741.

49. Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volckaert, G., & Ysebaert, M. (1976). Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature*, *260*(5551), 500–507.

50. Blanc, D. S., Magalhães, B., Koenig, I., Senn, L., & Grandbastien, B. (2020). Comparison of whole genome (wg-) and core genome (cg-) MLST (BioNumericsTM) versus SNP variant calling for epidemiological investigation of *Pseudomonas aeruginosa*. *Frontiers in Microbiology*, *11*.

51. Shimomura, Y., Okumura, K., Murayama, S. Y., Yagi, J., Ubukata, K., Kirikae, T., & Miyoshi-Akiyama, T. (2011). Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subspecies *equisimilis* strain causing streptococcal toxic shock syndrome (STSS). *BMC Genomics*, *12*(1), 17.

52. Watanabe, S., Kirikae, T., & Miyoshi-Akiyama, T. (2013). Complete genome sequence of *Streptococcus dysgalactiae* subspecies *equisimilis* 167 carrying Lancefield group C antigen and comparative genomics of *S. dysgalactiae* subspecies *equisimilis* strains. *Genome Biology and Evolution*, *5*(9), 1644–1651.

53. Agerhäll, M., Henrikson, M., Johansson Söderberg, J., Sellin, M., Tano, K., Gylfe, Å., & Berggren, D. (2021). High prevalence of pharyngeal bacterial pathogens among healthy adolescents and young adults. *APMIS: Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, *129*(12), 711–716.

54. Jaalama, M., Palomäki, O., Vuento, R., Jokinen, A., & Uotila, J. (2018). Prevalence and clinical significance of *Streptococcus dysgalactiae* subspecies *equisimilis* (groups C or G streptococci) colonization in pregnant women: A retrospective cohort study. *Infectious Diseases in Obstetrics and Gynecology*, *2018*, 1–7.

55. McDonald, M., Towers, R. J., Andrews, R. M., Carapetis, J. R., & Currie, B. J. (2007). Epidemiology of *Streptococcus dysgalactiae* subspecies *equisimilis* in tropical communities, Northern Australia. *Emerging Infectious Diseases*, *13*(11), 1694–1700.

56. Tartof, S. Y., Farrimond, F., de Matos, J. A., Reis, J. N., Ramos, R. T. T., Andrade, A. N., dos Reis, M. G., & Riley, L. W. (2011). Inverse association between Lancefield group G *Streptococcus* colonization and sore throat in slum and nonslum settings in Brazil. *Journal of Clinical Microbiology*, *49*(1), 409–412.

57. Eriksson, B. K. (1999). Anal colonization of group G beta-hemolytic streptococci in relapsing erysipelas of the lower extremity. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *29*(5), 1319–1320.

58. Semel, J. D., & Goldin, H. (1996). Association of athlete's foot with cellulitis of the lower extremities: Diagnostic value of bacterial cultures of ipsilateral interdigital space samples. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *23*(5), 1162–1164.

59. Riccardi, N., Rotulo, G. A., & Castagnola, E. (2019). Definition of opportunistic infections in immunocompromised children on the basis of etiologies and clinical features: A summary for practical purposes. *Current Pediatric Reviews*, *15*(4), 197–206.

60. Lin, J.-N., Chang, L.-L., Lai, C.-H., Lin, H.-H., & Chen, Y.-H. (2011). Clinical and molecular characteristics of invasive and noninvasive skin and soft tissue infections caused by group A *Streptococcus*. *Journal of Clinical Microbiology*, *49*(10), 3632–3637.

61. Hart, K. M. (2016). Rhinosinusitis and tonsillopharyngitis. *Family Medicine*, 519–526.

62. Turner, J. C., Hayden, G. F., Kiselica, D., Lohr, J., Fishburne, C. F., & Murren, D. (1990). Association of group C beta-hemolytic streptococci with endemic pharyngitis among college students. *JAMA*, *264*(20), 2644–2647.

63. Loubinoux, J., Plainvert, C., Collobert, G., Touak, G., Bouvet, A., & Poyart, C. (2013). Adult invasive and noninvasive infections due to *Streptococcus dysgalactiae* subspecies *equisimilis* in France from 2006 to 2010. *Journal of Clinical Microbiology*, *51*(8), 2724–2727.

64. Matsue, M., Ogura, K., Sugiyama, H., Miyoshi-Akiyama, T., Takemori-Sakai, Y., Iwata, Y., Wada, T., & Okamoto, S. (2020). Pathogenicity characterization of prevalent-type *Streptococcus dysgalactiae* subspecies *equisimilis* strains. *Frontiers in Microbiology*, *11*, 97.

65. Nevanlinna, V., Huttunen, R., Aittoniemi, J., Luukkaala, T., & Rantala, S. (2023). Incidence, seasonal pattern, and clinical manifestations of *Streptococcus dysgalactiae* subspecies *equisimilis* bacteremia; a population-based study. *European Journal of Clinical Microbiology & Infectious Diseases*, *42*(7), 819–825.

66. Rantala, S. (2014). *Streptococcus dysgalactiae* subspecies*equisimilis* bacteremia: An emerging infection. *European Journal of Clinical Microbiology & Infectious Diseases*, *33*(8), 1303–1310.

67. Oppegaard, O., Mylvaganam, H., Skrede, S., Jordal, S., Glambek, M., & Kittang, B. R. (2017). Clinical and molecular characteristics of infective β-hemolytic streptococcal endocarditis. *Diagnostic Microbiology and Infectious Disease*, *89*(2), 135–142.

68. Ekelund, K., Skinhøj, P., Madsen, J., & Konradsen, H. B. (2005). Invasive group A, B, C and G streptococcal infections in Denmark 1999–2002: Epidemiological and clinical aspects. *Clinical Microbiology and Infection*, *11*(7), 569–576.

69. Rantala, S., Vuopio-Varkila, J., Vuento, R., Huhtala, H., & Syrjänen, J. (2009). Clinical presentations and epidemiology of beta-haemolytic streptococcal bacteraemia: A population-based study. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *15*(3), 286–288.

70. Liao, C.-H., Liu, L.-C., Huang, Y.-T., Teng, L.-J., & Hsueh, P.-R. (2008). Bacteremia caused by Group G streptococci, Taiwan. *Emerging Infectious Diseases*, *14*(5), 837–840.

71. Lother, S. A., Jassal, D. S., Lagacé-Wiens, P., & Keynan, Y. (2017). Emerging group C and group G streptococcal endocarditis: A Canadian perspective. *International Journal of Infectious Diseases*, *65*, 128–132.

72. Kumar, A., Sandoe, J., & Kumar, N. (2005). Three cases of vertebral osteomyelitis caused by *Streptococcus dysgalactiae* subspeccies *equisimilis*. *Journal of Medical Microbiology*, *54*(11), 1103–1105.

73. Asplin, C. M., Beeching, N. J., & Slack, M. P. E. (1979). Osteomyelitis due to *Streptococcus equisimilis* (group C). *BMJ*, *1*(6156), 89–89.

74. Richette, P., Pizzuti, P., Quillard, A., Raskine, L., Naveau, B., & Lioté, F. (2001). A definite case of spondylodiscitis caused by *Streptococcus dysgalactiae*. *Clinical and Experimental Rheumatology*, *19*(5), 587–588.

75. Zarrif, H., Navarro-Navarro, R., Bolaños-Rivero, M., & Pisos-Álamo, E. (2014). Spondylodiscitis by *Streptococcus dysgalactiae* subspecies *equisimilis* (group G). *Revista Espanola De Quimioterapia: Publicacion Oficial De La Sociedad Espanola De Quimioterapia*, *27*(2), 130–131.

76. B. P., M.-C., Rodiguez, P. A., Blazquez, A., Castellon Perez, M., Garcia, C. S., Ruiz, A. M., Alfayate Minguelez, S., Requena, E. M., & Arias, L. G. (2016). Septic osteoarthritis due to *Streptococcus*

dysgalactiae subspecies *equisimilis* in a three-year-old girl. *Archives of Pediatric Infectious Diseases*, *4*(2). https://brieflands.com/articles/apid-20292.html

77. Erden, T. (2020). Periprosthetic joint infection with *Streptococcus dysgalactiae* subspecies *equisimilis*: Case report. *Joint Diseases and Related Surgery*, *31*(2), 399–402.

78. Pham, T.-T., Lazarevic, V., Gaia, N., Girard, M., Cherkaoui, A., Suva, D., & Schrenzel, J. (2020). Second Periprosthetic Joint Infection Caused by *Streptococcus dysgalactiae*: How Genomic Sequencing Can Help Defining the Best Therapeutic Strategy. *Frontiers in Medicine*, *7*, 53.

79. Kujath, P., & Kujath, C. (2010). Complicated skin, skin structure and soft tissue infections—Are we threatened by multi-resistant pathogens? *European Journal of Medical Research*, *15*(12), 544–553.

80. Bruun, T., Kittang, B. R., De Hoog, B. J., Aardal, S., Flaatten, H. K., Langeland, N., Mylvaganam, H., Vindenes, H. A., & Skrede, S. (2013). Necrotizing soft tissue infections caused by *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subspecies *equisimilis* of groups C and G in western Norway. *Clinical Microbiology and Infection*, *19*(12), E545–E550.

81. Schmitz, M., Roux, X., Huttner, B., & Pugin, J. (2018). Streptococcal toxic shock syndrome in the intensive care unit. *Annals of Intensive Care*, *8*(1), 88.

82. Ikebe, T., Okuno, R., Sasaki, M., Kanda, Y., Otsuka, H., Kawahara, R., Ohya, H., Suzuki, M., Uchida, K., Nihonmatsu, H., Ohnishi, M., & Working Group for Beta-Hemolytic Streptococci in Japan. (2018). Molecular characterization and antibiotic resistance of Streptococcus dysgalactiae subspecies equisimilis isolated from patients with streptococcal toxic shock syndrome. *Journal of Infection and Chemotherapy: Official Journal of the Japan Society of Chemotherapy*, *24*(2), 117–122.

83. Hashikawa, S., Iinuma, Y., Furushita, M., Ohkura, T., Nada, T., Torii, K., Hasegawa, T., & Ohta, M. (2004). Characterization of group C and G streptococcal strains that cause streptococcal toxic shock syndrome. *Journal of Clinical Microbiology*, *42*(1), 186–192.

84. Takahashi, T., Ubukata, K., & Watanabe, H. (2011). Invasive infection caused by *Streptococcus dysgalactiae* subspecies *equisimilis*: Characteristics of strains and clinical features. *Journal of Infection and Chemotherapy: Official Journal of the Japan Society of Chemotherapy*, *17*(1), 1–10.

85. Hagiya, H., Okita, S., Kuroe, Y., Nojima, H., Otani, S., Sugiyama, J., Naito, H., Kawanishi, S., Hagioka, S., & Morimoto, N. (2013). A fatal case of streptococcal toxic shock syndrome due to *Streptococcus dysgalactiae* subspecies *equisimilis* possibly caused by an intramuscular injection. *Internal Medicine*, *52*(3), 397–402.

86. Patti, J. M., Allen, B. L., McGavin, M. J., & Hook, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues. *Annual Review of Microbiology*, 585–617.

87. Dalton, C. J., & Lemmon, C. A. (2021). Fibronectin: Molecular structure, fibrillar structure and mechanochemical signaling. *Cells*, *10*(9), 2443.

88. Joh, D., Speziale, P., Gurusiddappa, S., Manor, J., & Höök, M. (1998). Multiple specificities of the staphylococcal and streptococcal fibronectin-binding microbial surface components recognizing adhesive matrix molecules. *European Journal of Biochemistry*, *258*(2), 897–905.

89. Kline, J. B., Xu, S., Bisno, A. L., & Collins, C. M. (1996). Identification of a fibronectin-binding protein (*Gfb*A) in pathogenic group G streptococci. *Infection and Immunity*, *64*(6), 2122–2129.

90. Kratovac, Z., Manoharan, A., Luo, F., Lizano, S., & Bessen, D. E. (2007). Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. *Journal of Bacteriology*, *189*(4), 1299–1310.

91. Okumura, K., Shimomura, Y., Murayama, S. Y., Yagi, J., Ubukata, K., Kirikae, T., & Miyoshi-Akiyama, T. (2012). Evolutionary paths of streptococcal and staphylococcal superantigens. *BMC Genomics*, *13*(1), 404.

92. Xie, O., Morris, J. M., Hayes, A. J., Towers, R. J., Jespersen, M. G., Lees, J. A., Ben Zakour, N. L., Berking, O., Baines, S. L., Carter, G. P., Tonkin-Hill, G., Schrieber, L., McIntyre, L., Lacey, J. A., James, T. B., Sriprakash, K. S., Beatson, S. A., Hasegawa, T., Giffard, P., … Davies, M. R. (2024). Inter-species gene flow drives ongoing evolution of *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subspecies *equisimilis*. *Nature Communications*, *15*(1), 2286.

93. Spellerberg, B., Rozdzinski, E., Martin, S., Weber-Heynemann, J., Schnitzler, N., Lütticken, R., & Podbielski, A. (1999). Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of *Streptococcus agalactiae* to human laminin. *Infection and Immunity*, *67*(2), 871–878.

94. Franken, C., Haase, G., Brandt, C., Weber-Heynemann, J., Martin, S., Lämmler, C., Podbielski, A., Lütticken, R., & Spellerberg, B. (2001). Horizontal gene transfer and host specificity of beta-haemolytic streptococci: The role of a putative composite transposon containing *scp*B and *lmb*. *Molecular Microbiology*, *41*(4), 925–935.

95. Metzgar, D., & Zampolli, A. (2011). The M protein of group A *Streptococcus* is a key virulence factor and a clinically relevant strain identification marker. *Virulence*, *2*(5), 402–412.

96. Påhlman, L. I., Mörgelin, M., Eckert, J., Johansson, L., Russell, W., Riesbeck, K., Soehnlein, O., Lindbom, L., Norrby-Teglund, A., Schumann, R. R., Björck, L., & Herwald, H. (2006). Streptococcal M protein: A multipotent and powerful inducer of inflammation. *Journal of Immunology (Baltimore, Md.: 1950)*, *177*(2), 1221–1228.

97. McMillan, D. J., Drèze, P.-A., Vu, T., Bessen, D. E., Guglielmini, J., Steer, A. C., Carapetis, J. R., Van Melderen, L., Sriprakash, K. S., & Smeesters, P. R. (2013). Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *19*(5), E222-229.

98. Fischetti, V. A. (1989). Streptococcal M protein: Molecular design and biological behavior. *Clinical Microbiology Reviews*, *2*(3), 285–314.

99. Schnitzler, N., Podbielski, A., Baumgarten, G., Mignon, M., & Kaufhold, A. (1995). M or M-like protein gene polymorphisms in human group G streptococci. *Journal of Clinical Microbiology*, *33*(2), 356–363.

100. Navarre, W. W., & Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiology and Molecular Biology Reviews: MMBR*, *63*(1), 174–229.

101. Khakzad, H., Happonen, L., Karami, Y., Chowdhury, S., Bergdahl, G. E., Nilges, M., Tran Van Nhieu, G., Malmström, J., & Malmström, L. (2021). Structural determination of *Streptococcus pyogenes* M1 protein interactions with human immunoglobulin G using integrative structural biology. *PLoS Computational Biology*, *17*(1), e1008169.

102. Nitsche, D. P., Johansson, H. M., Frick, I.-M., & Mörgelin, M. (2006). Streptococcal protein FOG, a novel matrix adhesin interacting with collagen I *in vivo*. *Journal of Biological Chemistry*, *281*(3), 1670–1679.

103. Lobner, E., Traxlmayr, M. W., Obinger, C., & Hasenhindl, C. (2016). Engineered IgG1-Fc – one fragment to bind them all. *Immunological Reviews*, *270*(1), 113–131.

104. Malke, H. (2019). Genetics and pathogenicity factors of group C and G streptococci. *Microbiology Spectrum*, *7*(2).

105. Castellino, F. J. (1979). A unique enzyme — protein substrate modifier reaction: Plasmin/streptokinase interaction. *Trends in Biochemical Sciences*, *4*(1), 1–5.

106. Lottenberg, R., Minning-Wenz, D., & Boyle, M. D. (1994). Capturing host plasmin(ogen): A common mechanism for invasive pathogens? *Trends in Microbiology*, *2*(1), 20–24.

107. Magassa, N., Chandrasekaran, S., & Caparon, M. G. (2010). *Streptococcus pyogenes* cytolysin-mediated translocation does not require pore formation by streptolysin O. *EMBO Reports*, *11*(5), 400–405.

108. Proft, T., & Fraser, J. D. (2007). Streptococcal superantigens. *Chemical Immunology and Allergy*, *93*, 1–23.

109. Brandt, C. M., Schweizer, K. G., Holland, R., Lütticken, R., & Freyaldenhoven, B. S. (2005). Lack of mitogenic activity of speG- and speG(dys)-positive *Streptococcus dysgalactiae* subspecies *equisimilis* isolates from patients with invasive infections. *International Journal of Medical Microbiology: IJMM*, *295*(8), 539–546.

110. Sachse, S., Seidel, P., Gerlach, D., Günther, E., Rödel, J., Straube, E., & Schmidt, K.-H. (2002). Superantigen-like gene(s) in human pathogenic *Streptococcus dysgalactiae* subspecies *equisimilis*: Genomic localisation of the gene encoding streptococcal pyrogenic exotoxin G (*spe*G(dys)). *FEMS Immunology and Medical Microbiology*, *34*(2), 159–167.

111. Korem, M., Hidalgo-Grass, C., Michael-Gayego, A., Nir-Paz, R., Salameh, S., & Moses, A. E. (2014). Streptococcal pyrogenic exotoxin G gene in blood and pharyngeal isolates of *Streptococcus dysgalactiae* subspecies *equisimilis* has a limited role in pathogenesis. *Journal of Microbiology, Immunology, and Infection = Wei Mian Yu Gan Ran Za Zhi*, *47*(4), 292–296.

112. Pillai, A., Thomas, S., & Williams, C. (2005). Clindamycin in the treatment of group G beta-haemolytic streptococcal infections. *The Journal of Infection*, *51*(4), e207-211.

113. Fuursted, K., Stegger, M., Hoffmann, S., Lambertsen, L., Andersen, P. S., Deleuran, M., & Thomsen, M. K. (2016). Description and characterization of a penicillin-resistant *Streptococcus dysgalactiae* subspecies *equisimilis* clone isolated from blood in three epidemiologically linked patients. *The Journal of Antimicrobial Chemotherapy*, *71*(12), 3376–3380.

114. Biedenbach, D. J., Toleman, M. A., Walsh, T. R., & Jones, R. N. (2006). Characterization of fluoroquinolone-resistant beta-hemolytic *Streptococcus* spp. Isolated in North America and Europe including the first report of fluoroquinolone-resistant *Streptococcus dysgalactiae* subspecies *equisimilis*: Report from the SENTRY Antimicrobial Surveillance Program (1997-2004). *Diagnostic Microbiology and Infectious Disease*, *55*(2), 119–127.

115. Szczuka, E., Makowska, N., Bosacka, K., Słotwińska, A., & Kaznowski, A. (2016). Molecular basis of resistance to macrolides, lincosamides and streptogramins in *Staphylococcus hominis* strains isolated from clinical specimens. *Folia Microbiologica*, *61*(2), 143–147.

116. Berbel, D., González-Díaz, A., López de Egea, G., Càmara, J., & Ardanuy, C. (2022). An overview of macrolide resistance in streptococci: Prevalence, mobile elements and dynamics. *Microorganisms*, *10*(12), 2316.

117. Leclercq, R. (2002). Mechanisms of resistance to macrolides and lincosamides: Nature of the resistance elements and their clinical implications. *Clinical Infectious Diseases*, *34*(4), 482–492.

118. Yin, J., Yu, S., Liu, X., Li, Y., Gao, W., Ma, X., & Yang, Y. (2011). Prevalence and mechanism of resistance to antimicrobial agents in group G streptococcal isolates from China. *Antimicrobial Agents and Chemotherapy*, *55*(1), 402–404.

119. Rojo-Bezares, B., Toca, L., Azcona-Gutiérrez, J. M., Ortega-Unanue, N., Toledano, P., & Sáenz, Y. (2021). *Streptococcus dysgalactiae* subspecies *equisimilis* from invasive and non-invasive infections in Spain: Combining epidemiology, molecular characterization, and genetic diversity. *European Journal of Clinical Microbiology & Infectious Diseases*, *40*(5), 1013–1021.

120. Brenciani, A., Bacciaglia, A., Vignaroli, C., Pugnaloni, A., Varaldo, P. E., & Giovanetti, E. (2010). Phim46.1, the main *Streptococcus pyogenes* element carrying *mef*(A) and *tet*(O) genes. *Antimicrobial Agents and Chemotherapy*, *54*(1), 221–229.

121. Varaldo, P. E., Montanari, M. P., & Giovanetti, E. (2009). Genetic elements responsible for erythromycin resistance in streptococci. *Antimicrobial Agents and Chemotherapy*, *53*(2), 343–353.

122. Berbel, D., Càmara, J., González-Díaz, A., Cubero, M., López de Egea, G., Martí, S., Tubau, F., Domínguez, M. A., & Ardanuy, C. (2021). Deciphering mobile genetic elements disseminating macrolide resistance in *Streptococcus pyogenes* over a 21 year period in Barcelona, Spain. *Journal of Antimicrobial Chemotherapy*, *76*(8), 1991–2003.

123. Ambroset, C., Coluzzi, C., Guédon, G., Devignes, M.-D., Loux, V., Lacroix, T., Payot, S., & Leblond-Bourget, N. (2016). New insights into the classification and integration specificity of *Streptococcus* integrative conjugative elements through extensive genome exploration. *Frontiers in Microbiology*, *6*.

124. Rubio-López, V., Valdezate, S., Álvarez, D., Villalón, P., Medina, M. J., Salcedo, C., & Sáez-Nieto, J.-A. (2012). Molecular epidemiology, antimicrobial susceptibilities and resistance mechanisms of *Streptococcus pyogenes* isolates resistant to erythromycin and tetracycline in Spain (1994–2006). *BMC Microbiology*, *12*, 215.

125. Palmieri, C., Magi, G., Creti, R., Baldassarri, L., Imperi, M., Gherardi, G., & Facinelli, B. (2013). Interspecies mobilization of an *erm*T-carrying plasmid of *Streptococcus dysgalactiae* subspecies *equisimilis* by a coresident ICE of the ICE*Sa2603* family. *The Journal of Antimicrobial Chemotherapy*, *68*(1), 23–26.

126. Xie, O., Zachreson, C., Tonkin-Hill, G., Price, D. J., Lacey, J. A., Morris, J. M., McDonald, M. I., Bowen, A. C., Giffard, P. M., Currie, B. J., Carapetis, J. R., Holt, D. C., Bentley, S. D., Davies, M. R., & Tong, S. Y. C. (2024). Overlapping *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subspecies *equisimilis* household transmission and mobile genetic element exchange. *Nature Communications*, *15*(1), 3477.

127. Mingoia, M., Morici, E., Marini, E., Brenciani, A., Giovanetti, E., & Varaldo, P. E. (2016). Macrolide resistance gene *erm*(TR) and *erm*(TR)-carrying genetic elements in *Streptococcus agalactiae*:

Characterization of ICE*SagTR7*, a new composite element containing IME*Sp2907*. *Journal of Antimicrobial Chemotherapy*, *71*(3), 593–600.

128. Morici, E., Simoni, S., Brenciani, A., Giovanetti, E., Varaldo, P. E., & Mingoia, M. (2017). A new mosaic integrative and conjugative element from *Streptococcus agalactiae* carrying resistance genes for chloramphenicol (*cat*Q) and macrolides [*mef*(I) and *erm*(TR)]. *Journal of Antimicrobial Chemotherapy*, *72*(1), 64–67.

129. Huang, J., Ma, J., Shang, K., Hu, X., Liang, Y., Li, D., Wu, Z., Dai, L., Chen, L., & Wang, L. (2016). Evolution and diversity of the antimicrobial resistance associated mobilome in *Streptococcus suis*: A probable mobile genetic elements reservoir for other streptococci. *Frontiers in Cellular and Infection Microbiology*, *6*.

130. Davies, M. R., Shera, J., Van Domselaar, G. H., Sriprakash, K. S., & McMillan, D. J. (2009). A Novel Integrative conjugative element mediates genetic transfer from group G *Streptococcus* to other β-hemolytic streptococci. *Journal of Bacteriology*, *191*(7), 2257–2265.

131. Shinohara, K., Murase, K., Tsuchido, Y., Noguchi, T., Yukawa, S., Yamamoto, M., Matsumura, Y., Nakagawa, I., & Nagao, M. (2023). Clonal expansion of multidrug-resistant *Streptococcus dysgalactiae* subspecies *equisimilis* causing bacteremia, Japan, 2005–2021. *Emerging Infectious Diseases*, *29*(3), 528–539.

132. Barros, R. R. (2021). Antimicrobial Resistance among Beta-Hemolytic *Streptococcus* in Brazil: An Overview. *Antibiotics*, *10*(8), 973.

133. López de Egea, G., González-Díaz, A., Guédon, G., Lao, J., Berbel, D., Casabella, A., Marimón, J. M., Cercenado, E., Fernández-Delgado, L., Chiapello, H., Lacroix, T., Domínguez, M. Á., Leblond-Bourget, N., & Ardanuy, C. (2023). A new integrative and mobilizable element is a major contributor to tetracycline resistance in *Streptococcus dysgalactiae* subspecies *equisimilis*. *Antibiotics (Basel, Switzerland)*, *12*(3), 579.

134. Liu, L.-C., Tsai, J.-C., Hsueh, P.-R., Tseng, S.-P., Hung, W.-C., Chen, H.-J., & Teng, L.-J. (2008). Identification of tet(S) gene area in tetracycline-resistant *Streptococcus dysgalactiae* subspecies *equisimilis clinical* isolates. *The Journal of Antimicrobial Chemotherapy*, *61*(2), 453–455.

135. Kim, S., Byun, J.-H., Park, H., Lee, J., Lee, H. S., Yoshida, H., Shibayama, A., Fujita, T., Tsuyuki, Y., & Takahashi, T. (2018). Molecular epidemiological features and antibiotic susceptibility patterns of *Streptococcus dysgalactiae* subspecies *equisimilis* isolates from Korea and Japan. *Annals of Laboratory Medicine*, *38*(3), 212–219.

136. Karyński, M., Łętowska, I., & Grzesiowski, P. (2012). Soft tissue infection caused by *Streptococcus dysgalactiae* subspecies *equisimilis* possessing group A antigen: A case report and review of the literature. *Advances in Dermatology and Allergology/Postępy Dermatologii i Alergologii*, *29*(4), 330–336.

137. Clewell, D. B., Flannagan, S. E., Jaworski, D. D., & Clewell, D. B. (1995). Unconstrained bacterial promiscuity: The Tn*916*–Tn*1545* family of conjugative transposons. *Trends in Microbiology*, *3*(6), 229–236.

138. Courvalin, P., & Carlier, C. (1987). Tn1545: A conjugative shuttle transposon. *Molecular and General Genetics MGG*, *206*(2), 259–264.

139. Pinho, M. D., Melo-Cristino, J., & Ramirez, M. (2010). Fluoroquinolone resistance in *Streptococcus dysgalactiae* subspecies *equisimilis* and evidence for a shared global gene pool with *Streptococcus pyogenes*. *Antimicrobial Agents and Chemotherapy*, *54*(5), 1769–1777.

140. Hooper, D. C. (2002). Fluoroquinolone resistance among Gram-positive cocci. *The Lancet. Infectious Diseases*, *2*(9), 530–538.

141. Aldred, K. J., Kerns, R. J., & Osheroff, N. (2014). Mechanism of Quinolone Action and Resistance. *Biochemistry*, *53*(10), 1565–1574.

142. Davies, J., & Wright, G. D. (1997). Bacterial resistance to aminoglycoside antibiotics. *Trends in Microbiology*, *5*(6), 234–240.

143. Galimand, M., Lambert, T., Gerbaud, G., & Courvalin, P. (1999). High-level aminoglycoside resistance in the beta-hemolytic group G *Streptococcus* isolate BM2721. *Antimicrobial Agents and Chemotherapy*, *43*(12), 3008–3010.

144. Breakpoint tables for interpretation of MICs and zone diameters. (2023). *The European Committee on Antimicrobial Susceptibility Testing*, *Version 13.1.* http://www.eucast.org

145. M100—Performance Standards for Antimicrobial Susceptibility Testing. (2023). *Clinical and Laboratory Standards Institute*, *33*.

146. Kimura, K., Wachino, J.-I., Kurokawa, H., Suzuki, S., Yamane, K., Shibata, N., & Arakawa, Y. (2009). Practical disk diffusion test for detecting group B streptococcus with reduced penicillin susceptibility. *Journal of Clinical Microbiology*, *47*(12), 4154–4157.

147. Figueira-Coelho, J., Ramirez, M., Salgado, M. j., & Melo-Cristino, J. (2004). *Streptococcus agalactiae* in a large Portuguese teaching hospital: Antimicrobial susceptibility, serotype distribution, and clonal analysis of macrolide-resistant isolates. *Microbial Drug Resistance*, *10*(1), 31–36.

148. Compain, F., Hays, C., Touak, G., Dmytruk, N., Trieu-Cuot, P., Joubrel, C., & Poyart, C. (2014). Molecular characterization of *Streptococcus agalactiae* isolates harboring small *erm*(T)-carrying plasmids. *Antimicrobial Agents and Chemotherapy*, *58*(11), 6928–6930.

149. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120.

150. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *19*(5), 455–477.

151. Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PloS One*, *9*(11), e112963.

152. Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069.

153. Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nature Methods*, *8*(10), 785–786.

154. Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, *35*(9), 3100–3108.

155. Clancy, J., Petitpas, J., Dib-Hajj, F., Yuan, W., Cronan, M., Kamath, A. V., Bergeron, J., & Retsema, J. A. (1996). Molecular cloning and functional analysis of a novel macrolide-resistance determinant, *mef*A, from *Streptococcus pyogenes*. *Molecular Microbiology*, *22*(5), 867–879.

156. Cai, Y., Kong, F., & Gilbert, G. L. (2007). Three new macrolide efflux (*mef*) gene variants in *Streptococcus agalactiae*. *Journal of Clinical Microbiology*, *45*(8), 2754–2755.

157. Tait-Kamradt, A., Clancy, J., Cronan, M., Dib-Hajj, F., Wondrack, L., Yuan, W., & Sutcliffe, J. (1997). *Mef*E is necessary for the erythromycin-resistant M phenotype in *Streptococcus pneumoniae*. *Antimicrobial Agents and Chemotherapy*, *41*(10), 2251–2255.

158. Amezaga, M. R., & McKenzie, H. (2006). Molecular epidemiology of macrolide resistance in β-haemolytic streptococci of Lancefield groups A, B, C and G and evidence for a new mef element in group G streptococci that carries allelic variants of *mef* and *msr*(D). *Journal of Antimicrobial Chemotherapy*, *57*(3), 443–449.

159. Cochetti, I., Vecchi, M., Mingoia, M., Tili, E., Catania, M. R., Manzin, A., Varaldo, P. E., & Montanari, M. P. (2005). Molecular characterization of pneumococci with efflux-mediated erythromycin resistance and identification of a novel *mef* gene subclass, *mef*(I). *Antimicrobial Agents and Chemotherapy*, *49*(12), 4999–5006.

160. Sangvik, M., Littauer, P., Simonsen, G. S., Sundsfjord, A., & Dahl, K. H. (2005). *mef*(A), *mef*(E) and a new *mef* allele in macrolide-resistant *Streptococcus* spp. Isolates from Norway. *The Journal of Antimicrobial Chemotherapy*, *56*(5), 841–846.

161. Petinaki, E., Guérin-Faublée, V., Pichereau, V., Villers, C., Achard, A., Malbruny, B., & Leclercq, R. (2008). Lincomycin Resistance Gene *lnu*(D) in *Streptococcus uberis*. *Antimicrobial Agents and Chemotherapy*, *52*(2), 626–630.

162. Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., Tai, C., Deng, Z., & Ou, H.-Y. (2019). ICEberg 2.0: An updated database of bacterial integrative and conjugative elements. *Nucleic Acids Research*, *47*(D1), D660–D665.

163. Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: A genome comparison visualizer. *Bioinformatics (Oxford, England)*, *27*(7), 1009–1010.

164. Hunter, P. R., & Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. *Journal of Clinical Microbiology*, *26*(11), 2465–2466.

165. Malbruny, B., Werno, A. M., Murdoch, D. R., Leclercq, R., & Cattoir, V. (2011). Cross-resistance to lincosamides, streptogramins A, and pleuromutilins due to the *lsa*(C) gene in *Streptococcus agalactiae* UCN70. *Antimicrobial Agents and Chemotherapy*, *55*(4), 1470–1474.

166. Khan, U. B., Portal, E. A. R., Sands, K., Lo, S., Chalker, V. J., Jauneikaite, E., & Spiller, O. B. (2023). Genomic analysis reveals new integrative conjugal elements and transposons in GBS conferring antimicrobial resistance. *Antibiotics*, *12*(3), Article 3.

167. Yao, T.-G., Li, B.-Y., Luan, R.-D., Wang, H.-N., & Lei, C.-W. (2020). Whole genome sequence of *Enterococcus gallinarum* EG81, a porcine strain harbouring the oxazolidinone-phenicol resistance gene optrA with chromosomal and plasmid location. *Journal of Global Antimicrobial Resistance*, *22*, 598–600.

168. Del Grosso, M., Camilli, R., Barbabella, G., Blackman Northwood, J., Farrell, D. J., & Pantosti, A. (2011). Genetic resistance elements carrying *mef* subclasses other than *mef*(A) in *Streptococcus pyogenes*. *Antimicrobial Agents and Chemotherapy*, *55*(7), 3226–3230.

169. Mingoia, M., Morici, E., Brenciani, A., Giovanetti, E., & Varaldo, P. E. (2015). Genetic basis of the association of resistance genes *mef*(I) (macrolides) and *cat*Q (chloramphenicol) in streptococci. *Frontiers in Microbiology*, *5*.

170. Paspaliari, D. K., Sarvikivi, E., Ollgren, J., & Vuopio, J. (2023). Invasive beta-haemolytic streptococcal infections, Finland, 2006 to 2020: Increase in Lancefield group C/G infections. *Eurosurveillance*, *28*(31).

171. Gajdács, M., Ábrók, M., Lázár, A., & Burián, K. (2020). Beta-haemolytic group A, C and G streptococcal infections in Southern Hungary: A 10-Year population-based retrospective survey (2008-2017) and a review of the literature. *Infection and Drug Resistance*, *Volume 13*, 4739–4749.

172. Takahashi, T., Sunaoshi, K., Sunakawa, K., Fujishima, S., Watanabe, H., & Ubukata, K. (2010). Clinical aspects of invasive infections with *Streptococcus dysgalactiae* ssp. *equisimilis* in Japan: Differences with respect to *Streptococcus pyogenes* and *Streptococcus agalactiae* infections. *Clinical Microbiology and Infection*, *16*(8), 1097–1103.

173. Broyles, L. N., Van Beneden, C., Beall, B., Facklam, R., Shewmaker, P. L., Malpiedi, P., Daily, P., Reingold, A., & Farley, M. M. (2009). Population-based study of invasive disease due to β-Hemolytic streptococci of groups other than A and B. *Clinical Infectious Diseases*, *48*(6), 706–712.

174. Pinho, M. D., Melo-Cristino, J., & Ramirez, M. (2006). Clonal Relationships between Invasive and noninvasive Lancefield group C and G streptococci and *emm*-specific differences in invasiveness. *Journal of Clinical Microbiology*, *44*(3), 841–846.

175. Lopardo, H. A., Vidal, P., Sparo, M., Jeric, P., Centron, D., Facklam, R. R., Paganini, H., Pagniez, N. G., Lovgren, M., & Beall, B. (2005). Six-month multicenter study on invasive infections due to *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subspecies *equisimilis* in Argentina. *Journal of Clinical Microbiology*, *43*(2), 802–807.

176. Kittang, B. R., Bruun, T., Langeland, N., Mylvaganam, H., Glambek, M., & Skrede, S. (2011). Invasive group A, C and G streptococcal disease in western Norway: Virulence gene profiles, clinical features and outcomes. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *17*(3), 358–364.

177. Leitner, E., Zollner-Schwetz, I., Zarfel, G., Masoud-Landgraf, L., Gehrer, M., Wagner-Eibel, U., Grisold, A. J., & Feierl, G. (2015). Prevalence of emm types and antimicrobial susceptibility of *Streptococcus dysgalactiae* subspecies *equisimilis* in Austria. *International Journal of Medical Microbiology: IJMM*, *305*(8), 918–924.

178. Traverso, F., Blanco, A., Villalón, P., Beratz, N., Sáez Nieto, J. A., Lopardo, H., & National Collaborative Group for the Study of Streptococci and Related Bacteria. (2016). Molecular characterization of invasive *Streptococcus dysgalactiae* subspecies *equisimilis*. Multicenter study: Argentina 2011-2012. *Revista Argentina De Microbiologia*, *48*(4), 279–289.

179. de Souza, J. P., Santos, A. R., de Paula, G. R., & Barros, R. R. (2016). Antimicrobial susceptibility and genetic relationships among *Streptococcus dysgalactiae* subspecies *equisimilis* isolates in Rio de Janeiro. *Infectious Diseases (London, England)*, *48*(9), 676–681.

180. DiPersio, L. P., DiPersio, J. R., Frey, K. C., & Beach, J. A. (2008). Prevalence of the *erm*(T) gene in clinical isolates of erythromycin-resistant group D *Streptococcus* and *Enterococcus*. *Antimicrobial Agents and Chemotherapy*, *52*(4), 1567–1569.

181. Woo, P. C. Y., To, A. P. C., Tse, H., Lau, S. K. P., & Yuen, K. (2003). Clinical and molecular epidemiology of erythromycin-resistant beta-hemolytic Lancefield group G streptococci sausing bacteremia. *Journal of Clinical Microbiology*, *41*(11), 5188–5191.

182. Kaci, A., Jonassen, C. M., Skrede, S., Sivertsen, A., The Norwegian Study Group on Streptococcus dysgalactiae, Steinbakk, M., Oppegaard, O., Fostervold, A., Jakovljev, A., Onken, A., Marvik, Å., Einar, N., Müller, F., Milad Chedid, G., Skov Simonsen, G., Zaragkoulias, K., Hjetland, R., Magne Baevre-Jensen, R., Åsheim, S., … Michael Leegaard, T. (2023). Genomic epidemiology of *Streptococcus dysgalactiae* subspecies *equisimilis* strains causing invasive disease in Norway during 2018. *Frontiers in Microbiology*, *14*.

183. Itzek, A., Weißbach, V., Meintrup, D., Rieß, B., Van Der Linden, M., & Borgmann, S. (2023). Epidemiological and clinical features of *Streptococcus dysgalactiae* ssp. *Equisimilis* stG62647 and other *emm* types in Germany. *Pathogens*, *12*(4), 589.

184. Oppegaard, O., Skrede, S., Mylvaganam, H., & Kittang, B. R. (2020). Emerging threat of antimicrobial resistance in β-hemolytic streptococci. *Frontiers in Microbiology*, *11*, 797.

185. Cochetti, I., Tili, E., Mingoia, M., Varaldo, P. E., & Montanari, M. P. (2008). *Erm*(B)-carrying elements in tetracycline-resistant pneumococci and correspondence between Tn*1545* and Tn*6003*. *Antimicrobial Agents and Chemotherapy*, *52*(4), 1285–1290.

186. Yi, S., Zhou, K., & Xu, X. (2024). Characterization of *erm*(B)-carrying Integrative and Conjugative Elements transferred from *Streptococcus anginosus* to other streptococci and enterococci. *Microbial Drug Resistance (Larchmont, N.Y.)*, *30*(6), 243–253.

187. DiPersio, L. P., DiPersio, J. R., Beach, J. A., Loudon, A. M., & Fuchs, A. M. (2011). Identification and characterization of plasmid-borne *erm*(T) macrolide resistance in group B and group A *Streptococcus*. *Diagnostic Microbiology and Infectious Disease*, *71*(3), 217–223.

188. Bley, C., van der Linden, M., & Reinert, R. R. (2011). *Mef*(A) is the predominant macrolide resistance determinant in *Streptococcus pneumoniae* and *Streptococcus pyogenes* in Germany. *International Journal of Antimicrobial Agents*, *37*(5), 425–431.

189. Gherardi, G., Imperi, M., Palmieri, C., Magi, G., Facinelli, B., Baldassarri, L., Pataracchia, M., & Creti, R. (2014). Genetic diversity and virulence properties of *Streptococcus dysgalactiae* subspecies *equisimilis* from different sources. *Journal of Medical Microbiology*, *63*(Pt 1), 90–98.

8. Supplementary material

Supplementary table 1 - List of participating hospital laboratories (2019-2022).

Centro Hospitalar Baixo Vouga
Centro Hospitalar de Cascais
Centro Hospitalar de Entre o Douro e Vouga - Hospital São Sebastião
Centro Hospitalar de Leiria - Hospital de Santo André
Centro Hospitalar e Universitário de Coimbra
Centro Hospitalar e Universitário do Algarve
Centro Hospitalar Gaia Espinho
Centro Hospitalar Lisboa Central
Centro Hospitalar Lisboa Norte
Centro Hospitalar Lisboa Ocidental
Centro Hospitalar Tâmega e Sousa
Centro Hospitalar Universitário de Santo António
Hospital Beatriz Ângelo - Loures
Hospital da Luz - Lisboa
Hospital de Aveiro
Hospital de Braga
Hospital Distrital da Figueira da Foz
Hospital do Espírito Santo - Évora
Hospital dos Lusíadas
Hospital Garcia de Orta - Almada
Laboratório José Júlio de Castro Fernandes
Synlab
Unidade Local de Saúde de Matosinhos - Hospital Pedro Hispano

Supplementary table 2 - Characterization of the elements found in the contigs containing the *erm*(B) gene (**Figure 18**).

Isolate	ST	MLST CC	*emm* type	Element sequence length (Kb)	SDSE insertion region	Element	Other resistance genes in element	Figure
SH13288	8	8	*stG643*	92.3	42	Full	*tet*(M)	(A)
SH6081	8	8	*stG11*	76.9	42	Full	*tet*(M)	(B)
SH9723	8	8	*stG11*	76.9	42	Full	*tet*(M)	(B)
SH8223	8	8	*stG11*	48.7	42	Incomplete at 3' end	*tet*(M)	(B)
SH12320	17	17	*stG485*	22.8	Undetermined	Incomplete at both ends	*tet(M)*	(B)
SH12723	17	17	*stG2078*	61.8	5	Full	.	(C)
SH7689	17	17	*stG2078*	74.4	5	Incomplete at 5' end	.	(C)
SH14853	34	34	*stGM220*	50.5	22	Full	*tet*(O)	(D)
SH8530	20	20	*stG62647*	74.7	22	Full	.	(E)
SH15607	15	15	*stG166b*	38.4	22	Incomplete at 3' end	.	(E)
SH6409	138	138	*stG652*	68.2	33	Incomplete at 5' end	*tet*(M)	(E)
SH8633	249	Singleton (ST249)	*stL1376*	71.9	33	Full	*tet*(O)	(F)
SH11911	499	Singleton (ST499)	*stG6*	19.6	33	Incomplete at 5' end	*tet*(M)	(F)
SH14293	499	Singleton (ST499)	*stG6*	19.6	33	Incomplete at 5' end	*tet*(M)	(F)
SH14515	269	269	*stG480*	23.3	Novel	Full	*tet*(M)	(G)

Supplementary table 3 - Characterization of the elements found in the contigs containing the *mef*(A) gene (**Figure 21**).

Isolate	ST	MLST CC	*emm* type	Element sequence length (Kb)	SDSE insertion region	Element	Other resistance genes in element	Figure
SH6079	15	15	*stG10*	113.2	33	Incomplete at 5' end	.	(A)
SH8633	249	Singleton (ST249)	*stL1376*	86.9	22	Full	*vat* (D) *lnu* (D)	(B)
SH16552	35	34	*stC7934*	63.7	33	Full	*tet* (M) *cat* (Q)	(C)
SH17123	35	34	*stCK401*	63.7	33	Incomplete at 5' end	*tet* (M) *cat* (Q)	(D)