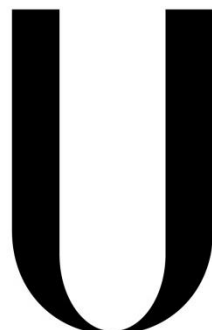


**UNIVERSIDADE DE LISBOA**

**FACULDADE DE MEDICINA**



**LISBOA**

---

UNIVERSIDADE  
DE LISBOA

**Caraterização genómica de *Streptococcus dysgalactiae*  
subsp. *equisimilis* responsável por infeção invasiva no Homem**

ANA CAROLINA SANTOS RODRIGUES DE CASTRO

Orientador: Prof. Doutor Marcos Daniel Caetano Borges de Pinho

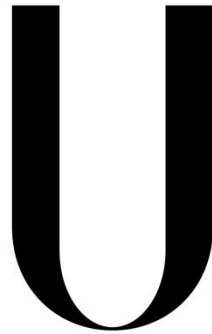
Dissertação especialmente elaborada para obtenção do grau de mestre em  
Microbiologia Clínica e Doenças Infeciosas Emergentes

2020



**UNIVERSIDADE DE LISBOA**

**FACULDADE DE MEDICINA**



**LISBOA**

---

UNIVERSIDADE  
DE LISBOA

**Caraterização genómica de *Streptococcus dysgalactiae*  
subsp. *equisimilis* responsável por infeção invasiva no Homem**

ANA CAROLINA SANTOS RODRIGUES DE CASTRO

Orientador: Prof. Doutor Marcos Daniel Caetano Borges de Pinho

Dissertação especialmente elaborada para obtenção do grau de mestre em  
Microbiologia Clínica e Doenças Infeciosas Emergentes

2020



**A impressão desta dissertação foi aprovada pelo Conselho Científico  
da Faculdade de Medicina de Lisboa em reunião de 30/06/2020.**



## AGRADECIMENTOS

---

Ao Prof. Doutor Mário Ramirez, agradeço a oportunidade de integração num projeto na sua equipa de investigação e por todos os conhecimentos transmitidos.

Ao Prof. Doutor Marcos Pinho, o meu orientador neste trabalho, agradeço a supervisão e toda a disponibilidade, atenção, confiança e partilha.

A todas as pessoas do laboratório MRamirez, em especial à Joana e à Soraia, agradeço o apoio, esclarecimento e a paciência em todos os momentos, à Inês Mendes, um obrigada pela constante disponibilidade e ajuda, e ao Bruno, agradeço todas as fotografias tiradas. Aos restantes colegas da sala “Bacteria lovers”, um obrigada pelos risos, momentos descontração e ajuda.

Às minhas amigas, colegas, companheiras e confidentes, Beatriz e Inês, obrigada pela amizade incondicional e que levo para a vida, sem vocês nada teria sido igual.

À minha família, em especial, aos meus pais, ao meu irmão, às minhas primas Rita e Nê e ao meu tio, por me proporcionarem, apoiarem e incentivarem, cada um à sua maneira, a seguir os meus sonhos e a concretizar mais um objetivo. Um obrigada do tamanho do mundo. A vocês, devo tudo. Agradeço também aos que, mesmo já não estando cá, continuam a olhar por mim e são a minha estrela-guia e inspiração.

Ao Rui, agradeço o amor, carinho, compreensão e apoio incondicionais, e a força e confiança transmitidas em todos os momentos. Nem com todas as palavras do mundo conseguirei expressar a minha gratidão, por tudo.

Do fundo do coração, obrigada.



*Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) é uma espécie de estreptococos beta-hemolíticos, dos grupos C ou G de Lancefield, de importância crescente como agente patogénico humano, principalmente em doentes idosos e com comorbidades. SDSE é agente etiológico de várias infeções da pele e tecidos moles, faringo-amigdalite e infeções mais graves, como bacteriemia, fascíte necrotizante ou a síndrome do choque tóxico estreptocócico (STSS).

O objetivo deste estudo foi a definição das linhagens clonais de SDSE predominantes em infeção invasiva em Portugal, com ênfase na identificação de fatores que possam influenciar a prevalência e a virulência dessas linhagens. Neste sentido, uma coleção de 316 estirpes SDSE de infeções invasivas no Homem, isoladas entre 2011 a 2017 em Portugal, foram caracterizadas através de métodos fenotípicos – determinação do grupo de Lancefield e suscetibilidade a antimicrobianos – e genotípicos – tipagem do gene *emm* e sequenciação de alto débito. Esta última foi usada para determinar perfis alélicos de “multilocus sequence type” (MLST) convencional e “core genome” (cgMLST), fatores de virulência e determinantes de resistência.

As infeções invasivas por SDSE em Portugal afetaram maioritariamente doentes idosos (média de 69,5 anos) do sexo masculino (58,2%). A maioria das estirpes pertencia ao grupo G de Lancefield (n=226, 71,5%), seguido dos grupos C 89 (n=89, 28,2%) e L (n=1, 0,3%). A tipagem do gene *emm* e a análise por MLST evidenciaram uma elevada diversidade genética das estirpes, com 25 tipos *emm* e 76 tipos de sequência (STs) definidos por MLST. Os STs foram agrupados em 16 complexos clonais (CCs), dos quais CC17 (n=101, 32,0%), CC20 (n=78, 24,7%) e CC29 (n=56, 17,7%) se destacaram como os mais frequentes. Enquanto os CC17 e CC29 incluíam estirpes do grupo G de Lancefield com vários tipos *emm*, o complexo clonal CC20 era composto por estirpes exclusivamente do grupo C de Lancefield e maioritariamente do tipo *emm* *stG62647*, o mais frequente (26,6%) do estudo. A análise por cgMLST confirmou estes três CCs como as três principais linhagens clonais entre as estirpes de SDSE responsáveis por infeção invasiva em Portugal.

A pesquisa de genes associados à virulência nos genomas de SDSE mostrou que os genes *emm*, *fbp54*, *hasC*, *gapC* e o operão da SLS estavam presentes em todas as estirpes e os genes *lmb*, *ska*, *slo* e *nga* na grande maioria. O gene *speG* foi o único gene de exotoxinas pirogênicas detetado, estando presente em 177 (56,0%) estirpes e apresentando uma distribuição clonal. Entre os sistemas reguladores identificados, *covRS* e *fasCAX* estavam presentes em todas as estirpes, enquanto o *sil* tinha uma presença variável, estando ausente nos dois principais CCs identificados, CC17 e CC20.

As estirpes invasivas de SDSE apresentaram resistência à eritromicina (30,1%), tetraciclina (16,1%), clindamicina (8,2%), levofloxacina (3,8%) e aminoglicosídeos (3,2%). A maioria das estirpes resistentes à eritromicina apresentava um fenótipo iMLS<sub>B</sub> conferido pelo gene *ermA* e estava associada às estirpes do CC17 (p<0,001). Observou-se uma redução da taxa de resistência à tetraciclina ao longo do tempo (p=0,01), a qual foi maioritariamente conferida pelo gene *tetM* e associada a CC15 (p<0,001),

A análise genómica das estirpes de SDSE demonstrou uma diversidade genética considerável das estirpes causadoras de infeção invasiva em Portugal, em que um número limitado de linhagens clonais altamente prevalentes na população se destacam. Diferenças na distribuição de genes associados à virulência e nos reguladores desses genes foram detetados entre linhagens clonais. Entre as principais linhagens clonais identificadas, algumas apresentavam resistências expressivas aos antimicrobianos, justificando a sua monitorização futura.

**Palavras-chave:** *Streptococcus dysgalactiae* subsp. *equisimilis*; infeção invasiva; sequenciação de alto débito; tipagem; antimicrobianos.

*Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) is a species of beta-haemolytic streptococci belonging to Lancefield groups C or G of increasing importance as a human pathogen, especially in elderly patients and in those with comorbidities. SDSE is the etiologic agent of various skin and soft tissue infections, pharyngo-tonsillitis and invasive infections, such as bacteraemia, necrotizing fasciitis or streptococcal toxic shock syndrome.

The aim of the current study was to define the main clonal lineages of SDSE causing invasive infections in Portugal, and to identify the factors that may influence the prevalence and virulence of those lineages. For this purpose, a collection of 316 SDSE invasive isolates from human infections, isolated between 2011 and 2017 in Portugal, were characterized by phenotypic – Lancefield group determination and antimicrobial susceptibility testing - and genotypic – *emm* typing and high-throughput sequencing – methods. The latter was used to determine “multilocus sequence type” (MLST) and “core genome” (cgMLST) allelic profiles, virulence factors and resistance determinants.

Invasive SDSE infections in Portugal affected mostly elderly (average of 69.5 years) male (58.2%) patients. Most isolates belonged to Lancefield group G (n=226, 71.5%), followed by groups C 89 (n=89, 28.2%) and L (n=1, 0.3%). *emm* typing and MLST showed a high genetic diversity of the population, with 25 *emm* types and 76 sequence types (STs) defined by MLST being found. STs were grouped into 16 clonal complexes (CCs), of which CC17 (n = 101, 32.0%), CC20 (n = 78, 24.7%) and CC29 (n = 56, 17.7%) were the most frequent. While clonal complexes CC17 and CC29 included Lancefield group G isolates of several *emm* types, CC20 only included group C isolates, mostly with *emm* type *stG62647*, the most frequent (26.6%) of the study. Analysis by cgMLST confirmed these three CCs as the three main clonal lineages among SDSE causing invasive infections in Portugal.

The search for genes associated with virulence in the SDSE genomes showed that *emm*, *fbp54*, *hasC*, *gapC* genes and SLS operon were present in all isolates, while almost all possessed *lmb*, *ska*, *slo* and *nga* genes. *speG* gene was the only among pyrogenic exotoxin genes to be detected, being present in 177 (56.0%) isolates and showed a clonal

distribution. Among the regulatory systems identified, *covRS* and *fasCAX* were present in all strains, while *sil* had a variable presence, being absent from CC17 and CC20, the two main CCs identified.

Invasive SDSE isolates showed resistance to erythromycin (30.1%), tetracycline (16.1%), clindamycin (8.2%), levofloxacin (3.8%) and aminoglycosides (3.2%). Most isolates resistant to erythromycin had an iMLS<sub>B</sub> phenotype, conferred by an *ermA* gene, and were associated with CC17 ( $p < 0.001$ ). Resistance to tetracycline, mostly conferred by the *tetM* gene and associated with CC15 ( $p < 0,001$ ), decreased over time ( $p = 0,01$ ).

Genomic analysis of SDSE isolates highlighted the considerable genetic diversity of the isolates causing invasive infection in Portugal. Nevertheless, a limited number of highly prevalent clonal lineages in the population were found. Clonal lineages showed differences in the distribution of virulence associated genes and in the regulators of these genes. Relevant resistance to antimicrobials was found among the main clonal lineages identified, justifying their future monitoring.

**Keywords:** *Streptococcus dysgalactiae* subsp. *equisimilis*; invasive infection; high throughput sequencing; typing; antimicrobials.

<b>AGRADECIMENTOS</b> .....	<b>i</b>
<b>RESUMO</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>v</b>
<b>ÍNDICE</b> .....	<b>vii</b>
<b>LISTA DE FIGURAS</b> .....	<b>ix</b>
<b>LISTA DE TABELAS</b> .....	<b>xi</b>
<b>LISTA DE ABREVIATURAS</b> .....	<b>xiii</b>
<b>1. INTRODUÇÃO</b> .....	<b>1</b>
1.1 <i>Streptococcus dysgalactiae</i> subsp. <i>equisimilis</i> (SDSE).....	1
1.2 Características morfológicas e identificação laboratorial .....	3
1.3 Reservatórios de SDSE e origem da infecção .....	5
1.4 Principais infecções .....	6
1.5 Características clínicas e epidemiológicas das infecções por SDSE.....	7
1.6 Suscetibilidade a antimicrobianos e tratamento das infecções .....	8
1.7 Resistência a antimicrobianos.....	9
1.8 Fatores de virulência .....	12
1.8.1 Proteína M.....	14
1.8.2 Superantigénios .....	16
1.9 Métodos de tipagem.....	17
1.9.1 Tipagem do gene <i>emm</i> .....	18
1.9.2 “Multilocus sequence typing” (MLST).....	18
1.9.3 “Whole genome sequencing” (WGS) .....	21
<b>2. OBJETIVOS</b> .....	<b>23</b>
<b>3. MATERIAIS E MÉTODOS</b> .....	<b>25</b>

---

3.1	Estirpes bacterianas.....	25
3.2	Condições de crescimento e morfologia das colónias .....	25
3.3	Classificação de Lancefield .....	26
3.4	Teste de suscetibilidade aos antimicrobianos .....	26
3.5	Tipagem do gene <i>emm</i> .....	27
3.6	Sequenciação genómica de alto débito .....	28
3.6.1	Preparação do ADN genómico.....	28
3.6.2	Sequenciação, montagem e anotação dos genomas .....	29
3.6.3	Análise gene-por-gene e relações genéticas entre as estirpes .....	30
3.6.4	Determinantes de resistência e fatores de virulência .....	31
3.7	Análise estatística .....	31
<b>4.</b>	<b>RESULTADOS.....</b>	<b>33</b>
4.1	Caraterísticas das estirpes e dos doentes.....	33
4.2	Tipagem do gene <i>emm</i> .....	34
4.3	“Multilocus sequence typing” (MLST) .....	36
4.4	Correlação entre tipo <i>emm</i> , STs e CCs .....	40
4.5	cgMLST .....	43
4.6	Fatores de virulência .....	46
4.7	Teste de suscetibilidade a antimicrobianos.....	49
<b>5.</b>	<b>DISCUSSÃO.....</b>	<b>53</b>
5.1.	Caraterização molecular.....	54
5.2.	Fatores de virulência .....	58
5.3.	Suscetibilidade a antimicrobianos.....	60
<b>6.</b>	<b>CONCLUSÃO .....</b>	<b>65</b>
6.1	Perspetivas futuras.....	66
<b>7.</b>	<b>REFERÊNCIAS .....</b>	<b>67</b>
<b>8.</b>	<b>ANEXOS .....</b>	<b>87</b>

---

## LISTA DE FIGURAS

---

<b>Figura 1:</b> Morfologia das colónias e padrão de hemólise.....	3
<b>Figura 2:</b> Fenótipos de resistência aos macrólidos e lincosamidas. ....	10
<b>Figura 3:</b> Esquema da estrutura da proteína M de GAS.....	15
<b>Figura 4:</b> Exemplo de diagrama utilizando o algoritmo goeBURST.....	20
<b>Figura 5:</b> Distribuição das estirpes de SDSE isoladas de infeção invasiva ao longo do período de estudo (2011-2017).....	33
<b>Figura 6:</b> Distribuição de tipos <i>emm</i> e grupos de Lancefield das 316 estirpes invasivas de SDSE isoladas em Portugal (2011-2017). ....	35
<b>Figura 7:</b> Distribuição de tipos <i>emm</i> ao longo do período de estudo (2011-2017).....	36
<b>Figura 8:</b> Distribuição dos principais CCs ao longo do período de estudo (2011-2017). .....	37
<b>Figura 9:</b> Diagrama goeBURST das relações entre as 316 estirpes invasivas de SDSE com diferentes grupos de Lancefield.....	38
<b>Figura 10:</b> Diagrama goeBURST das relações entre as 316 estirpes invasivas de SDSE com diferentes tipos <i>emm</i> . ....	42
<b>Figura 11:</b> Árvore de extensão mínima das 316 estirpes invasivas isoladas em Portugal com base no esquema cgMLST.....	45
<b>Figura 12:</b> Evolução das taxas de resistência aos antimicrobianos (2011-2017).....	50

### ANEXOS

<b>Figura A1:</b> Árvore de extensão mínima de 1775 estirpes de <i>S. dysgalactiae</i> isoladas de diferentes hospedeiros com base no esquema MLST.....	89
<b>Figura A2:</b> Árvore de extensão mínima de 456 estirpes invasivas isoladas de diferentes países com base no esquema cgMLST. ....	90



## LISTA DE TABELAS

---

<b>Tabela 1:</b> Nomenclatura das estirpes atualmente incluídas no taxon <i>S. dysgalactiae</i> .....	2
<b>Tabela 2:</b> Distribuição de tipos <i>emm</i> pelos CC definidos por MLST. ....	39
<b>Tabela 3:</b> Diferenças alélicas por cgMLST. ....	43
<b>Tabela 4:</b> Genes de virulência das 316 estirpes.....	47
<b>Tabela 5:</b> Estirpes resistentes e intermédias aos antimicrobianos.....	49
<b>Tabela 6:</b> Alterações nos aminoácidos de GyrA e ParC de estirpes de SDE não suscetíveis levofloxacina. ....	51

### ANEXOS

<b>Tabela A1:</b> Lista de laboratórios hospitalares. ....	87
<b>Tabela A2:</b> Diferenças entre CCs. ....	88



## LISTA DE ABREVIATURAS

---

**ADN** – Ácido desoxirribonucleico

**ARN** – Ácido ribonucleico

**ATCC** – American Type Culture Collection

**AW** – Coeficiente ajustado de Wallace (“adjusted Wallace coefficient”)

**CARD** – “Comprehensive Antibiotic Resistance Database”

**CBM** – Concentração bactericida mínima

**CC(s)** – Complexo(s) clonal(is)

**CDC** – “Centers for Disease Control and Prevention”

**cgMLST** – “core genome MLST”

**CIM** – Concentração inibitória mínima

**CLSI** – “Clinical & Laboratory Standards Institute”

**cMLS<sub>B</sub>** – Fenótipo constitutivo de resistência a macrólidos, lincosamidas e estreptograminas B

**DLVs** – “Double *locus* variants”

**EGCG** – Estreptococos pertencentes aos grupos C e G de Lancefield formadores de colônias grandes (>0,5 mm de diâmetro)

**FDR** – Taxa de falsos positivos (“false discovery rate”)

**IC95%** – Intervalo de confiança de 95%

**iMLS<sub>B</sub>** – Fenótipo induzido de resistência a macrólidos, lincosamidas e estreptograminas B

**MALDI-TOF MS** – “Matrix assisted laser desorption ionization-time of flight mass spectrometry”

**MLST** – Tipagem de sequência multilocus (“multilocus sequence type”)

**NCBI** – “National Center for Biotechnology Information”

**OR** – “odds ratio”

**ORFs** – “Open reading frames”

**PBPs** – Proteínas de ligação à penicilina (“penicillin binding proteins”)

**PCR** – Reação em cadeia de polimerase (“polymerase chain reaction”)

**PFGE** – Eletroforese em gel de campo pulsado (“pulsed-field gel electrophoresis”)

**QRDRs** – Regiões determinantes da resistência às quinolonas (“quinolone resistance-determining regions”)

**SDSE** – *Streptococcus dysgalactiae* subsp. *equisimilis*

**SID** – Índice da diversidade de Simpson (“Simpson’s index diversity”)

***sil*** – “streptococcal invasive locus”

**SLO** – Estreptolisina O (“streptolysin O”)

**SLS** – Estreptolisina S (“streptolysin S”)

**SLVs** – “Single *locus* variants”

**ST** – Tipo de sequência (“sequence type”)

**STSS** – Síndrome do choque tóxico estreptocócico (“streptococcal toxic shock syndrome”)

**tARN** – Ácido ribonucleico de transferência

**TLVs** – “Triple *locus* variants”

**VFDB** – “Virulence Factors Database”

**WGS** – “Whole genome sequencing”

# 1. INTRODUÇÃO

---

## 1.1 *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE)

*Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) é uma bactéria que pertence ao grupo dos estreptococos beta-hemolíticos dos grupos C e G de Lancefield formadores de colônias grandes (EGCG) (1–3). Este grupo inclui várias espécies com diversos hospedeiros animais, das quais SDSE é a mais frequentemente isolada do Homem (4). SDSE é filogeneticamente próximo de *Streptococcus pyogenes* (grupo A de Lancefield) (5–8) e, em conjunto com este e com *Streptococcus agalactiae* (grupo B de Lancefield), é uma das espécies de estreptococos beta-hemolíticos causadoras de maior número de infecções humanas (4).

Apesar de se encontrar desde há muito identificado como importante agente de infecção em animais, SDSE foi, durante muitos anos, considerado um microrganismo de baixa patogenicidade para o Homem, sendo descrito neste último como parte da microbiota comensal (9,10). No entanto, a partir dos anos 80 do século XX, foi notado um aumento da incidência das infecções humanas causadas por esta bactéria, o que levou ao seu reconhecimento como agente patogénico emergente no Homem (3,9–11) e à necessidade de uma clarificação da sua taxonomia.

As estirpes incluídas atualmente na espécie *Streptococcus dysgalactiae* foram, durante muito tempo, divididas em pelo menos cinco grupos, com base nas suas diferentes características fenotípicas (grupo de Lancefield, padrão de hemólise e vários testes bioquímicos) e nos hospedeiros a partir dos quais eram isoladas (3,12–14). Em 1996, Vandamme *et al.*, baseando-se em testes fisiológicos e “whole-cell protein analysis”, sugeriu pela primeira vez a divisão desta espécie em duas subespécies, *equisimilis* e *dysgalactiae*, a primeira incluindo as estirpes isoladas de humanos e a segunda as estirpes isoladas de animais (13). No entanto, dois anos mais tarde, um estudo baseado na hibridação ADN-ADN e “multilocus enzyme electrophoresis” (MLEE), conduzido por Vieira *et al.* (14), propôs algumas modificações a esta divisão, separando as estirpes em função da hemólise e não do hospedeiro (**Tabela 1**).

**Tabela 1:** Nomenclatura das estirpes atualmente incluídas no taxon *S. dysgalactiae*.

Divisões taxonómicas	Caraterísticas das estirpes
Grupos reconhecidos inicialmente (3)	Estreptococos do grupo C de Lancefield, alfa-hemolíticos ou não hemolíticos, isolados de bovinos ( <i>S. dysgalactiae</i> *)
	Estreptococos do grupo C de Lancefield, beta-hemolíticos, isolados de porcos ( <i>S. equisimilis</i> *)
	Estreptococos do grupo L de Lancefield, beta-hemolíticos, isolados de vários animais (porcos, cães e gado) e, raramente, de humanos
	Estreptococos do grupo C de Lancefield, beta-hemolíticos, isolados de humanos ( <i>S. equisimilis</i> *)
	Estreptococos do grupo G de Lancefield, beta-hemolíticos, isolados de humanos
Divisão de Vandamme <i>et al.</i> (13)	<i>S. dysgalactiae</i> subsp. <i>equisimilis</i> subsp. nov.: estreptococos dos grupos C e G de Lancefield, beta-hemolíticos, formadores de colónias grandes, isolados de humanos
	<i>S. dysgalactiae</i> subsp. <i>dysgalactiae</i> subsp. nov.: estreptococos dos grupos C e L de Lancefield, hemólise variável, formadores de colónias grandes, isolados de animais
Divisão de Vieira <i>et al.</i> (14)	<i>S. dysgalactiae</i> subsp. <i>equisimilis</i> : estreptococos beta-hemolíticos dos grupos C, G e L formadores de colónias grandes, isolados de humanos e animais
	<i>S. dysgalactiae</i> subsp. <i>dysgalactiae</i> : estreptococos alfa-hemolíticos ou não hemolíticos do grupo C, formadores de colónias grandes, isolados de animais

\*Nome da espécie pelo qual eram referidas as estirpes pertencentes ao grupo indicado. Adaptado de (3).

Atualmente, a divisão proposta por Vieira *et al.* é a mais frequentemente utilizada. No entanto, a classificação de algumas estirpes de origem animal permanece dúbia, conquanto as suas características genóticas foram alvo de um limitado número de estudos e os resultados obtidos nem sempre congruentes (5,15,16).

## 1.2 Características morfológicas e identificação laboratorial

SDSE, à semelhança das outras espécies que compõem o género *Streptococcus*, é composto por cocos (células esféricas ou ovoides) de Gram-positivo normalmente dispostos em pares ou cadeias (1,2). SDSE é uma bactéria aeróbia facultativa, catalase-negativa e possui necessidades nutricionais complexas, exigindo o uso de meios enriquecidos com sangue ou soro para o seu isolamento laboratorial (1,2).

A identificação é, primariamente, feita com base na observação do crescimento das estirpes em placas de ágar suplementadas com sangue e na determinação do grupo de Lancefield (4). Estirpes humanas de SDSE possuem, maioritariamente, os antigénios dos grupos C e G de Lancefield e formam colónias grandes (diâmetro maior do que 0,5mm) – o que as distingue de estreptococos que podem possuir os mesmos grupos de Lancefield, mas que formam colónias pequenas (grupo “*anginosus*” ou “*milleri*”) – rodeadas por uma ampla zona de beta-hemólise após 24h de incubação (1–3,17) (**Figura 1**). Raramente, podem apresentar outro padrão de hemólise ou possuir os grupos de Lancefield A ou L (1–3,17,18).



**Figura 1:** Morfologia das colónias e padrão de hemólise.

Estirpe de SDSE beta-hemolítica e formadora de colónias grandes (>0,5mm) após crescimento durante 24 horas a 35 °C em placas de ágar suplementadas com sangue.

O uso dos padrões de hemólise para classificar os estreptococos remonta ao início do século XX, quando Schottmüller, em 1903, utilizou meios enriquecidos com sangue para diferenciar as estirpes beta-hemolíticas das não beta-hemolíticas (18,19). As espécies estreptocócicas apresentam beta-hemólise quando as colónias, após crescimento em meio sólido contendo sangue, ficam rodeadas por um halo translúcido, correspondente a uma zona de lise total dos eritrócitos (hemólise) (19) (**Figura 1**). Outros padrões de hemólise são a alfa-hemólise, uma hemólise parcial que ocorre devido a uma redução da hemoglobina nos eritrócitos que origina um halo esverdeado, e a gama-hemólise que corresponde à ausência de hemólise (19).

A classificação de Lancefield refere-se ao esquema desenvolvido, em 1933, por Rebecca Lancefield, que dividia os estreptococos beta-hemolíticos em grupos, por meio de uma reação de precipitação (1,20). Esta técnica serológica tem por base a deteção de antígenos específicos existentes na parede celular dos estreptococos, resultantes da variabilidade do carboidrato C, uma substância solúvel de natureza polissacarídea (1,20).

A caracterização com base na hemólise, tamanho das colónias e grupo de Lancefield, é insuficiente para obter a identificação da espécie, devido à sobreposição de características entre espécies que podem causar infeção em humanos (4,18). Embora menos frequentes como agentes de infeção neste hospedeiro, estirpes de *Streptococcus canis* e *Streptococcus equi* subsp. *zooepidemicus* podem ser confundidas com SDSE, pois possuem o antígeno G e C de Lancefield, respetivamente (2,18). A distinção pode ser feita com base em testes bioquímicos (18,21), incluindo o recurso a “kits” comerciais e sistemas automatizados – apesar de nem sempre fornecerem uma classificação correta (3,22).

Outros métodos têm sido aplicados, entre os quais se destaca “matrix-assisted laser desorption ionization–time of flight mass spectrometry” (MALDI-TOF MS), que alia rapidez e baixo custo (21,23–25). Esta tecnologia, baseada na composição proteica das células microbianas, especificamente proteínas ribossomais, tem tido uma aplicação exponencial nos últimos anos (23,25). A utilização de MALDI-TOF MS permite a identificação das principais espécies de EGCG (23,26), apesar de estarem descritas dificuldades na diferenciação de *S. dysgalactiae*, *S. pyogenes* e *S. canis* (24) e a identificação de *S. dysgalactiae* só ser habitualmente possível a nível da espécie (21,23,26).

A utilização de métodos genotípicos para a identificação de SDSE encontra-se igualmente descrita na literatura, embora a sua utilização para este fim não seja prática corrente em laboratórios que façam diagnóstico clínico. Vários estudos reportaram métodos baseados na amplificação por reação em cadeia de polimerase (PCR) e/ou sequenciação de genes específicos, focando-se na distinção de *S. pyogenes* e SDSE (27,28). Os métodos de sequenciação do ADN, principalmente tendo como alvo o gene 16S rRNA – gene altamente conservado dentro de uma espécie e entre espécies do mesmo género – tornaram-se o método padrão para a identificação das bactérias ao nível da espécie (3,22,29,30). A sequenciação do gene do 16S rRNA permite a distinção de SDSE de outras espécies beta-hemolíticas, mas a discriminação a nível da subespécie de *S. dysgalactiae* nem sempre é possível (31). Para este fim, há a necessidade de se aplicar esquemas de sequenciação multilocus, que utilizam múltiplos genes, dado que a informação obtida a partir de genes individuais nem sempre é congruente (5).

### **1.3 Reservatórios de SDSE e origem da infeção**

Em humanos, SDSE pode ser encontrado como parte da microbiota, colonizando a pele e as mucosas dos tratos respiratório superior, gastrointestinal e genitourinário feminino (3,32,33).

Estudos que determinaram a presença assintomática de EGCG na faringe revelaram taxas de colonização que variam com a região geográfica (34,35). Apesar de em muitos destes estudos não realizarem a identificação da espécie de EGCG, as taxas destes estudos refletirão maioritariamente a presença de SDSE, dado que SDSE é a única espécie deste grupo para a qual a colonização é reconhecida. Taxas de 10% em crianças da Índia (35) e superiores a 20% nas comunidades aborígenes do Norte da Austrália (34) foram reportadas. Já no Japão (36) e nos EUA (37,38) as taxas apresentam-se menores, variando entre os 2% e os 4%, em crianças e adultos, respetivamente. Um estudo brasileiro mostrou que a presença assintomática de SDSE ocorria maioritariamente em jovens adultos (6,3%) e que a persistência era considerável (37,5% dos portadores assintomáticos de SDSE após 12 meses) (39). Relativamente a outros locais de colonização, um estudo finlandês obteve uma taxa de colonização de 2,9% após a realização de exsudados vaginais e retais em grávidas (40).

Assim, o principal reservatório de SDSE é o hospedeiro humano e a maioria das infecções tem, provavelmente, origem endógena, ocorrendo transmissão pessoa-pessoa através do contacto direto com secreções respiratórias, principalmente aerossóis, ou com lesões na pele infetadas (2,3,22,32). Estão documentados surtos de faringite por SDSE com presumível origem no consumo de alimentos contaminados por portadores desta bactéria (41). Alguns estudos detetaram a presença de estirpes de SDSE com características genóticas idênticas em humanos e animais, incluindo cavalos (15,42) e cães (43). Embora estas observações sugiram que algumas infecções humanas possam ocorrer no contexto de uma zoonose, tal deverá corresponder a uma minoria dos casos de infeção por SDSE.

#### **1.4 Principais infeções**

SDSE é agente etiológico de várias doenças que variam consideravelmente em termos de gravidade, num espectro de infeções semelhante ao de *S. pyogenes* (3,14,44,45). As infeções mais comuns estão associadas aos locais de colonização, sobretudo a pele e os tecidos moles e o trato respiratório superior (3,9,11,32,46,47). Estas infeções incluem celulite, impetigo, erisipela, infeção de feridas, faringite, endoftalmite e otite média (3,4,32,48). As infeções mais graves incluem bacteriemia, pneumonia, meningite, osteomielite, artrite séptica, endocardite, fascíte necrotizante, síndrome do choque tóxico estreptocócico (“streptococcal toxic shock syndrome” - STSS), entre outras (9,10,17,44,49–51).

SDSE é um reconhecido agente de faringo-amigdalite em jovens adultos (3,37,52,53). A infeção em crianças é menos comum e alguns estudos não conseguiram estabelecer uma associação estatisticamente significativa entre a presença de SDSE na orofaringe e a presença de sintomatologia (4,35,36,38). Surtos epidémicos de faringite causados por SDSE afetam não só adultos, mas também crianças (3,38).

As infeções invasivas (isto é, isolamento de SDSE de um local normalmente estéril, como o sangue, o líquido cefalorraquidiano, o líquido pleural, entre outros) causadas por SDSE, com ocorrência de bacteriemia, têm sido largamente, e cada vez mais, reportadas em todo o mundo (3,4,17,33,44,54–56). A origem mais frequente são os locais de

colonização e as infecções locais (3,32). Na maioria dos casos, existe uma disrupção da barreira cutânea que proporciona a invasão do microrganismo (47,50), sendo as infecções na pele e tecidos moles (3,47,49,50,57), especialmente a celulite (22,47,49,58,59), mais vezes associadas à ocorrência de bacteriemia. A ocorrência de bacteriemia primária, ou seja, sem fonte de infecção conhecida é, por vezes, descrita (49,58,60,61) e casos de recorrência foram verificados por alguns autores (49,59). O primeiro caso de STSS causado por SDSE foi descrito em 1996 (62) e desde então vários têm sido os estudos que descrevem SDSE como causador deste síndrome, com elevadas taxas de mortalidade (30,44,63). SDSE foi também implicado no desenvolvimento de síndromes não supurativos tipicamente causados por *S. pyogenes*, como a febre reumática aguda ou a glomerulonefrite, embora não exista ainda uma ligação clara (7,35,64,65).

### **1.5 Características clínicas e epidemiológicas das infecções por SDSE**

A prevalência das infecções causadas por SDSE tem aumentado nos últimos anos, de acordo com estudos conduzidos em diferentes locais do mundo (17,55). Nalgumas regiões, a prevalência das infecções por SDSE é já comparável à prevalência das infecções por *S. pyogenes* e *S. agalactiae* (4,55), habitualmente considerados os principais estreptococos beta-hemolíticos causadores de infecções no Homem (4). Alguns estudos indicam que SDSE é isolado mais frequentemente que estas duas espécies (34,35,66). Estudos do Canadá (67) e EUA (4) têm reportado taxas de incidência de infecções invasivas de 2,24 casos e 3,2 casos/100.000 habitantes, respetivamente, mas a Finlândia (47) e a Noruega (68) têm reportado taxas mais elevadas, de 4,3 casos/100.000 habitantes e 6,3 casos/100.000 habitantes, respetivamente.

A maioria das infecções por SDSE são adquiridas na comunidade (45,50,66). No entanto, alguns estudos reportam até um terço das infecções como estando associadas a cuidados de saúde (4,11,22,69). As infecções invasivas por SDSE são mais frequentes em doentes idosos (49,66,69), existindo uma estreita relação entre aumento da idade e risco de desenvolver doença invasiva (56,67) e as taxas de infecção são maiores em homens do que em mulheres (49,58,67,68). A maioria dos doentes que desenvolve infecção invasiva tem, pelo menos, uma comorbidade (4,32,60,66,70), entre as quais se incluem diabetes *mellitus*, doença oncológica, doenças cardiovasculares, imunossupressão, doenças dos

ossos e articulações, alcoolismo com desenvolvimento de cirrose, uso de drogas ilícitas, doença renal, hipertensão ou outras condições que comprometam a integridade da pele (11,22,32,45,50,58,59,61,70). Aliado ao aumento das taxas de incidência têm sido reportadas taxas de mortalidade e morbidade elevadas (4), com taxas de mortalidade que variam desde 9% a 25% (11,22,66,68) e taxas de seqüela entre 5,3% (17) e 25% (45). Assim, a idade (3,60), bem como as condições subjacentes (3,11), são consideradas fatores de risco e é previsível que a mortalidade aumente com o aumento da idade e a presença de doenças subjacentes (11).

Embora a idade e a presença de comorbidades sejam considerados fatores de risco para as infecções invasivas por SDSE, as alterações imunológicas associadas à idade têm sido propostas como o fator major (14,50,56). A doença invasiva ocorre pois a degeneração da barreira cutânea – barreira física à infecção – permite a invasão de agentes colonizadores da pele, como SDSE e, como as defesas do hospedeiro estão comprometidas, o sistema imunitário torna-se incapaz de controlar a infecção local (50). Logo, SDSE é considerado um agente patogênico oportunista, associado à senescência do sistema imunitário (50).

### **1.6 Suscetibilidade a antimicrobianos e tratamento das infecções**

As estirpes de SDSE são suscetíveis aos antimicrobianos habitualmente usados no tratamento de infecções por outros estreptococos beta-hemolíticos (9,71). O antibiótico de primeira escolha para o tratamento de infecções por SDSE é a penicilina, pois, à semelhança de *S. pyogenes*, a generalidade das estirpes de SDSE continuam a apresentar suscetibilidade aos beta-lactâmicos (3,9,11). No passado, foram reportados casos de tolerância à penicilina – definida como uma concentração bactericida mínima (CBM), pelo menos 32 vezes superior à concentração inibitória mínima (CIM) – mas o impacto clínico desta observação permanece desconhecido (71–74). Em 2016, na Dinamarca, foi descrito o primeiro caso de resistência à penicilina em SDSE (CIM=0,5 mg/L) (74). Este estudo descreveu o isolamento de quatro estirpes que possuíam alterações nas proteínas de ligação à penicilina (“penicillin binding proteins”, PBPs), nomeadamente na PBP2x, PBP1a e PBP1b, resultando presumivelmente numa redução da afinidade para os beta-

lactâmicos (74). As estirpes eram epidemiologicamente relacionadas e a análise genotípica das mesmas mostrou que pertenciam a um único clone (74).

A combinação de um beta-lactâmico com um aminoglicosídeo é uma hipótese no tratamento de infecções invasivas graves por SDSE, pois têm efeito sinérgico (3,9,11,32,73). Em doentes com STSS, fascíte necrotizante ou formas graves de celulite, a administração de penicilina em conjunto com a clindamicina (uma lincosamida) é também uma hipótese (3,32,75). Estas combinações parecem ser preferíveis ao aumento da dose de penicilina, dado que não se verificou vantagem terapêutica nesta última medida (71).

Em doentes com alergia à penicilina e outros beta-lactâmicos, os macrólidos, as lincosamidas e a vancomicina são opções terapêuticas alternativas (3,11,75). SDSE continua a ser uniformemente suscetível aos glicopéptidos, apesar de, à semelhança do caso da penicilina, estirpes de SDSE com tolerância à vancomicina terem sido descritas no passado (9,76). SDSE é igualmente suscetível a antimicrobianos de outras classes, tais como a linezolid (uma oxazolidinona) (3,77,78). A nova geração de fluoroquinolonas tem cada vez mais uso, pois têm uma excelente cobertura e as resistências são menos frequentes (3). Por outro lado, as tetraciclinas são o grupo de antibióticos ao qual estão reportadas mais resistências em estreptococos beta-hemolíticos, o que faz com que não sejam habitualmente uma opção no tratamento (3).

Tendo em conta que as estirpes de SDSE têm diferentes padrões de suscetibilidade, podendo apresentar resistências a diferentes classes de antibióticos, testes de suscetibilidade devem ser realizados para determinar a terapêutica apropriada (79,80).

### **1.7 Resistência a antimicrobianos**

As estirpes de SDSE são frequentemente resistentes à tetraciclina (4,44,81) e taxas de resistência expressivas a macrólidos e lincosamidas também têm sido reportadas (4,44,67,79,82). Estirpes de SDSE resistentes a fluoroquinolonas (4,66,79,83) e com resistência de alto nível a aminoglicosídeos (81) foram igualmente descritas, apesar de menos frequentes que as anteriores.

Os macrólidos e lincosamidas, assim como as estreptograminas B, têm o mesmo alvo na célula bacteriana, atuando pela ligação reversível à subunidade 50S do ribossoma. Esta ligação bloqueia as reações de translocação e alongamento da cadeia peptídica, inibindo, assim, a síntese proteica (84–86). A resistência a estes antimicrobianos, à semelhança de outros estreptococos beta-hemolíticos, é majoritariamente conferida por dois mecanismos distintos (79,84–86). Por um lado, a presença de uma metilase, codificada pelos genes *erm*, a qual modifica o local de ligação no ribossoma e confere resistência cruzada às três classes de antibióticos, dado que partilham o mesmo local de ligação. Os genes *ermA*, *ermB* e, raramente, *ermT*, foram previamente detetados em SDSE (64,79,82,87,88). Este mecanismo corresponde ao fenótipo MLS<sub>B</sub> (resistência a macrólidos, lincosamidas e estreptogramina B), que pode ser induzido (iMLS<sub>B</sub>) ou constitutivo (cMLS<sub>B</sub>) (79,84–86). O outro mecanismo é a presença de uma bomba de efluxo, codificada pelos genes *mef*, que confere resistência apenas aos macrólidos e corresponde ao fenótipo M (79,84–86). SDSE possui os genes *mefA* ou *mefE* (64,79,82,88).

Os vários fenótipos de resistência a macrólidos e lincosamidas podem ser distinguidos laboratorialmente através da realização de um antibiograma em que se colocam discos de eritromicina (macrólido) e clindamicina (lincosamida) em proximidade (89) (**Figura 2**). O fenótipo cMLS<sub>B</sub> ocorre quando se observa resistência a ambos os antimicrobianos (**Figura 2 – A**). O fenótipo iMLS<sub>B</sub> observa-se quando a resistência à clindamicina só ocorre na presença da eritromicina, ou seja, é induzida por esta, resultando numa deformação do halo de inibição da clindamicina na proximidade do disco de eritromicina (em “D”) (**Figura 2 – B**). O fenótipo M é definido por resistência à eritromicina e halo de suscetibilidade à clindamicina, sem observação de qualquer interação (**Figura 2 - C**).



**Figura 2:** Fenótipos de resistência aos macrólidos e lincosamidas.

(A) Fenótipo cMLS<sub>B</sub>; (B) Fenótipo iMLS<sub>B</sub>; (C) Fenótipo M.  
E – Disco de eritromicina; DA – Disco de clindamicina.

A utilização de macrólidos e lincosamidas no tratamento das infecções causadas por SDSE é limitada pelas taxas de resistência variáveis que têm sido reportadas em estudos de vários países, tais como 11% no Canadá (67), 13% na Áustria (78), 23% no Japão (82), 24% na China (79), 26,4% em França (44) e 29% nos EUA (4).

As tetraciclina atuam pela ligação reversível à subunidade 30S do ribossoma, não permitindo a ligação do ácido ribonucleico de transferência (tARN), o que inibe a síntese proteica (84,90,91). A aquisição de genes *tet* confere resistência a esta classe de antibióticos, tal como acontece em outros cocos de Gram-positivo. Estes genes codificam bombas de efluxo ou enzimas modificadoras do local de ligação no ribossoma, ou seja, proteínas de proteção ribossomal que promovem uma alteração no local de ligação das tetraciclina no ribossoma impedindo a ligação destas (84,90,91). Os genes *tetL* (bomba de efluxo) e os genes *tetM*, *tetO* e *tetS* (enzimas modificadoras do local de ligação) foram previamente detetados em SDSE (44,70,81,92). As tetraciclina não são habitualmente utilizadas no tratamento empírico de infecções por SDSE (3), devido às elevadas taxas de resistência reportadas por diferentes países, tais como 35% em França (44) e 40% na Argentina (81). A presença de genes de resistência à tetraciclina e aos macrólidos e lincosamidas nos mesmos elementos genéticos móveis foi descrita em algumas espécies de *Streptococcus* (93,94).

As fluoroquinolonas atuam pela ligação à ADN girase (composta por duas subunidades, GyrA e GyrB) e à topoisomerase IV (composta por duas subunidades, ParC e ParE), enzimas que são essenciais para o enrolamento do ADN bacteriano e para todos os processos que requerem esse enrolamento (replicação, transcrição, entre outros) (80,83,84). O mecanismo de resistência mais comum em bactérias de Gram-positivo é a alteração das enzimas alvo de ligação do antibiótico, através de mutações nas designadas regiões determinantes da resistência às quinolonas (“quinolone resistance-determining regions” – QRDRs) nas sequências de ADN de *gyrA* e *parC* (80,83,84). Apesar de terem sido igualmente detetadas alterações de aminoácidos nas subunidades GyrB e ParE, o papel destas no desenvolvimento de resistência não se encontra bem estabelecido (80,83,84). Este mecanismo é o único que confere resistência de alto nível, mas a resistência também pode ser conferida por bombas de efluxo (80,83,84). Embora estudos dos EUA (4), da China (79), do Japão (66) e da maioria da Europa (3) reportem baixas taxas de resistência, rondado os 1% em SDSE (3,4,17,66,79), um estudo de Portugal

reportou uma taxa de resistência de 12% em estirpes isoladas de infecções invasivas e não invasivas entre os anos de 1998 e 2005 (83).

Os aminoglicosídeos inibem a síntese proteica pela ligação reversível à subunidade 30S do ribossoma (84,95,96). São uma das classes de antibióticos à qual foram reportadas menor número de estirpes de SDSE resistentes. Os estreptococos são intrinsecamente resistentes a baixas concentrações destes antimicrobianos, à semelhança de *Enterococcus* spp., devido à baixa permeabilidade da parede celular, o que limita o transporte através da membrana citoplasmática e, conseqüentemente, a internalização do antibiótico (81,84,95,96). Resistências de alto nível (definidas como CIM>512 µg/mL para a gentamicina e CIM>2000 µg/mL para a estreptomicina) são pouco comuns em SDSE (81,84,89,95,97). Os mecanismos que promovem resistência de alto nível aos aminoglicosídeos são a modificação do alvo de ligação e a modificação enzimática do antibiótico, sendo este último o mecanismo mais comum que ocorre por enzimas que modificam o grupo amino ou hidroxilo do antibiótico (84,95,96).

### **1.8 Fatores de virulência**

O genoma de SDSE codifica várias proteínas que podem potencialmente contribuir para a virulência desta bactéria (47,98). Entre estas incluem-se produtos bacterianos que contribuem para a adesão aos tecidos do hospedeiro ou para escapar à imunidade do hospedeiro, toxinas ou proteases, entre outros (3,17,33). Foram igualmente identificados nos genomas destas bactérias reguladores que controlam a expressão de vários genes de virulência e que atuam durante as várias fases da infecção (17,33).

Muito do que se sabe atualmente acerca dos fatores de virulência de SDSE provém da pesquisa de genes previamente identificados em *S. pyogenes* (6,17,32,99). A existência de um ancestral comum entre *S. pyogenes* e SDSE justifica a sobreposição de propriedades de virulência, sendo muitos dos fatores de virulência de SDSE semelhantes aos de *S. pyogenes* e codificados por genes homólogos (5–8). Acresce o facto de poder existir transferência genética horizontal entre as duas espécies envolvendo alguns destes genes (7,100).

Entre os fatores de virulência que SDSE partilha com *S. pyogenes*, incluem-se a proteína M, a exotoxina pirogênica SpeG, várias proteínas extracelulares – tais como as estreptolisinas O e S (“streptolysin O” [SLO] e “streptolysin S” [SLS], respetivamente), a estreptoquinase e a estreptodornase – e proteínas de superfície celular, como a C5a peptidase (gene *scpA*) e as proteínas de ligação à fibronectina, ao plasminogénio, à laminina e ao colagénio (por exemplo, produtos dos genes *fbp54*, *lmb*, *gapC* e *cbp*, respetivamente) (3,6,17,100–104).

As SLO e SLS são citolisinas responsáveis pelo fenótipo de beta-hemólise de SDSE devido à sua atividade como hemolisinas (105,106). Ambas são codificadas em operões idênticos aos encontrados nos genomas de *S. pyogenes* e a atividade destas citolisinas foi relacionada com a capacidade das estirpes de SDSE causarem infeções graves da pele e tecidos moles (105,106). O operão que inclui o gene da SLO (*slo*) contém igualmente o gene *nga*, que codifica uma NAD(+)-glicohidrolase (NADase) (107). Em *S. pyogenes*, crê-se que a SLO forma poros nas membranas das células do hospedeiro, facilitando a entrada da NADase no citoplasma onde exerce o seu efeito (108). A estreptoquinase (gene *ska*) é uma enzima extracelular, semelhante à estreptoquinase A de *S. pyogenes*, que ativa o plasminogénio e, após essa ativação, exerce uma atividade fibrinolítica (104,109). A estreptodornase (gene *sda*) é uma DNase que, em SDSE, à semelhança da estreptodornase D de *S. pyogenes*, é codificada num fago (103,104). Em *S. pyogenes*, esta nuclease provavelmente contribui para a mutação de um sistema de regulação (*covRS*) promovendo assim a conversão de *S. pyogenes* num fenótipo mais virulento (103,104).

Duas regiões do tipo FCT (para ligação da fibronectina, colagénio e antigénio T) foram identificadas nos genomas de SDSE, as quais codificam estruturas do tipo *pili* em *S. pyogenes* e *S. agalactiae* (104).

SDSE não só partilha muitos genes de virulência com *S. pyogenes*, mas também possui homólogos dos sistemas reguladores, tais como *covRS*, *fasCAX* e *sil* (“streptococcal invasion locus”) (109–111). Foi descrito que os reguladores *covRS* e *sil* exercem um controlo transcricional negativo da expressão de fatores de virulência e que mutações em genes desses reguladores levam a um aumento da expressão desses fatores, com consequente aumento da virulência das estirpes (109–112). Enquanto *covRS* está presente em praticamente todas as estirpes de *S. pyogenes* e SDSE, *sil* tem uma presença variável, sendo mais comum em SDSE que em *S. pyogenes* (cerca de 80% e 25%, respetivamente) (112). Este regulador é composto por seis “open reading frames” (ORFs), *silA* e *silB*

codificam um sistema de dois componentes, *silD* e *silE* codificam um transportador e *silC* e *silCR* codificam péptidos associados a virulência e competência, respetivamente (111,112). Um estudo norueguês recente descreveu um aumento das infeções invasivas causadas por estirpes SDSE com o grupo C de Lancefield, com o tipo *emm stG62647* e pertencentes ao complexo clonal CC20 (111). O estudo genómico destas estirpes mostrou que o gene *silB* era interrompido pela inserção de uma transposase, IS1548 (111). Foi sugerido pelos autores do estudo que a perda de funcionalidade do regulador associada a esta disrupção do gene *silB* poderia contribuir para uma maior virulência desta linhagem de SDSE, apesar de não terem sido feitos quaisquer estudos funcionais para comprovar esta hipótese (111). Infeções por este clone na Noruega foram associadas a formas graves de doença invasiva, incluindo STSS, infeções cutâneas necrotizantes e endocardite (111).

Apesar da partilha de um grande número de genes, SDSE carece do regulador *ropB* e de vários genes que codificam importantes fatores de virulência em *S. pyogenes*. Entre estes contam-se a maioria dos genes das exotoxinas pirogénicas, ou a protease de cisteína *speB* (6,33,103,104,113). Vários estudos indicam igualmente que o operão *hasABC*, necessário para a síntese da cápsula de ácido hialurónico em *S. pyogenes*, não é conservado em SDSE (6,33,103,104).

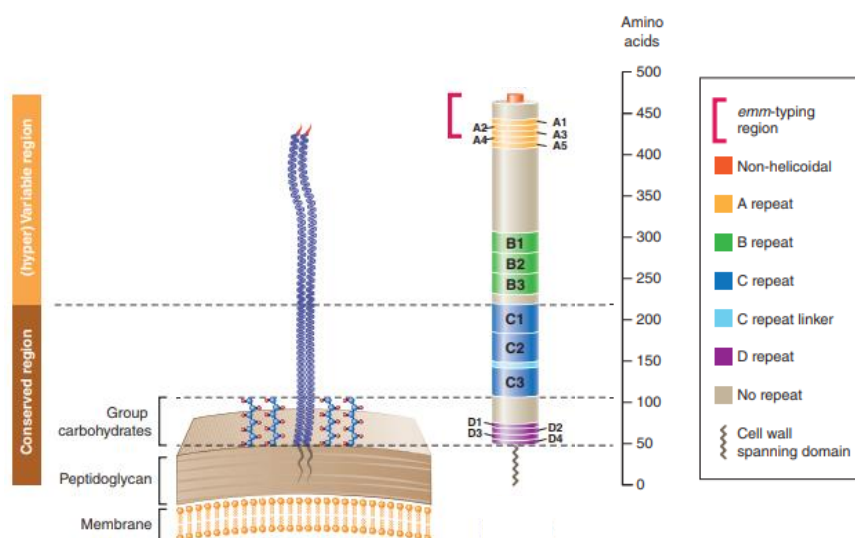
### 1.8.1 Proteína M

A proteína M, codificada pelo gene *emm* (17,32), foi identificada há mais de 90 anos em *S. pyogenes* por Rebecca Lancefield e, mais tarde, definida como sendo um fator de virulência major nesta bactéria, principalmente devido à sua atividade antifagocítica (3,114,115).

A semelhança existente entre *S. pyogenes* e SDSE sugeriu que esta última espécie também possuísse a proteína M (116). A presença desta proteína em SDSE e noutras espécies de EGCG, foi confirmada através de estudos nos quais foi demonstrada reação cruzada com anticorpos policlonais para a proteína M de *S. pyogenes* e pela sua capacidade de resistir à fagocitose e de se multiplicar em sangue humano (114,116,117). Tal como em *S. pyogenes*, crê-se que a proteína M é um fator de virulência chave em SDSE e, nestas espécies, este fator de virulência parece ser funcional e estruturalmente semelhante (114,118). Estendendo-se desde a superfície celular, a proteína M apresenta-se como um dímero fibrilar em espiral enrolada alfa-helicoidal e está ancorada na parede

celular, estrutura que contribui para a sua multifuncionalidade (115,117,119). Funcionalmente, a proteína M inibe a opsonofagocitose, interferindo com a ativação das vias do complemento na superfície bacteriana, pela ligação à proteína reguladora do complemento (fator H), ao fibrinogénio e à albumina (7,32,115,120,121). Esta molécula também se liga a outros componentes do hospedeiro, como colagénio e plasminogénio, para facilitar a adesão aos tecidos e invasão do microrganismo (7,121,122). No entanto, sabe-se que nem todas as variantes da proteína exibem as mesmas funções (115).

A sua estrutura básica (**Figura 3**) inclui um péptido sinal, uma região N-terminal hipervariável e não helicoidal, um domínio central e uma região C-terminal altamente conservada e ancorada na parede celular (114,115,119,121). A região hipervariável é seguida de um número de regiões de repetição (A, B, C e D) com tamanho e composição variáveis (114,115,119,121).



**Figura 3:** Esquema da estrutura da proteína M de GAS.

Adaptado de (123).

O elevado grau de polimorfismo da região N-terminal entre diversas proteínas M confere a especificidade serológica, enquanto a região C-terminal contém uma sequência que é partilhada entre a maioria dos serotipos, mesmo entre espécies diferentes (21,100,115,117,118). Apesar da organização geral das proteínas M ser conservada, existem variações de tamanho – dependente do número e tamanho das repetições – contribuindo para o não reconhecimento pelos anticorpos do hospedeiro (115,119,121,122). A presença de duas classes de proteína M (classe I ou II) foram

definidas em *S. pyogenes*, com base nas variações na estrutura de segmentos repetidos na região C-terminal conservada (114,115). O domínio conservado de SDSE mostrou ser estruturalmente semelhante às moléculas de classe I de *S. pyogenes* (118,119,124).

Foi demonstrado que pode induzir imunidade protetora (adquirida), pois possui epítomos que levam à produção de anticorpos opsonizantes contra serotipos específicos desta proteína (3,115,117,124,125).

### 1.8.2 Superantigénios

Em *S. pyogenes*, as exotoxinas pirogénicas são um conjunto de proteínas que podem atuar como superantigénios, isto é, proteínas que estimulam a proliferação de linfócitos (células T), induzem febre e também aumentam a suscetibilidade a reações letais mediadas por endotoxinas (126,127). Esta estimulação de linfócitos já foi implicada na patogénese de doenças infecciosas (126,127). A presença de superantigénios em *S. pyogenes* foi associada ao desenvolvimento de STSS em doentes com infeções por esta bactéria (128). O crescente número de casos descrito de doentes com infeções invasivas por SDSE que desenvolvem STSS (17,32,68) sugeriu que esta espécie também possuísse superantigénios. Além disso, a maioria dos genes que codificam superantigénios estreptocócicos estão associados a bacteriófagos, elementos genéticos móveis que são provavelmente o mecanismo primário de transferência lateral de genes entre *S. pyogenes* e SDSE (119). Apesar de vários estudos terem tentado identificar estes fatores de virulência em SDSE, o único identificado repetidamente em SDSE foi o gene *speG* (104,119,126,127,129,130). No entanto, o gene que codifica este superantigénio não está localizado num elemento genético móvel em *S. pyogenes* e, portanto, a sua presença em SDSE deve-se à presença de um ancestral comum entre as espécies e não devido a transferência horizontal (127,131). Outros destes genes, tais como *speA*, *speC*, *speH*, *speI*, *speJ*, *speM*, *ssa* e *smeZ*, foram detetados esporadicamente em estirpes de SDSE isoladas de doentes sem evidência de STSS (3,92,119,129).

A contribuição do produto do gene *speG* – ortólogo do gene *speG* de *S. pyogenes* – na patogénese de infeções por SDSE em humanos é controversa (104,119,126,127,129–131). Alguns estudos não detetaram qualquer atividade mitogénica em estirpes de SDSE com este gene (54,78,126,127). O gene *speG* não está presente em todas as estirpes de SDSE, com estudos de diferentes regiões geográficas a reportar percentagens variáveis

de estirpes positivas para este gene, desde 43% no Japão (70), 58% em Israel (129), 63% na Alemanha (130) e 66% na Noruega (119) e no Canadá (98). Uma vez que a presença de *speG* não é superior em estirpes de SDSE isoladas de doentes com STSS e outros superantigénios conhecidos não são habitualmente detetados, pode sugerir que SDSE possui outros superantigénios que ainda não foram identificados (54,77,78,129).

### 1.9 Métodos de tipagem

A utilização de um método de tipagem tem como objetivo a determinação de relações clonais entre diferentes estirpes com base em caracteres fenotípicos ou genotípicos específicos, permitindo elucidar as fontes e as vias de propagação da bactéria, bem como identificar estirpes ou clones emergentes dentro de uma espécie (132). Além disso, a informação obtida pode ser relevante para a taxonomia, a ecologia e o estudo da patogénese (132).

Entre os métodos fenotípicos utilizados para caracterizar as estirpes de SDSE incluem-se a técnica serológica de Lancefield (20), testes bioquímicos (133), a serotipagem T (134) ou a tipagem de fagos e bacteriocinas (9). Com exceção da classificação de Lancefield, estes métodos não são atualmente utilizados na tipagem das estirpes de SDSE. Vários métodos genotípicos foram aplicados na caracterização epidemiológica de estirpes de SDSE. Estes incluem os baseados na separação eletroforética de fragmentos de ADN, como a eletroforese em gel de campo pulsado (“Pulsed-field gel electrophoresis” – PFGE) (29,135), e os baseados na análise da sequência de um ou mais genes, como a tipagem do gene *emm* ou a “multilocus sequence typing” (MLST), respetivamente (6,136,137). A técnica de PFGE foi, durante muito tempo, considerada o “gold standard” dos métodos de tipagem, devido, principalmente, ao seu elevado poder discriminatório (132,135,138,139). No entanto, o método mais utilizado na caracterização epidemiológica de estirpes de SDSE tem sido a tipagem do gene *emm* (3,44,68,99). A introdução da técnica de MLST (132) e, mais recentemente, da sequenciação de alto débito (98,140), ganharam especial relevância nos últimos anos.

### 1.9.1 Tipagem do gene *emm*

A tipagem do gene *emm* foi inicialmente desenvolvida para caracterizar estirpes de *S. pyogenes*, como alternativa genotípica à serotipagem da proteína M (141). Sendo a proteína M um fator de virulência major em *S. pyogenes* e SDSE, como descrito anteriormente, a informação obtida pela tipagem desta proteína ou do seu gene, tem o potencial de fornecer informações clínicas e epidemiológicas relevantes.

O gene *emm*, à semelhança do que acontece em *S. pyogenes*, está presente em virtualmente todas as estirpes de SDSE e apresenta polimorfismo substancial na sua porção terminal 5' (118). Assim, a base deste método de tipagem é a amplificação por PCR e subsequente sequenciação desta região do gene, que codifica a porção N-terminal da proteína M, responsável pela especificidade serológica (17,99,119). As variações na sequência de nucleótidos da porção terminal 5' do gene *emm* – nomeadamente nos 180pb correspondentes a 10 codões do péptido sinal seguidos dos primeiros 50 codões da proteína M madura – permitem classificar as sequências obtidas de diferentes estirpes em tipos *emm* (7,120,125). O “Centers for Disease Control and Prevention” (CDC) fornece os protocolos para a tipagem *emm* e uma base de dados de referência para atribuição de uma designação de tipo *emm* para a sequência submetida (<https://www.cdc.gov/streplab/groupa-strep/index.html>).

Mais de 90 tipos *emm* foram anteriormente identificados em SDSE (6,99). Alguns estudos reportaram associações significativas entre certos tipos *emm* e doença invasiva ou não invasiva, tal como previamente descrito em *S. pyogenes* (3,29,88,120).

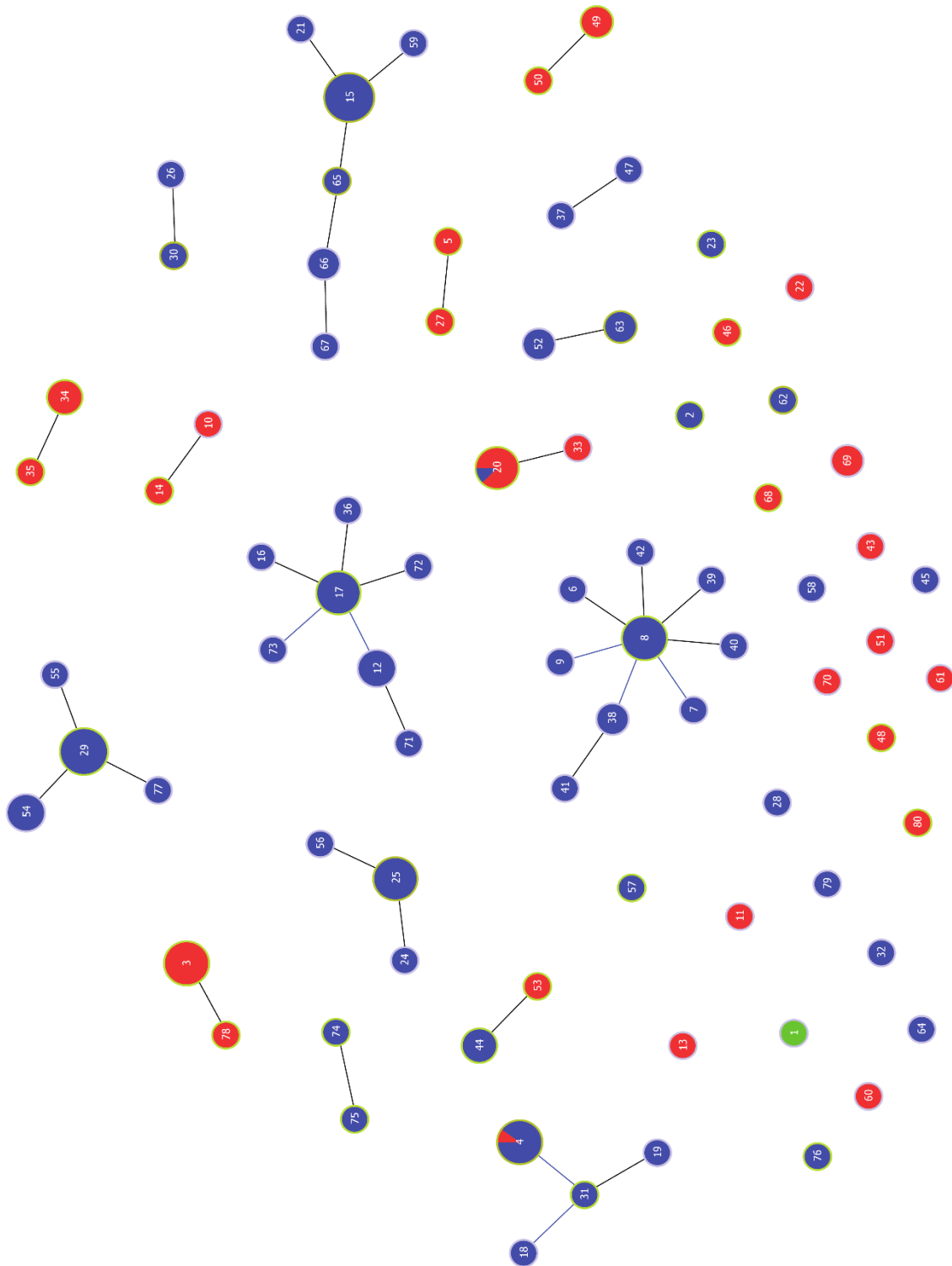
### 1.9.2 “Multilocus sequence typing” (MLST)

A MLST, desenvolvida em 1998, analisa a variação nas sequências de nucleótidos de fragmentos internos de genes “housekeeping” que, como codificam proteínas com funções celulares básicas essenciais, estão presentes em todas as estirpes de uma determinada espécie bacteriana (142–144).

As sequências nucleotídicas obtidas para cada um dos *loci* usados num esquema de MLST são comparadas com uma base de dados com alelos conhecidos (132,143–145). A cada sequência de cada *locus* é atribuído um número de alelo e a combinação dos números dos vários *loci* forma o perfil alélico da estirpe, ao qual é atribuído um tipo de sequência

(“sequence type”, ST) (132,143–145). Utilizando algoritmos de agrupamento, como o eBURST (144) ou o goeBURST (143), é possível definir as relações entre vários STs, baseadas no número de diferenças do perfil alélico, agrupando-os em complexos clonais (CCs) (143–145). Para cada CC é identificado um potencial ST fundador, definido como o ST que difere apenas num gene em relação ao maior número de STs incluídos nesse CC (145). O ST fundador corresponde ao genótipo que aumenta de frequência na população e a partir do qual ocorre diversificação gradual (por mutações ou recombinação) (143,144). Estes STs fruto da diversificação são chamados de “single locus variants” (SLVs), quando diferem em apenas um dos sete *loci* em relação ao fundador e “double locus variants” (DLVs), “triple locus variants” (TLVs), quando diferem dois e três, respectivamente, e assim sucessivamente (143,144) (**Figura 4**).

O atual esquema de MLST de SDSE foi estabelecido em 2009 e utiliza sete genes “housekeeping”: “glucose kinase” (*gki*), “glutamine transporter protein” (*gtr*), “glutamate racemase” (*murI*), “DNA mismatch repair protein” (*mutS*), “transketolase” (*recP*), “xanthine phosphoribosyl transferase” (*xpt*) e “acetyl-CoA acetyltransferase” (*atoB*, também chamado de *yqiZ*) (136,137). A utilização deste esquema para estudar a estrutura populacional de SDSE evidenciou a disseminação internacional de algumas linhagens clonais de SDSE, algumas das quais associadas com múltiplos tipos *emm*, e mostrou que a recombinação intra e interespecíes envolvendo os genes “housekeeping” ocorre com frequência (6,120,137).



**Figura 4:** Exemplo de diagrama utilizando o algoritmo goeBURST.

Relações entre 178 estirpes de SDSE isoladas em diferentes regiões geográficas. STs que pertencem ao mesmo CC estão unidos por linhas retas e diferem entre si em apenas um dos setes *loci*, isto é, são SLVs.

O tamanho de cada círculo é proporcional ao número de estirpes com esse ST específico numa escala logarítmica. Círculos azuis representam estirpes do grupo G, círculos vermelhos estirpes do grupo C e o círculo verde uma estirpe do grupo L.

Adaptado de (137).

### 1.9.3 “Whole genome sequencing” (WGS)

A utilização da WGS, tem sido, cada vez mais, adotada com o surgimento das novas tecnologias de sequenciação de alto débito (140,146,147). Estas tecnologias têm um alto poder discriminatório, superior ao de outras técnicas de tipagem que só analisam uma região específica do genoma, permitindo a diferenciação entre estirpes geneticamente muito próximas (140,146–149). Com recurso a diversas ferramentas de bioinformática consegue-se obter a partir dos dados de WGS as informações de tipagem que têm sido conseguidas através das técnicas mencionadas anteriormente, como o perfil alélico MLST ou o tipo *emm* (150). Além disso, pode-se investigar a presença de genes específicos de virulência e de resistência aos antibióticos e adquirir novos conhecimentos sobre a transmissão de doenças (146,147,149–151). Através destas informações, é possível compreender a evolução dos microrganismos e estudar as disseminações de linhagens clonais (140,150) e, quando aplicada em tipagem epidemiológica, a vigilância e investigação de surtos (148,149).

Com a sequenciação de alto débito do genoma é possível desenhar esquemas de genoma core MLST (“core genome MLST”, cgMLST), que fornecem uma maior resolução em relação ao MLST tradicional (152). A cgMLST tem o mesmo princípio da MLST, mas faz uma comparação dos genes do genoma core (isto é, genes que estão presentes em todas as estirpes dentro de um grupo taxonómico) entre estirpes da mesma espécie (152,153).

No “National Center for Biotechnology Information” (NCBI) estão disponíveis 19 sequências completas do genoma de *S. dysgalactiae*, dos quais 15 pertencem a SDSE (<https://www.ncbi.nlm.nih.gov/genome/genomes/823>, acessido a 28 de janeiro de 2020). O primeiro genoma completo de SDSE foi reportado em 2011, pertencente a uma estirpe do grupo G de Lancefield, causadora de STSS (104).

O genoma de SDSE consiste num cromossoma circular único de tamanho médio de 2,16398Mb (2,0764Mb – 2,2699Mb) e conteúdo G+C médio de 39,4% (<https://www.ncbi.nlm.nih.gov/genome/823>, acessido a 28 de janeiro de 2020) (103,104). Análises das sequências completas dos genomas de SDSE mostraram que estes têm uma sobreposição de cerca de 90% uns com os outros e que parte considerável da diversidade observada se deve à presença de profagos (103). Para além disto, têm uma sobreposição de 61-72% com os genomas de *S. pyogenes* (17,88,104).



## 2. OBJETIVOS

---

O aumento das infeções invasivas por SDSE descrito em vários locais do mundo e, conseqüentemente, a consciencialização do papel patogénico deste microrganismo reforçaram a importância da sua deteção e vigilância. A aplicação de métodos de sequenciação genómica de alto débito possibilitará uma análise mais detalhada da diversidade clonal das estirpes de SDSE responsáveis por infeção invasiva, bem como dos fatores presentes nos genomas dessas estirpes.

Com este estudo, propõe-se:

- Realizar a caracterização genómica das estirpes de SDSE responsáveis por infeção invasiva em Portugal;
- Identificar as linhagens clonais predominantes em infeção invasiva no país;
- Identificar as principais diferenças de conteúdo genético entre as linhagens clonais detetadas, com particular ênfase nos fatores que possam contribuir para diferenças na sua prevalência, virulência e resistência a agentes antimicrobianos;
- Comparar as estirpes invasivas de SDSE com as estudadas anteriormente em Portugal e noutras regiões geográficas, através dos dados obtidos com outros métodos de tipagem – tipagem do gene *emm* e MLST – já descritos na literatura;
- Analisar a congruência entre os diferentes métodos de tipagem;
- Determinar a suscetibilidade das estirpes invasivas de SDSE aos antimicrobianos.



#### 3.1 Estirpes bacterianas

A coleção utilizada neste estudo incluiu 316 estirpes de SDSE isoladas de infeções invasivas no Homem em Portugal. As estirpes de infeção invasiva, definidas no presente trabalho como o isolamento de uma estirpe de SDSE de um produto biológico habitualmente estéril, foram isoladas a partir do sangue (n=289, 91,5%), líquido sinovial (n=19, 6,0%), líquido ascítico (n=6, 1,9%), líquido cefalorraquidiano (n=1, 0,3%) e de biópsia óssea (n=1, 0,3%). As estirpes foram isoladas entre janeiro de 2011 e dezembro de 2017, com a seguinte distribuição anual: 28 em 2011, 35 em 2012, 38 em 2013, 56 em 2014, 54 em 2015, 45 em 2016 e 60 em 2017.

As estirpes foram isoladas e identificadas em 24 laboratórios hospitalares portugueses (**Tabela A1 - Anexos**) e enviadas ao Instituto de Microbiologia da Faculdade de Medicina da Universidade de Lisboa, onde foi realizada toda a caracterização posterior. Apenas uma estirpe por doente foi incluída no estudo.

#### 3.2 Condições de crescimento e morfologia das colónias

As estirpes foram inoculadas em placas de “Tryptone Soya Agar” (TSA) (Oxoid, Basingstoke, Reino Unido) suplementadas com 5% de sangue desfibrinado de carneiro (Probiológica, Lisboa, Portugal) e incubadas a 35 °C em atmosfera não controlada. Após 24 horas de incubação, foi observada a morfologia das colónias para determinação do tamanho (confirmação visual de diâmetro superior a 0,5 mm) e do tipo de hemólise das mesmas.

Todas as estirpes foram armazenadas a –80 °C em “Tryptic Soy Broth” (TSB) (bioMérieux, Marcy-l'Étoile, França) suplementado com 15% (V/V) de glicerol (Sigma-Aldrich, St. Louis, Missouri, EUA), a partir de culturas bacterianas com 24 horas de crescimento nas condições acima descritas.

### 3.3 Classificação de Lancefield

O grupo de Lancefield foi determinado usando uma técnica de aglutinação com látex comercial, “Oxoid Streptococcal Grouping” (Oxoid, Basingstoke, Reino Unido), de acordo com as instruções do fabricante. Todas as estirpes foram testadas com os soros dos grupos de Lancefield C, G e A. As estirpes para as quais não se obteve aglutinação por esta técnica foram testadas com o soro do grupo L (Statens Serum Institut, Copenhaga, Dinamarca).

### 3.4 Teste de suscetibilidade aos antimicrobianos

A suscetibilidade das estirpes de SDSE a um conjunto de 12 antimicrobianos foi testada através do método de Kirby-Bauer (154), de acordo com as normas do “Clinical and Laboratory Standards Institute” (CLSI) para os estreptococos beta-hemolíticos (89). Foram utilizados discos de penicilina, cefotaxima, vancomicina, eritromicina, clindamicina, tetraciclina, cloranfenicol, gentamicina, estreptomicina, levofloxacina, linezolid e quinupristina-dalfopristina (Oxoid, Basingstoke, Reino Unido). Resumidamente, efetuou-se uma suspensão de colónias em 0,85% de NaCl com uma turbidez semelhante a 0,5 McFarland, a qual foi inoculada uniformemente em placas de Mueller-Hinton (Oxoid, Basingstoke, Reino Unido) suplementadas com 5% de sangue desfibrinado de carneiro (Probiológica, Lisboa, Portugal). Após colocação dos discos, as placas foram incubadas a 35 °C em atmosfera enriquecida com 5% de CO<sub>2</sub>, durante 18 a 24 horas.

Os fenótipos de resistência aos macrólidos foram identificados colocando os discos de eritromicina e clindamicina em proximidade, como indicado nas normas do CLSI (89). Para as estirpes nas quais foi detetada resistência intermédia à eritromicina pelo método de difusão com disco, determinou-se a CIM por Etest® (bioMérieux, Marcy-l'Étoile, França). A deteção de resistência de alto nível aos aminoglicosídeos (gentamicina e estreptomicina) foi realizada de acordo com as recomendações do CLSI para *Enterococcus* spp. (89), devido à inexistência de normas específicas para a deteção nos estreptococos. As estirpes de *Streptococcus pneumoniae* ATCC 49619 e *Enterococcus*

*faecalis* ATCC 29212 foram utilizadas como controlos na realização dos antibiogramas (89).

### 3.5 Tipagem do gene *emm*

A tipagem do gene *emm* foi feita com base nos protocolos disponibilizados pelo CDC (<https://www.cdc.gov/streplab/groupa-strep/emm-typing-protocol.html>) para a extração de ADN, amplificação por PCR e sequenciação de um fragmento do gene *emm*.

Apesar da realização da sequenciação de alto débito às 316 estirpes e posterior definição do tipo *emm* a partir do genoma, fez-se a tipagem do gene *emm* pois, inicialmente, não iria ser realizada a sequenciação de alto débito a todas as estirpes, servindo os resultados desta tipagem para selecionar as estirpes para a sequenciação.

A extração do ADN foi realizada a partir de uma cultura em placa com 24 horas de incubação. Cerca de metade de uma ansa de 10 µL com colónias foi ressuspensa em 300 µL de NaCl 0,85% (m/V) e aquecida a 70 °C durante 15 minutos. Após centrifugação a 14000 rpm durante dois minutos e remoção do líquido sobrenadante, o depósito de células foi ressuspensa em 50 µL tampão TE (10 mM Tris [Sigma-Aldrich, St. Louis, Missouri, EUA], 1 mM EDTA [Sigma-Aldrich, St. Louis, Missouri, EUA], pH 7), contendo 0,6 unidades/µL de mutanolisina (Sigma-Aldrich, St. Louis, Missouri, EUA) e 1,2 µg/µL de hialuronidase (Sigma-Aldrich, St. Louis, Missouri, EUA). Esta suspensão foi colocada a incubar a 37 °C durante 30 minutos e no final aquecida a 100 °C durante 10 minutos. Os lisados celulares preparados deste modo foram armazenados a -20 °C até à realização da PCR.

A PCR consistiu, resumidamente, na adição de 1,5 µl do lisado celular preparado à mistura de PCR contendo 1U polimerase ADN “GoTaq G2 Flexi” (Promega, Madison, WI, United States), 1X “Green GoTaq Flexi Buffer” (Promega, Madison, WI, United States), 1,5 mM MgCl<sub>2</sub> (Promega, Madison, WI, United States), 200 mM desoxirribonucleotídeos fosfatados (Thermo Fisher Scientific, Waltham, MA, United States) e 0,4 mM “primers” (*emm*1: 5’ – TATT(C/G)GCTTAGAAAATTAA – 3’ e *emm*2: 5’ – GCAAGTTCTTCAGCTTGTTT – 3’), num volume final de 50 µL. As condições da reação para amplificação foram as seguintes: 94 °C durante um minuto,

seguido de 30 ciclos de 94 °C durante 15 segundos, 46,5 °C durante 30 segundos e 72 °C durante um minuto e 15 segundos (com um incremento de 10 segundos no tempo de extensão em cada um dos últimos 19 ciclos) e um período de extensão final a 72 °C durante 10 minutos. A reação foi efetuada num termociclador Biometra T gradient (Goettingen, Alemanha). Um volume de 5 µL do produto da reação foi utilizado para resolução por eletroforese em gel de agarose (Bio-Rad, München, Alemanha) a 1% (m/V) em tampão Tris-acetato-EDTA (TAE) 1X (Bio-Rad, München, Alemanha). As condições de separação foram 120V durante 90 minutos. Foi utilizado “1 Kb plus” (Invitrogen, Carlsbad, California) como marcador de peso molecular. Após confirmação da presença de um único fragmento com o tamanho apropriado, os produtos de amplificação foram purificados utilizando o “High Pure PCR Product Purification Kit” (Roche, Mannheim, Alemanha) e enviados para sequenciação, de acordo com o protocolo.

Um tipo *emm* para cada estirpe foi atribuído após a submissão das sequências à base de dados do CDC (<https://www.cdc.gov/streplab/groupa-strep/emm-background.html>).

### **3.6 Sequenciação genómica de alto débito**

#### **3.6.1 Preparação do ADN genómico**

A extração do ADN genómico das 316 estirpes foi realizada usando o “kit” de extração “Invitrogen Purelink Genomic DNA” (Thermo Fisher Scientific Inc., Waltham, Massachusetts, EUA) de acordo com as instruções do fabricante, com pequenas modificações.

Uma suspensão bacteriana foi obtida por inoculação de uma colónia em 5 mL de meio líquido “Todd-Hewitt Broth” (Becton Dickinson, Nova Jérсия, EUA) e incubação a 35 °C durante a noite. Um volume de 1 mL desta suspensão foi centrifugado a 15000 rpm durante cinco minutos e o sobrenadante removido. Ao tampão de lise com lisozima, adicionou-se 75 U de mutanolisa (Sigma-Aldrich, St. Louis, Missouri, EUA) e, após o tratamento com proteinase K (20mg/mL), procedeu-se à adição de 400 µg de RNase (disponibilizada no “kit”) com uma incubação de 1 minuto à temperatura ambiente.

A integridade do ADN foi verificada pela utilização de 5 µl do DNA eluído para resolução por eletroforese em gel de 1% (m/V) de agarose (Bio-Rad, Hercules, Califórnia, EUA) em tampão TBE 0,5X (Bio-Rad, Hercules, Califórnia, EUA). Foi utilizado “1 Kb plus” (Invitrogen, Carlsbad, California) como marcador de peso molecular. A pureza do ADN foi avaliada pela medição, num espectrofotómetro NanoDrop 2000 (Thermo Fisher Scientific Inc., Waltham, Massachusetts, EUA), das razões de absorvância 260/280 nm e 260/230 nm. A concentração do ADN foi determinada usando o “kit Invitrogen Qubit<sup>TM</sup> dsDNA HS” (Thermo Fisher Scientific Inc., Waltham, Massachusetts, EUA). Por fim, a concentração de ADN das amostras foi acertada a 10 ng/µL através da diluição em Tris-HCl 10 mM pH8 (Sigma-Aldrich, St. Louis, Missouri, EUA).

### 3.6.2 Sequenciação, montagem e anotação dos genomas

A sequenciação completa do genoma foi realizada no Instituto Gulbenkian de Ciência – Genomics Unit (Oeiras, Portugal) e a análise bioinformática foi realizada pelo grupo de bioinformática do Instituto de Microbiologia da Faculdade de Medicina da Universidade de Lisboa.

As bibliotecas de sequenciação de alto débito foram preparadas usando o “kit Nextera® XT DNA Library Prep paired-end”, o “kit Index v2” (Illumina®, San Diego, Califórnia, EUA) e sequenciadas no sistema Illumina NextSeq® 500 (Illumina®) usando o “kit NextSeq® 500/550 Mid-Output” (versão 2) (300 ciclos). A avaliação e controlo da qualidade e subsequente montagem *de novo* das “reads paired-end” obtidas foi utilizada a pipeline INNUca (<https://github.com/B-UMMI/INNUca>). A pipeline INNUca (versão 3.1) foi corrida usando a imagem Docker “ummidock/innuca:3.1” (<https://hub.docker.com/r/ummidock/innuca/>), fornecendo como “input” sequências de adaptadores Nextera XT para remoção de adaptadores e um tamanho de genoma previsto de 2,1 Mb. Nesta pipeline, a qualidade das “reads” é avaliada com a ferramenta FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) e, posteriormente, processada com o Trimmomatic (155). A montagem *de novo* das “reads” processadas em genomas rascunho é obtida com o SPAdes (156) e, posteriormente, polida com o Pilon (157).

Os genomas rascunho polidos foram anotados com a pipeline Prokka v1.12 (158) usando a imagem Docker “ummidock/prokka:1.12”

(<https://hub.docker.com/r/ummidock/prokka/>). As ferramentas SignalP (versão 4.1) (159), usado para identificar o péptido sinal nos CDS, e RNAmmer (versão 1.2) (160), usado para identificar ARN ribossomal, foram externamente fornecidos à imagem Docker. O Prokka foi corrido usando uma base de dados de género criada com genomas completos de SDSE, usando os seguintes parâmetros: `--addgenes --usegenus --rfam --rnammer --gram pos --increment 10 --mincontiglen 1 --gcode 11 --kingdom Bacteria --genus Streptococcus --species dysgalactiae`.

### 3.6.3 Análise gene-por-gene e relações genéticas entre as estirpes

A relação genética das 316 estirpes foi determinada com base no esquema de MLST disponível para *S. dysgalactiae* (<https://pubmlst.org/sdysgalactiae/>). Para obter o perfil alélico (genes *gki*, *gtr*, *murI*, *mutS*, *recP*, *xpt* e *atoB*) e o ST correspondente a partir do genoma rascunho de cada estirpe, usou-se o *mlst* (<https://github.com/tseemann/mlst>) presente na pipeline INNUca. Os novos alelos ou perfis identificados foram submetidos à base de dados da PubMLST para atribuição de um número. Para definição dos CCs, usou-se o algoritmo goeBURST implementado no PHYLOViZ (161), incluindo na análise os perfis alélicos de todas as estirpes de *S. dysgalactiae* presentes na base de dados a 1 de dezembro de 2019 (<https://pubmlst.org/sdysgalactiae/>). Os CCs foram definidos ao nível SLV.

A análise gene-por-gene dos genomas rascunho das 316 estirpes invasivas de SDSE foi realizada usando o chewBBACA (162). Um esquema de cgMLST foi criado utilizando as 316 estirpes invasivas de Portugal, 122 estirpes de infeção invasiva descritas num estudo canadiano (98) e 18 estirpes de infeção invasiva isoladas na Noruega (111). Com a mesma ferramenta, foi obtido um perfil alélico para cada estirpe com base em 1481 genes presentes em todas as 456 estirpes incluídas neste esquema. Este perfil foi utilizado para determinar a relação genética das 316 estirpes isoladas em Portugal, validar as partições obtidas na análise por MLST convencional e comparar com as estirpes invasivas isoladas noutros locais geográficos que foram alvo de estudo genómico. Para avaliar as relações genéticas entre estirpes, foi usada uma abordagem baseada em rede, usando uma extensão do algoritmo goeBURST (goeBURST full MST) que gera uma árvore de extensão mínima, implementada no PHYLOViZ online (163,164).

### 3.6.4 Determinantes de resistência e fatores de virulência

Utilizaram-se os dados da sequenciação genômica de alto débito para a determinação de fatores de virulência através da base de dados Virulence Factors Database (VFDB) (165). Adicionalmente, foram pesquisados manualmente, através do algoritmo BLAST no Geneious (versão 8.1.9), os genes de fatores de virulência operão da SLS, *gapC*, *gfbA*, e sistemas reguladores *sil*, *covRS* e *fasCAX*.

Os genes de resistência a antimicrobianos, macrólidos e lincosamidas, tetraciclina e aminoglicosídeos, foram obtidos das bases de dados ARG-ANNOT (166), ResFinder (167) e Comprehensive Antibiotic Resistance Database (CARD) (168). Com o objetivo de definir as mutações presentes nas estirpes de SDSE resistentes e intermédias à levofloxacina, as sequências correspondentes às QRDRs nos genes *gyrA* e *parC*, destas estirpes foram extraídas dos genomas e comparadas com as sequências de estirpes de SDSE resistentes às fluoroquinolonas previamente descritas (83). As sequências foram traduzidas nas respectivas proteínas e alinhadas pelo algoritmo MUSCLE presente no Geneious (versão 8.1.9).

### 3.7 Análise estatística

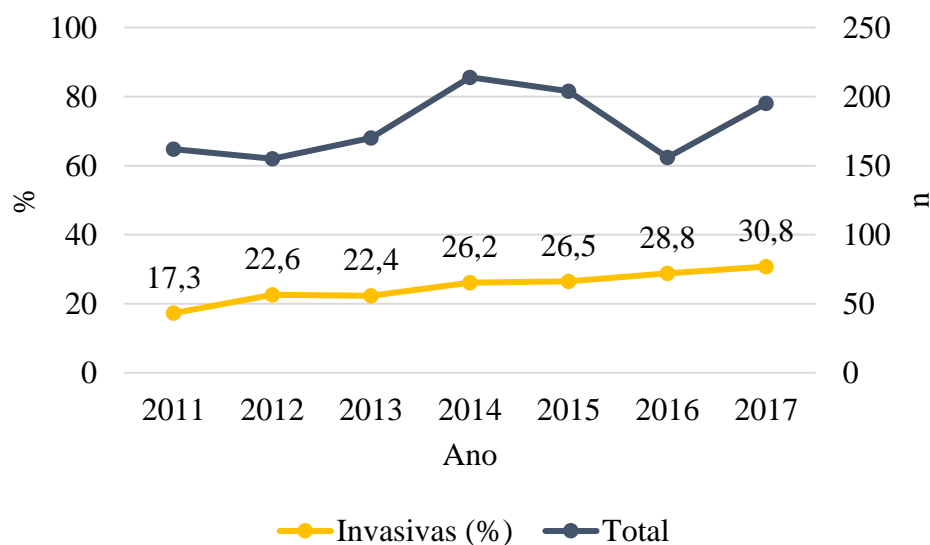
A diversidade da população foi avaliada através do cálculo do índice de diversidade de Simpson (“Simpson’s index of diversity” – SID) e respectivos intervalos de confiança de 95% (IC95%) (<http://www.comparingpartitions.info>). Este índice expressa a probabilidade de duas estirpes selecionadas aleatoriamente da população pertencerem a dois grupos diferentes e pode ser utilizado como uma medida do poder discriminatório dos métodos de tipagem (169). O valor do SID varia de zero a um, sendo que quanto maior for maior é a diversidade da população.

O coeficiente ajustado de Wallace (“Adjusted Wallace” – AW) e respectivos IC95% foram calculados para comparar a congruência entre dois métodos de tipagem (<http://www.comparingpartitions.info>). Este coeficiente compara as partições obtidas por dois métodos de tipagem, indicando a probabilidade dos resultados de um método de tipagem poderem ser previstos por outro (170). O valor do AW varia de zero a um, sendo a congruência dos dois métodos mais alta quanto mais o valor de AW se aproxima de um.

Calcularam-se as *odds ratio* (OR), de modo a determinar associações entre variáveis, e os *p-value* obtidos foram corrigidos usando a taxa de descoberta de falsos positivos (FDR) para testes múltiplos. Estes foram determinados usando as bibliotecas “Epitools” e “multtest” da linguagem “R”. O teste Cochran-Armitage foi usado para avaliar tendências entre duas variáveis. Valores de OR e respectivos IC95% maiores que um e *p-value* < 0,05 mesmo após a correção FDR foram considerados estatisticamente significativos.

### 4.1 Características das estirpes e dos doentes

As 316 estirpes invasivas de SDSE caracterizadas neste estudo representaram 25,2% do total de estirpes de SDSE isoladas durante este período de 7 anos (n=1256), tendo-se observado um aumento da proporção de estirpes invasivas isoladas durante o período do estudo (p=0,001) (**Figura 5**).



**Figura 5:** Distribuição das estirpes de SDSE isoladas de infecção invasiva ao longo do período de estudo (2011-2017).

Total: número total de estirpes (invasivas e não invasivas) em cada ano.  
n: número de estirpes.

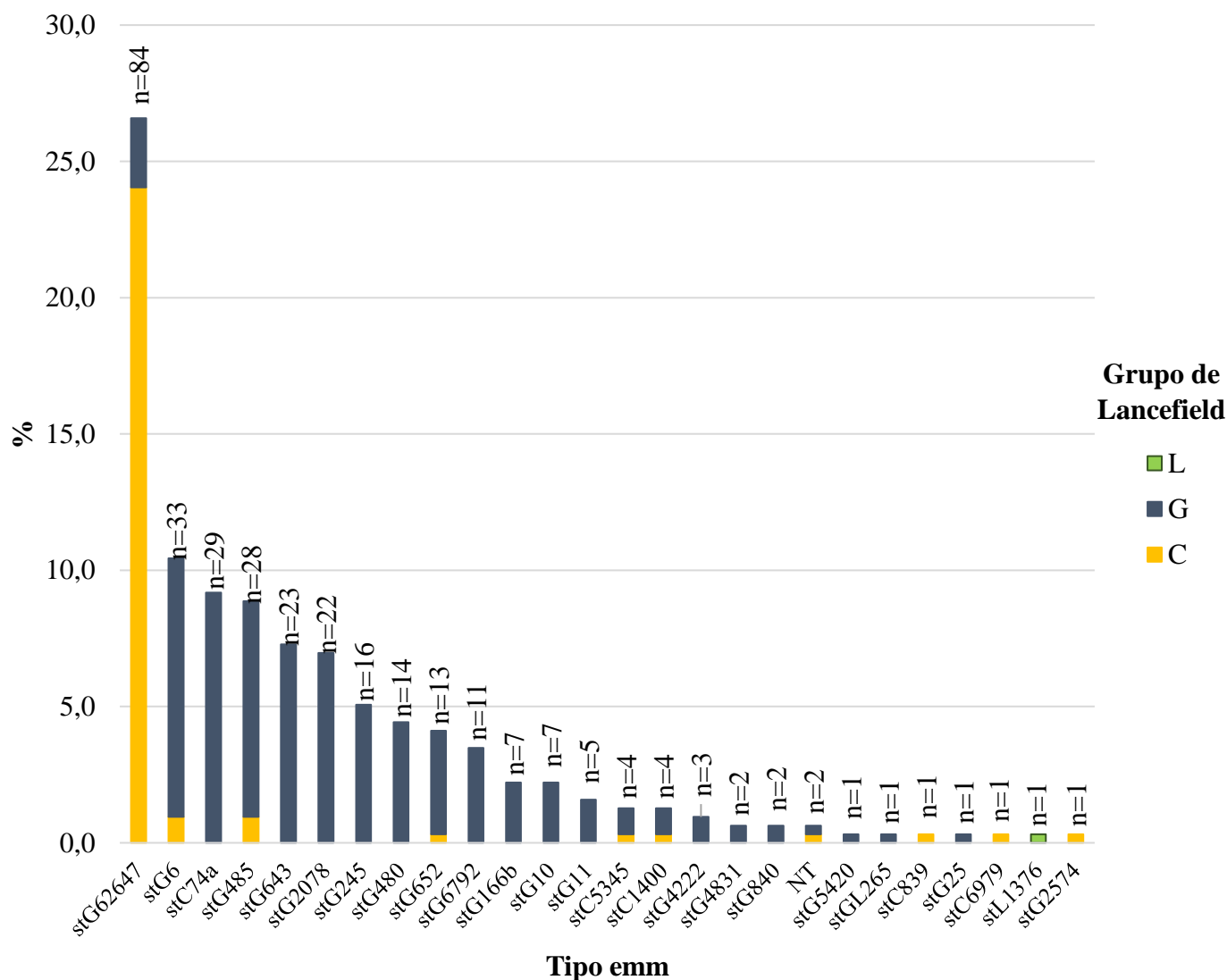
A média de idades dos doentes a partir dos quais as estirpes de SDSE foram isoladas foi de 69,5 anos, tendo variado entre os 5 dias e os 100 anos. Com a exceção de dois doentes, de 5 dias e 3 meses de idade, todos eram adultos a partir dos 33 anos e 90,5% (n=286) tinham idade igual ou superior a 50 anos. As estirpes foram recolhidas maioritariamente de doentes do sexo masculino, 58,2% das estirpes (n=184).

Todas as estirpes incluídas neste estudo formavam colónias >0,5 mm após 24 horas de crescimento e o padrão de hemólise foi, para 99,7% das estirpes (n=315) beta-hemólise, após esse período de incubação. Uma única estirpe apresentou alfa-hemólise após 24

horas de crescimento, tendo sido observado beta-hemólise apenas após incubação prolongada até às 48 horas. Quanto ao grupo de Lancefield, 226 estirpes (71,5%) pertenciam ao grupo G, 89 (28,2%) ao grupo C e uma (0,3%) ao grupo L.

#### 4.2 Tipagem do gene *emm*

A tipagem do gene *emm* das 316 estirpes invasivas resultou em 25 tipos *emm* diferentes e duas estirpes não tipáveis (SID=0,886; IC95%: 0,865-0,907) (**Figura 6**). Os 10 tipos *emm* mais frequentes (>10 estirpes) incluíam quase 90% das estirpes, sendo o tipo *emm stG62647* o mais comum, estando presente em cerca de um quarto das estirpes (n=84, 26,6%). A maioria das estirpes com este tipo *emm* pertencia ao grupo C de Lancefield (n=76; 90,5%), totalizando 85,4% das estirpes com grupo C.

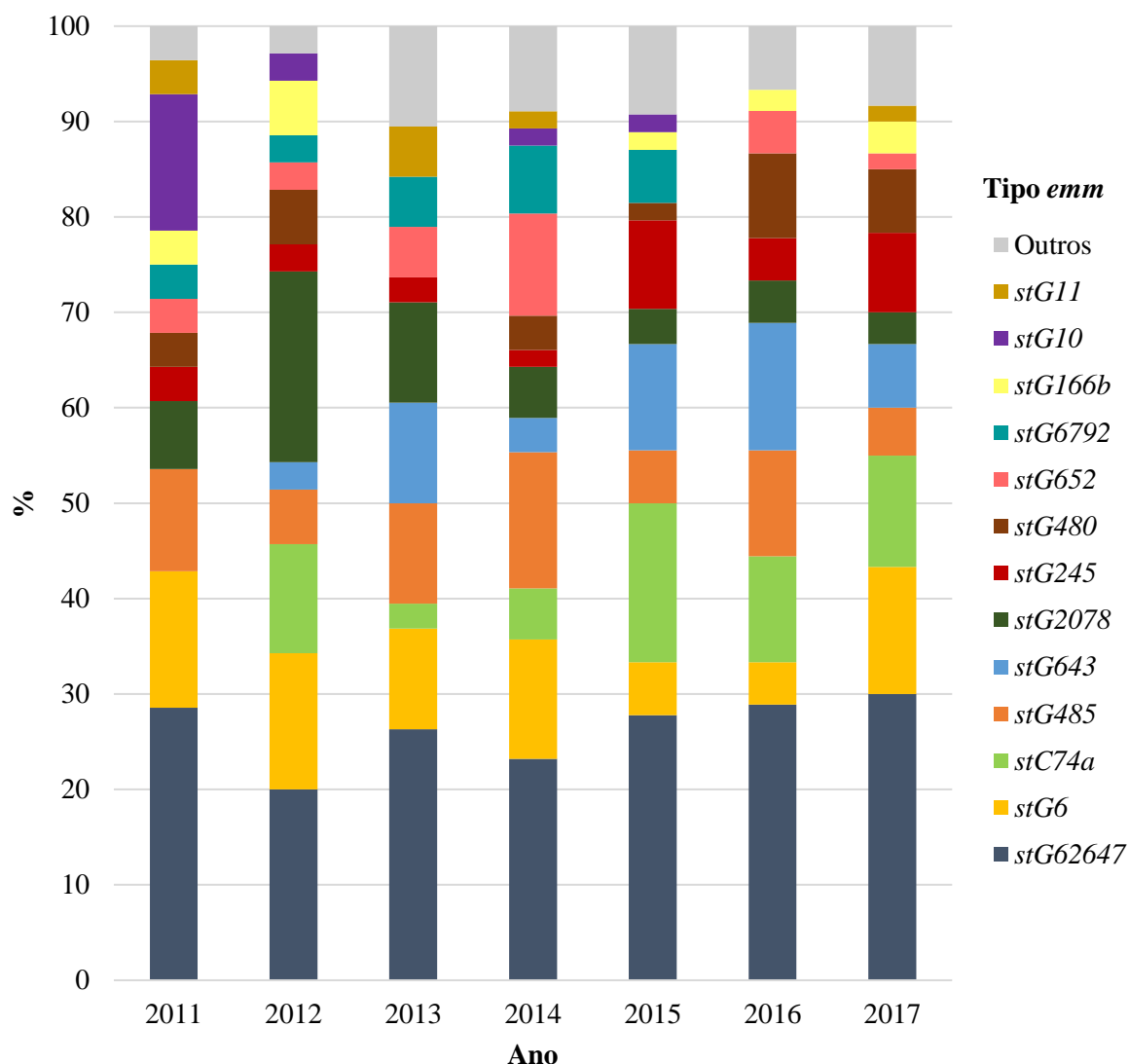


**Figura 6:** Distribuição de tipos *emm* e grupos de Lancefield das 316 estirpes invasivas de SDSE isoladas em Portugal (2011-2017).

NT – não tipável

Seis dos 25 tipos *emm* identificados estavam associados a estirpes com diferentes grupos de Lancefield (AW=0,661; IC95%: 0,499-0,824). Apesar de se ter observado algumas oscilações dos diferentes tipos *emm* ao longo do tempo, a maioria dos tipos *emm* mais frequentes foram detetados em todos os anos do estudo. As estirpes do tipo *emm* *stG62647* foram as mais frequentes em todos os anos do estudo. O tipo *emm* *stG2078*, apesar de ser apenas o sexto mais frequente globalmente, igualou o tipo *emm* *stG62647* no ano de 2012 (ambos com 20% das estirpes) e o tipo *emm* *stG10*, que em 2011 representava quase 15% das estirpes isoladas, decresceu em prevalência ao longo do

tempo ( $p=0,004$ ), não sendo detetadas estirpes com este tipo *emm* nos últimos dois anos (Figura 7).



**Figura 7:** Distribuição de tipos *emm* ao longo do período de estudo (2011-2017).

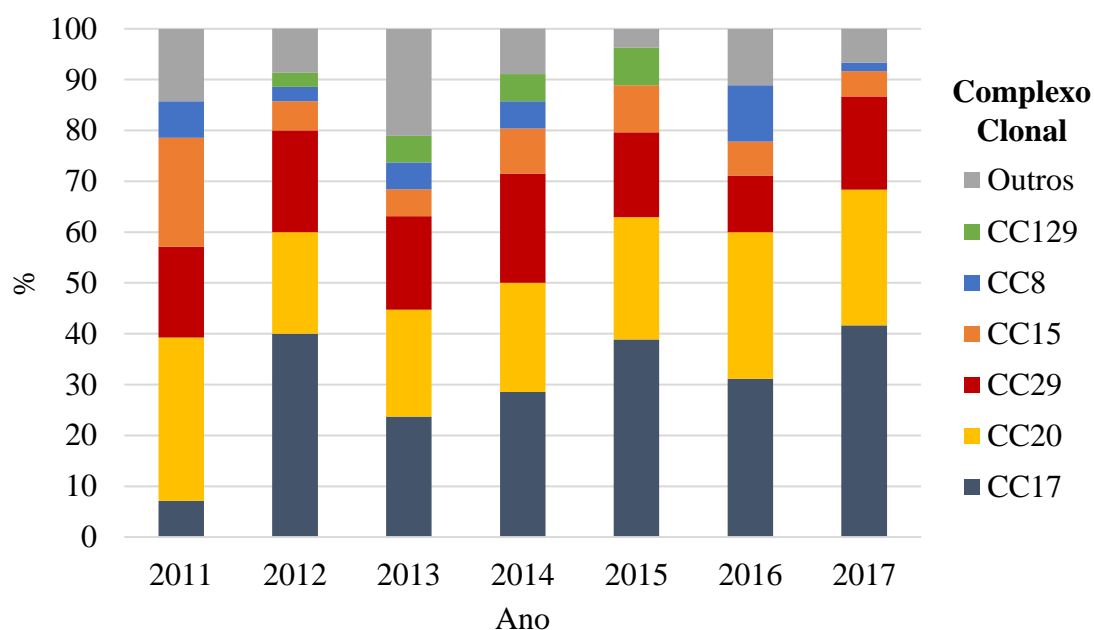
Outros inclui tipos de *emm* com menos de cinco estirpes (*stC5345*, *stC1400*, *stG4222*, *stG4831*, *stG840*, *stG5420*, *stGL265*, *stC839*, *stG25*, *stC6979*, *stL1376*, *stG2574* e não tipáveis).

#### 4.3 “Multilocus sequence typing” (MLST)

A caracterização molecular das 316 estirpes por MLST revelou 76 STs diferentes (SID=0,877; IC95%: 0,852-0,901) agrupados em 16 CCs e oito “singletons” (SID=0,802; CI95%: 0,778-0,826) (Figura 8). Aproximadamente 75% das estirpes estavam incluídas

em três CCs principais (>10% estirpes), CC17 (n=101, 32,0%), CC20 (n=78, 24,7%) e CC29 (n=56, 17,7%).

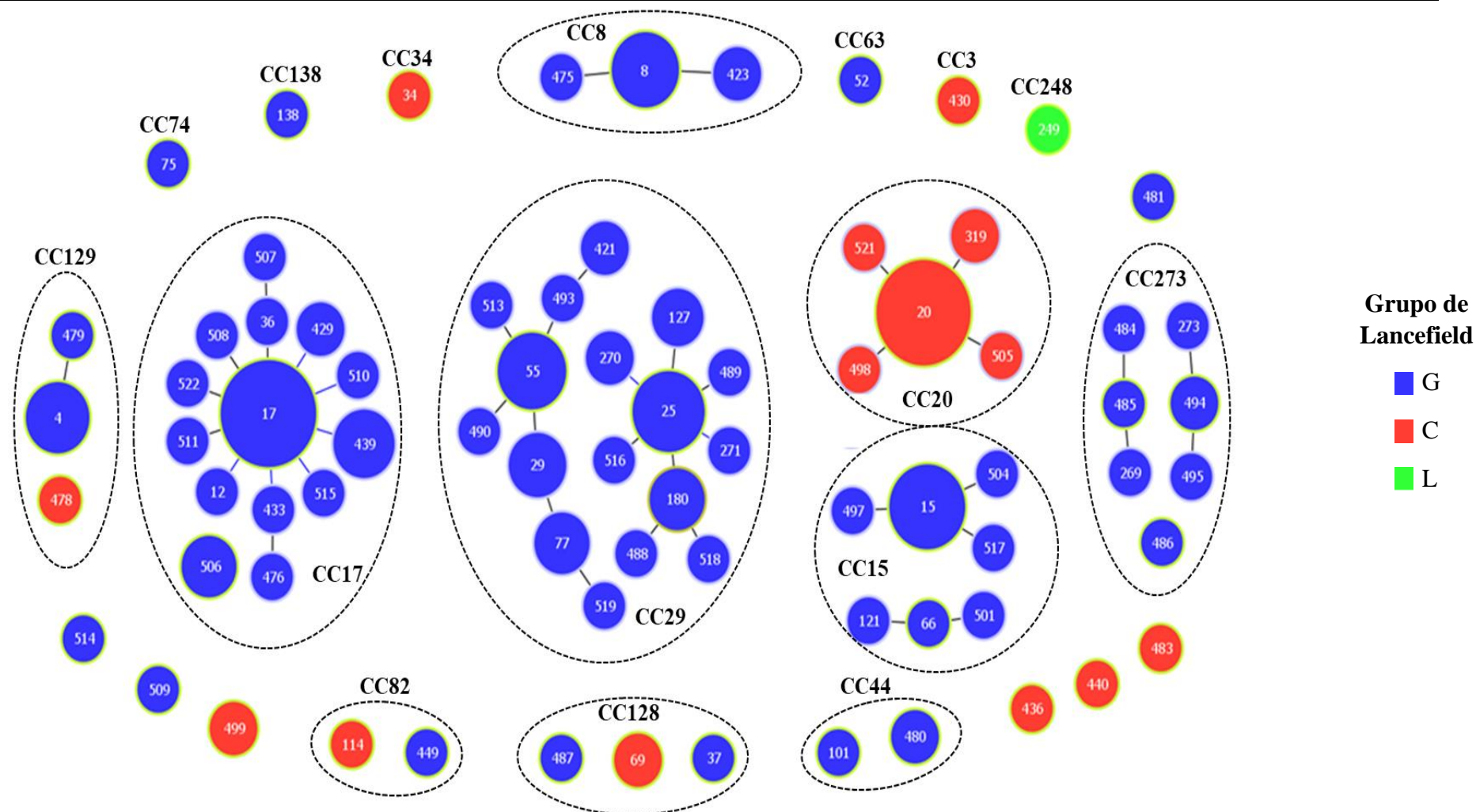
Não houve diferenças significativas na diversidade da população em termos de ST/CC ao longo do período de estudo e a maioria dos principais STs e CCs foram detetados ao longo dos vários anos de estudo (**Figura 8**). O complexo clonal CC17, que incluía apenas 7% das estirpes isoladas em 2011, foi o mais frequente no último ano do estudo (41,7%).



**Figura 8:** Distribuição dos principais CCs ao longo do período de estudo (2011-2017).

Outros inclui CCs com menos de 10 estirpes (CC273, CC128, CC44, CC82, CC248, CC138, CC34, CC75, CC34, CC63 e oito “singletons”).

Cada um dos 76 STs identificados incluía apenas estirpes de um único grupo de Lancefield (AW=1,000; IC95%: 1,000-1,000). Apesar de três dos 16 CCs englobarem estirpes com grupos de Lancefield distintos (AW=0,964; IC95%: 0,943-0,985), não foi detetada nenhuma estirpe de grupo de Lancefield diferente ligada ao nível de SLV entre as 316 analisadas (**Figura 9**).



**Figura 9:** Diagrama goeBURST das relações entre as 316 estirpes invasivas de SDSE com diferentes grupos de Lancefield.

Os CCs foram definidos com base na análise goeBURST usando todos os STs disponíveis na base de dados a 1 de dezembro de 2019 (<https://pubmlst.org/sdysgalactiae/>).

O tamanho dos círculos é proporcional ao número de estirpes com o respectivo ST, numa escala logarítmica.

STs unidos por uma linha são SLVs. STs pertencentes ao mesmo CC estão circundados pela linha tracejada.

**Tabela 2:** Distribuição de tipos *emm* pelos CC definidos por MLST.

<i>Tipo emm</i>	CC	CC17	CC20	CC29	CC15	CC8	CC129	CC273	CC128	CC44	CC82	Outros <sup>1</sup>	Total (nº de estirpes [%])
<i>stG62647</i>		8	76										<b>84 (26,6)</b>
<i>stG6</i>		2	1	23	4							3	<b>33 (10,4)</b>
<i>stC74a</i>		15		13								1	<b>29 (9,2)</b>
<i>stG485</i>		11		12					4			1	<b>28 (8,9)</b>
<i>stG643</i>		18				5							<b>23 (7,3)</b>
<i>stG2078</i>		21									1		<b>22 (7,0)</b>
<i>stG245</i>		10		2	3			1					<b>16 (5,1)</b>
<i>stG480</i>		9			1	3				1			<b>14 (4,4)</b>
<i>stG652</i>				1	7	1		2			1	1	<b>13 (4,1)</b>
<i>stG6792</i>							9	1				1	<b>11 (3,5)</b>
Outros <sup>2</sup>		7	1	5	11	5	1	4		2		7	<b>43 (13,6)</b>
Total (nº de estirpes [%])		<b>101 (32,0)</b>	<b>78 (24,7)</b>	<b>56 (17,7)</b>	<b>26 (8,2)</b>	<b>14 (4,4)</b>	<b>10 (3,2)</b>	<b>8 (2,5)</b>	<b>4 (1,3)</b>	<b>3 (0,9)</b>	<b>2 (0,6)</b>	<b>14 (4,4)</b>	<b>316 (100,0)</b>

<sup>1</sup>Outros inclui seis CCs (CCs 3, 34, 63, 75, 138 e 248) com uma única estirpe identificada no conjunto das 316 estirpes analisadas e oito estirpes pertencentes a STs “singletons”.

<sup>2</sup>Outros inclui 15 tipos de *emm* com menos de 10 estirpes (*stG166b*, *stG10*, *stG11*, *stC5345*, *stC1400*, *stG4222*, *stG4831*, *stG840*, *stG5420*, *stGL265*, *stC839*, *stG25*, *stC6979*, *stL1376*, *stG2574* e as duas estirpes não tipáveis).

A análise por goeBURST incluindo não só as estirpes invasivas de Portugal, mas também todas as estirpes de *S. dysgalactiae* presentes na base de dados (<https://pubmlst.org/sdysgalactiae/>) – para definição dos CCs – mostrou que duas das 316 estirpes, com ST249 (grupo L de Lancefield) e ST483 (grupo C de Lancefield) agruparam juntamente com estirpes isoladas de animais presentes na base de dados num grupo distinto do que incluía a vasta maioria das estirpes isoladas de humanos, sugerindo que estas estão mais relacionadas com as estirpes animais do que com as estirpes humanas (**Figura A1 – Anexos**).

#### 4.4 Correlação entre tipo *emm*, STs e CCs

A maioria dos STs e CCs mais frequentes incluía estirpes de diferentes tipos *emm*, pelo que tanto o ST como o CC revelaram uma baixa capacidade de prever o tipo *emm* de uma estirpe ( $AW_{ST \rightarrow \text{tipo } emm} = 0,448$ ; IC95%: 0,408-0,487 e  $AW_{CC \rightarrow \text{tipo } emm} = 0,288$ ; IC95%: 0,259-0,318) (**Tabela 2 e Figura 10**). As estirpes pertencentes aos complexos clonais CC17 (SID=0,874; IC95%: 0,851-0,897) e CC15 (SID=0,858; IC95%: 0,792-0,925) foram as que apresentaram a maior diversidade de tipos *emm*, em linha com a variedade de tipos *emm* associados ao ST fundador deste CC (12 tipos *emm* nas estirpes ST17 e sete tipos *emm* nas estirpes ST15, respetivamente). Pelo contrário, as estirpes pertencentes aos complexos clonais CC20 (SID=0,051; IC95%: 0,000-0,120) e CC129 (SID=0,200; IC95%: 0,000-0,504) apresentaram a mais baixa diversidade de tipos *emm*, dado que, no caso do CC20, apenas duas das 73 estirpes não tinham o tipo *emm stG62647* e, no caso do CC129, apenas uma das 10 estirpes não apresentava o tipo *emm stG6792*.

Todos os principais tipos *emm* identificados neste estudo foram detetados em estirpes com STs de diferentes CCs ( $AW_{\text{tipo } emm \rightarrow ST} = 0,487$ ; IC95%: 0,395-0,579 e  $AW_{\text{tipo } emm \rightarrow CC} = 0,551$ ; IC95%: 0,454-0,648). Os tipos *emm stG6* e *stG652* foram os encontrados em maior número de CCs diferentes (seis CCs cada, SID=0,506; IC95%: 0,307-0,704 e SID=0,718; IC95%: 0,475-0,961, respetivamente), enquanto os tipos *emm stG62647*, *stG643* e *stG2078* foram, dentro dos principais aqueles associados com menor número de CCs (dois CCs cada, SID=0,174; IC95%: 0,071-0,278, SID=0,356; IC95%: 0,161-0,550 e SID=0,091; IC95%: 0,000-0,252, respetivamente).

A linhagem clonal CC20/*stG62647* (n=76), incluía o maior número de estirpes de um determinado CC com o mesmo tipo *emm*, sendo que a maioria das estirpes apresentavam o mesmo ST (ST20). Todas as estirpes *stG62647* pertencentes a CC20 apresentavam o grupo C de Lancefield, enquanto as estirpes deste tipo *emm* com o grupo G de Lancefield agruparam num CC distinto (CC17) (**Figura 10**). Os tipos *emm* mais prevalentes estavam associados aos quatro CCs principais, CC20, CC17, CC29 e CC15, com exceção do tipo *emm stG6792* associado maioritariamente ao CC129 (**Tabela 2**).



#### 4.5 cgMLST

A relação entre as estirpes invasivas foi avaliada com base num esquema cgMLST contendo 1481 genes. A separação das estirpes na árvore de extensão mínima gerada revelou boa concordância geral com os CCs definidos por MLST convencional, com raras exceções (**Figura 11**). Três estirpes pertencentes ao complexo clonal CC15, com os STs 66, 121 e 501, não agruparam juntamente com as restantes estirpes deste CC, não obstante serem DLVs (STs 66 e 501) e TLVs (ST121) de ST15 (**Figura 11**). De igual modo, a análise por cgMLST não deu suporte à inclusão das duas estirpes com STs 114 e 449 no complexo clonal CC82, dado que, apesar de geneticamente distantes, ambas apresentavam menos diferenças para uma estirpe do CC273 do que entre si mesmas (**Figura 11 e Tabela 3**). Além destes casos, foi possível identificar um número limitado de estirpes dentro dos CCs definidos por MLST com uma distância muito superior ao dos restantes membros do CC (**Figura 11 e Tabela 3**). A análise do cgMLST permitiu detetar a proximidade genética entre alguns CCs e “singletons” não evidenciados pela análise MLST, como foi o caso de uma estirpe com ST509 que apresentava apenas 33 diferenças de uma estirpe pertencente ao CC17 (**Figura 11 e Tabela A2 - Anexos**). A distância média entre as estirpes incluídas no CC20 foi baixa, em consonância com a uniformidade em termos de tipo *emm* e a baixa diversidade de STs incluídos neste complexo clonal (**Tabela 3**).

**Tabela 3:** Diferenças alélicas por cgMLST.

CC	Número de estirpes	Média diferenças	Maior diferença	Menor diferença	Menor diferença para outro CC
CC17	101	223,2	771	4	948* <sup>1</sup>
CC20	78	89,3	259	2	748* <sup>2</sup>
CC29	56	703,2	1134	5	253
CC15	26	387,5	1169	18	748
CC8	14	255,5	482	7	1093
CC129	10	239,4	889	50	943
CC273	8	677,0	957	227	796
CC128	4	258,8	311	136	259* <sup>3</sup>
CC44	3	540,0	749	124	980* <sup>4</sup>
CC82	2	901,0	901	901	796

\*<sup>1</sup>Diferença menor para um “singleton” (33).

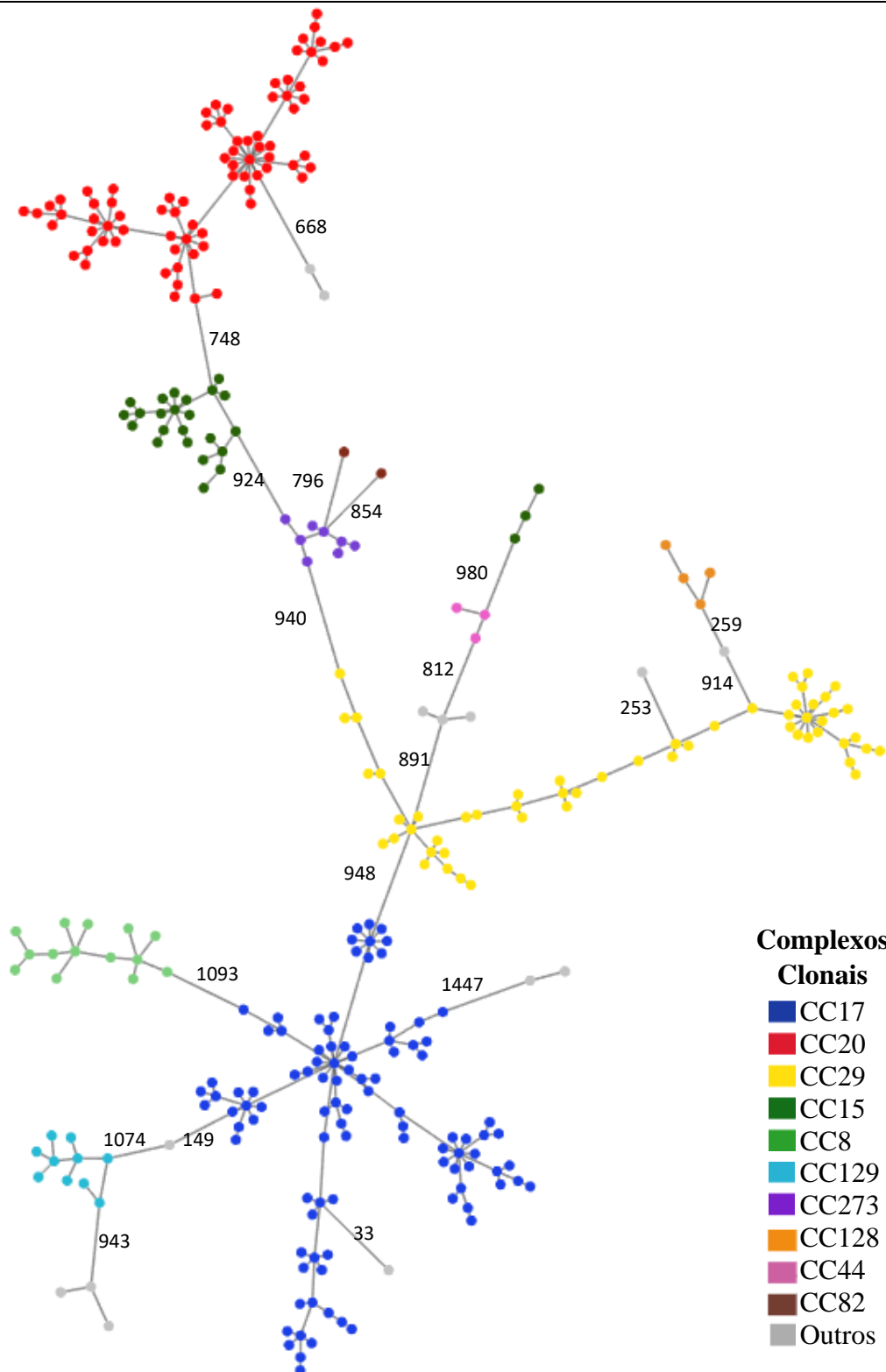
\*<sup>2</sup>Diferença menor para um “singleton” (668).

\*<sup>3</sup>A única ligação é a um “singleton”.

\*<sup>4</sup>Diferença menor para um “singleton” (812).

As estirpes potencialmente de origem animal identificadas pela análise por MLST, com os STs 249 e 483, apresentaram o maior número de diferenças para as restantes estirpes. A estirpe do grupo L (ST249) tinha alelos diferentes em 1477 dos 1481 *loci* incluídos no esquema (99,7%) para a estirpe de origem humana mais próxima (CC17, ST17) e diferia da estirpe ST483 em 1415 *loci*.

A árvore de extensão mínima construída incluindo 122 estirpes de infecção invasiva isoladas no Canadá (98) e 18 estirpes *stG62647* isoladas na Noruega (111), mostrou a ocorrência dos mesmos CCs nos diferentes países e confirmou a proximidade genética das estirpes CC20 *stG62647* do grupo C de Lancefield isoladas nos três países (**Figura A2 – Anexos**).



**Figura 11:** Árvore de extensão mínima das 316 estirpes invasivas isoladas em Portugal com base no esquema cgMLST.

Representa a distribuição dos CCs definidos por MLST.

Outros inclui seis CCs (CCs 3, 34, 63, 75, 138 e 248) com uma única estirpe identificada no conjunto das 316 estirpes analisadas e oito estirpes pertencentes a STs “singletons”.

Os números indicados nos ramos da árvore indicam a distância expressa em número de diferenças entre duas estirpes.

#### 4.6 Fatores de virulência

Com os resultados da sequenciação de alto débito foi possível detetar a presença de genes que codificam fatores de virulência nas 316 estirpes.

Os genes *emm*, *fbp54*, *hasC*, *gapC* e o operão da SLS foram detetados em todas as 316 estirpes (**Tabela 4**). A presença de um gene *emm* foi confirmada nas duas estirpes que eram não tipáveis por PCR. A estirpe ST15 não tipável possui um gene *emm* com uma sequência similar ao das estirpes *stG652* do mesmo ST. No entanto, possui uma deleção de 7 pb na região do péptido sinal que coincide com o local de ligação de um dos “primers” utilizado na amplificação por PCR, justificando assim a incapacidade de obter uma amplificação por PCR nesta estirpe. Acresce ainda a presença de uma segunda deleção, de 138 pb, que inclui parte da região utilizada para definir o tipo *emm* desta estirpe. Dada a impossibilidade de obter um tipo *emm* pelo método convencional (PCR) optou-se por classificar esta estirpe como não tipável. A outra estirpe não tipável, com ST483 e presumivelmente de origem animal, possui um gene anotado como “proteína M”, mas os locais de ligação dos “primers” não são conservados e uma pesquisa por BLAST não evidencia semelhanças com os genes detetados nas outras estirpes invasivas.

O operão da SLS foi detetado em todas as estirpes, incluindo na estirpe que apresentava um fenótipo de alfa-hemólise após 24 horas de crescimento. Pesquisou-se por alterações no operão desta estirpe através da comparação da sequência de nucleótidos com outras estirpes do mesmo ST apresentando beta-hemólise, mas não foram encontradas diferenças.

Os genes *lmb*, *ska*, *slo* e *nga* estavam presentes em todas as estirpes, exceto nas duas presumivelmente de origem animal (**Tabela 4**). Os genes *gbs0628*, *gbs0630*, *gbs0631*, *gbs0632*, *scpA*, *scpB*, *gfbA*, *speG* e *mf3* estavam presentes num número variável de estirpes (**Tabela 4**). Não se detetou a presença de qualquer outro gene de exotoxinas pirogénicas além do gene *speG* acima referido.

**Tabela 4:** Genes de virulência das 316 estirpes.

<b>Gene de Virulência (Produto)</b>	<b>Número de estirpes (%)</b>	<b>Origem da sequência*</b>
<b><i>fbp54</i></b> (proteína de ligação à fibronectina)	316 (100,0%)	VFDB
<b><i>gapC</i></b> (gliceraldeído-3-fosfato desidrogenase)	316 (100,0%)	X97788
<b><i>hasC</i></b> (foforilase glucose UDP)	316 (100,0%)	VFDB
<b>Operão da SLS</b> (estreptolisina S)	316 (100,0%)	VFDB
<b><i>lmb</i></b> (proteína de superfície de ligação à laminina)	314 (99,4%)	VFDB
<b><i>nga</i></b> (NAD-glicohidrolase)	314 (99,4%)	VFDB
<b><i>ska</i></b> (precursor da estreptoquinase A)	314 (99,4%)	VFDB
<b><i>slo</i></b> (precursor da estreptolisina O)	314 (99,4%)	VFDB
<b><i>gbs0630</i></b> (proteína não caracterizada)	291 (92,1%)	VFDB
<b><i>gbs0631</i></b> (proteína não caracterizada)	291 (92,1%)	VFDB
<b><i>gbs0632</i></b> (proteína não caracterizada)	291 (92,1%)	VFDB
<b><i>scpB</i></b> (peptidase C5a estreptocócica)	284 (89,9%)	VFDB
<b><i>speG</i></b> (precursor da exotoxina G estreptocócica)	177 (56,0%)	VFDB
<b><i>gfbA</i></b> (proteína de ligação à fibronectina)	150 (47,5%)	AJ605744
<b><i>scpA</i></b> (precursor da C5a peptidase)	48 (15,2%)	VFDB

<b><i>mf3</i></b> (desoxirribonuclease)	11 (3,5%)	VFDB
<b><i>gbs0628</i></b> (proteína não caracterizada)	3 (0,9%)	VFDB

\* VFDB: “Virulence Factors Database” ; X97788: “Accession number” NCBI.

O gene *speG* estava presente em todas as estirpes pertencentes aos CC20, CC15, CC8 e CC129, e uniformemente ausente das estirpes do CC17. O gene *scpA* estava significativamente associado aos CCs 17 ( $p=0,04$ ) e 129 ( $p<0,001$ ) e o gene *scpB* a CC20 ( $p<0,001$ ). Os genes *gbs0630*, *gbs0631* e *gbs0632* estavam presentes em todas as estirpes com *stG62647* ( $p=0,001$ ), estando o último associado ao CC17 ( $p=0,04$ ). O gene *mf3* estava estatisticamente associado com o CC129 ( $p<0,001$ ).

Os sistemas reguladores *covRS*, *fasCAX* e *sil* foram igualmente pesquisados nas 316 estirpes por BLAST. Genes do *covRS* (“accession number” AY075106) e do *fasCAX* (“accession number” AY075107) estavam presentes nas 316 estirpes. O conjunto de genes do regulador *sil*, de *silA* a *silE*, foi integralmente detetado em apenas 109 estirpes (34,5%). Em 207 estirpes a análise do genoma não permitiu confirmar a presença ou a integridade do conjunto de genes que compõe este regulador. Destas últimas, salientaram-se 103 estirpes em que nenhum dos genes foi detetado, incluindo a totalidade das 101 estirpes pertencentes ao CC17. Em 77 estirpes, todas pertencentes ao CC20, o gene *silB* encontrava-se dividido em dois “contigs”, com tamanho dos fragmentos consentâneo com a interrupção previamente descrita para as estirpes deste CC (111). Uma transposase de sequência idêntica à descrita por Oppegaard *et al* (111) estava presente no genoma de todas estas estirpes, em “contigs” em que foi a única ORF detetada. Em 14 estirpes apenas um fragmento parcial de *silE* foi detetado, incluindo todas as estirpes do CC273 e a estirpe CC20 que, tal como as restantes estirpes deste CC, não possuía o regulador completo. Dez das 56 estirpes CC29 apresentavam o *silE* dividido em dois *contigs*. Em resumo, estes dados apontam para a inexistência de um regulador *sil* funcional em todas as estirpes dos CCs 17, 20 e 273.

#### 4.7 Teste de suscetibilidade a antimicrobianos

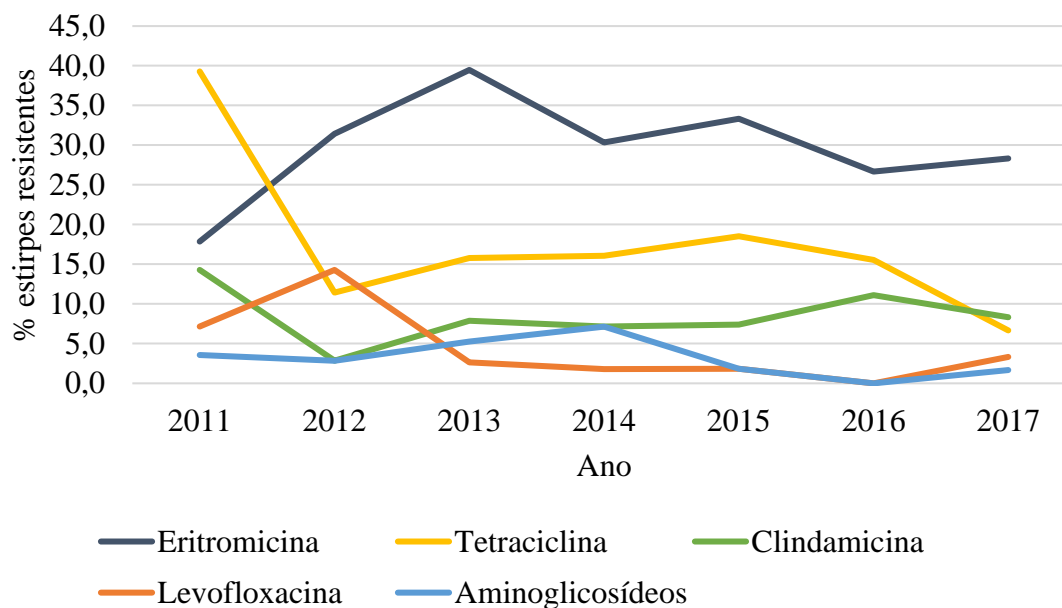
As 316 estirpes analisadas neste estudo eram suscetíveis à penicilina, à cefotaxima, à vancomicina, ao cloranfenicol, à linezolide e à quinupristina-dalfopristina. No total, 134 (42,4%) eram resistentes a, pelo menos, um antibiótico e 17 (5,4%) resistentes a, pelo menos, três antibióticos. A suscetibilidade das 316 estirpes aos antimicrobianos testados encontra-se resumida na **Tabela 5**.

**Tabela 5:** Estirpes resistentes e intermédias aos antimicrobianos.

Antimicrobiano	Padrão de Resistência [nº de estirpes (%)]		
	Resistente	Intermédio	Suscetível
Eritromicina	95 (30,1) <sup>1</sup>	0 (0,0)	221 (69,9)
Tetraciclina	51 (16,1)	21 (6,7)	244 (77,2)
Clindamicina	26 (8,2)	0 (0,0)	290 (91,8)
Levofloxacina	12 (3,8)	3 (1,0)	301 (95,3)
Estreptomicina	9 (2,9)	0 (0,0)	307 (97,2)
Gentamicina	1 (0,3)	0 (0,0)	315 (99,7)

<sup>1</sup>Inclui duas estirpes classificadas como intermédias no teste de difusão em disco, mas com valores de CIM determinados por Etest que as classificam como resistentes à eritromicina (4 µg/mL).

A evolução das resistências aos antimicrobianos ao longo do período em estudo está ilustrada na **Figura 12**. Observou-se uma redução da taxa de resistência à tetraciclina ao longo do tempo ( $p=0,01$ ), de 39,3% ( $n=11$ ) em 2011 para 6,7% ( $n=4$ ) em 2017. A única estirpe resistente à gentamicina foi recuperada em 2011, não tendo sido detetadas estirpes resistentes a este aminoglicosídeo nos anos mais recentes.



**Figura 12:** Evolução das taxas de resistência aos antimicrobianos (2011-2017).

A resistência à eritromicina e à clindamicina foi encontrada em 30,1% (n=95) e 8,2% (n=26) das estirpes, respetivamente. A maioria das estirpes expressou o fenótipo de resistência iMLS<sub>B</sub> (n=67, 70,5%), conferido pelos genes *ermA* (n=64), *ermT* (n=2) e *ermB* (n=1). O fenótipo cMLS<sub>B</sub> foi expresso por 26 estirpes (27,4%), 14 das quais possuíam o gene *ermA* (detetado em conjunto com *lsaC* numa das estirpes) e 12 o gene *ermB* (detetado em conjunto com *lnuD*, *mef* e *msrD* numa das estirpes). Das duas estirpes que expressavam o fenótipo M (2,1%), uma possuía os genes *mefA* e *msrD*, enquanto na outra não foi detetado qualquer gene de resistência aos macrólidos. A resistência à eritromicina estava associada com CC17 (p<0,001), complexo clonal que incluía 56% (n=53) das estirpes resistentes à eritromicina, a maioria (n=46) expressando um fenótipo iMLS<sub>B</sub>. Metade das estirpes com o fenótipo cMLS<sub>B</sub> pertenciam ao CC17 ou ao CC8, o qual estava estatisticamente associado à resistência à clindamicina (p <0,001).

A resistência à tetraciclina foi encontrada em 16,1% das estirpes (n=51) que possuíam maioritariamente o gene *tetM* (n=40), seguido dos genes *tetO* (n=10), *tetL* (n=4, detetado concomitantemente com *tetM*) e *tetW* (n=1). Entre as estirpes com resistência intermédia à tetraciclina, apenas numa se detetou um destes genes (*tetW*), enquanto o gene *tetM* foi detetado em quatro das estirpes suscetíveis à tetraciclina. A resistência à tetraciclina associou-se significativamente com CC15 (p<0,001), que incluía mais de um terço (n=22) de todas as estirpes resistentes à tetraciclina.

As 12 estirpes (3,8%) resistentes à levofloxacina apresentavam diferentes substituições de aminoácidos nas subunidades GyrA e ParC (**Tabela 6**), sendo as mais comuns S81F em GyrA e S79F em ParC. As estirpes com resistência intermédia à levofloxacina (n=3, 1,0%) apresentaram mutações apenas em ParC. A resistência à levofloxacina foi encontrada maioritariamente entre estirpes do CC17 (n=8/12, p<0,001).

**Tabela 6:** Alterações nos aminoácidos de GyrA e ParC de estirpes de SDSE resistentes e intermédias à levofloxacina.

Fenótipo* (nº de estirpes)	Substituições de aminoácidos na posição indicada									
	GyrA					ParC				
	81	85	142	151	79	80	83	91	140	166
R (4)	S81F				S79F			N91D	P140S	
R (2)	E81K				S79F					
R (1)		E85G			S79F					
R (1)	S81F		D142N		S79F					
R (1)	S81F		D142N	E151K	S79F					
R (1)	S81F		D142N				D83V			
R (1)	S81F				S79F			N91D		
R (1)	S81Y					S80P		N91D	P140S	
I (1)							D83G			
I (1)							D83G	N91D	P140S	
I (1)										

\*R=estirpes resistentes; I=estirpes com resistência intermédia.

A resistência aos aminoglicosídeos foi detetada em 10 estirpes (3,2%). Entre as nove estirpes resistentes à estreptomicina foi detetada a combinação dos genes *ant6-Ia*, *aph(3'')-III* e *sata4* (n=6), a combinação dos genes *ant6-Ia* e *aph(3'')-III* (n=1) e o gene *ant6-Ia* isoladamente (n=2). O gene *aac(6')-Ie-aph(2'')-Ia* foi detetado na única estirpe resistente à gentamicina. Uma estirpe suscetível a ambos os aminoglicosídeos tinha presente o gene *aac6-aph2*. Seis CCs distintos e um “singleton” estavam representados nas estirpes resistentes aos aminoglicosídeos.



O aumento de casos de infeções invasivas por SDSE reportado em diferentes estudos tem levado a um maior reconhecimento do potencial patogénico desta bactéria para o hospedeiro humano. Neste trabalho caracterizaram-se as estirpes de SDSE responsáveis por infeção invasiva no Homem, em Portugal, durante um período de sete anos, entre 2011 e 2017. Durante este período foram recolhidas um total de 316 estirpes em vários hospitais do país, maioritariamente de amostras de sangue. A caracterização fenotípica das estirpes incluiu a determinação do tipo de hemólise, do grupo de Lancefield e da suscetibilidade a um conjunto de antimicrobianos. A caracterização genotípica das estirpes envolveu a tipagem do gene *emm* e a sequenciação dos genomas das 316 estirpes, o que permitiu definir as principais linhagens genéticas e estudar a distribuição de fatores de virulência e de determinantes de resistência antimicrobiana na população de SDSE causadora de infeção invasiva em Portugal.

Um aumento na incidência das infeções invasivas causadas por SDSE tem sido notado em diversos países do Mundo nas décadas mais recentes (54–56). Em linha com estas observações, registou-se um aumento da proporção de estirpes de infeção invasiva isoladas em Portugal no âmbito deste trabalho durante o período de estudo, entre 2011 e 2017. As infeções invasivas por SDSE no nosso país atingiram maioritariamente indivíduos do sexo masculino (58,2%) com uma média de idade de 69,5 anos, indicando que a infeção invasiva em Portugal afeta maioritariamente doentes masculinos idosos. Estes resultados são congruentes com o previamente observado noutros países, da Europa (44,54,56,59), Ásia (22,61,66) e América do Norte (4,171). Apesar de um maior reconhecimento do potencial patogénico de SDSE poder contribuir para uma melhor monitorização das infeções causadas por esta bactéria e alguns autores sugerirem que o aumento na incidência das infeções invasivas poderá, em parte, advir de um maior número de hemoculturas realizadas aos doentes (47,68,171), os motivos para este aumento não são claros. As infeções invasivas por SDSE incidem maioritariamente nos idosos (49,66,69) e naqueles que apresentam comorbidades (4,32,60,66,70), pelo que a sobrevivência prolongada de indivíduos com condições médicas subjacentes pode igualmente ser um fator importante (32,47,56,99). Tomando em consideração o envelhecimento das populações observados em Portugal e na generalidade dos países

mais industrializados, é natural que este aumento reflita uma crescente população de hospedeiros particularmente suscetíveis a SDSE e que a tendência de aumento da incidência destas infecções se mantenha. Apesar da possível contribuição dos fatores acima descritos, o aumento das infecções por SDSE pode igualmente refletir um aumento da virulência destas bactérias. Neste contexto, o aparecimento e disseminação de linhagens clonais com um potencial patogénico acrescido, mediado pela presença de fatores de virulência específicos, poderá ter um papel importante (32,47,98).

As estirpes de SDSE estudadas pertenciam, maioritariamente, ao grupo G de Lancefield (72,2%), seguido do grupo C (27,5%) e L (0,3%). Estes resultados são concordantes com os reportados em estudos anteriores, em Portugal (29) e em muitos outros países (4,54,55,66,92), que indicam que a maior parte das estirpes de SDSE isoladas de humanos possuem o grupo G de Lancefield, enquanto as estirpes de SDSE isoladas de animais apresentam predominantemente os grupos C e L (14). De acordo com esta observação, a caracterização genotípica da única estirpe com o grupo L isolada mostrou uma maior proximidade genética com as estirpes de SDSE isoladas habitualmente de animais, sugerindo que esta possa ter uma origem zoonótica. Durante o período de estudo não foram detetadas estirpes de SDSE com o grupo A de Lancefield, em linha com a presença rara do carbohidrato específico deste grupo de Lancefield em estirpes de SDSE, o qual parece restrito a uma linhagem clonal pouco prevalente em infeção humana (172–174).

### **5.1. Caraterização molecular**

A tipagem *emm* tem sido o método de tipagem mais amplamente utilizado na caracterização epidemiológica das estirpes de SDSE (44,59,88,99,119,175). Deste modo, em muitos estudos anteriores o tipo *emm* é a única informação genotípica disponível e o marcador epidemiológico mais comumente usado para comparar a prevalência de estirpes e linhagens clonais de SDSE em diferentes estudos. Também em Portugal, este é o método de tipagem que permite acompanhar a evolução das estirpes invasivas de SDSE por um período mais longo, fruto da sua utilização em estirpes isoladas desde 1998 (176).

A caracterização das 316 estirpes invasivas isoladas entre 2011 e 2017 mostrou uma diversidade elevada de tipos *emm* ( $SID > 0,8$ ), com a presença de 25 tipos *emm* distintos e apenas duas estirpes não tipáveis. Os tipos *emm* mais frequentes foram, na sua generalidade, idênticos aos principais tipos *emm* identificados nas estirpes invasivas isoladas no país entre os anos 1998 e 2010 (176). No entanto, observaram-se algumas diferenças entre estes dois períodos que sugerem uma possível mudança nalgumas linhagens clonais circulantes em Portugal, à imagem do reportado noutras localizações geográficas.

O tipo *emm stG62647* destacou-se claramente como o mais frequente ao estar presente em mais de um quarto das estirpes de infeção invasiva (26,6%). Este tipo *emm*, encontrado em somente 3,5% de todas as estirpes de SDSE isoladas em Portugal até 2005 (83), tornou-se no tipo *emm* mais prevalente no país a partir do ano 2007 (176). De igual modo, estudos da Finlândia (59), Noruega (111) e dos EUA (4), que abrangeram a primeira década dos anos 2000, reportaram uma baixa prevalência deste tipo *emm* entre estirpes de SDSE isoladas de infeção invasiva. O surgimento de uma linhagem de SDSE com o tipo *emm stG62647* do grupo C de Lancefield foi reportada num estudo norueguês, o qual identificou, no período entre 2013 e 2015, uma prevalência deste tipo *emm* em mais de 20% das estirpes, tornando-se no tipo *emm* predominante entre as estirpes invasivas de SDSE isoladas naquela região (111). Em período idêntico, este tipo *emm* foi identificado como um dos mais frequentes num estudo realizado na Suécia (177), sugerindo um aparecimento recente e simultâneo de estirpes com este tipo *emm* nos países nórdicos.

O primeiro estudo a descrever a vigilância epidemiológica das infeções por SDSE em Portugal, abrangendo o período entre 1998 e 2004, identificou dois tipos *emm*, *stG2078* e *stG10* estatisticamente associados a infeção invasiva (29). Embora o tipo *emm stG2078* se tenha mantido como um dos mais frequentes entre estirpes de infeção invasiva, o presente estudo notou um decréscimo na prevalência das estirpes *stG10* ( $p=0,004$ ) até não serem detetadas nos últimos dois anos. As estirpes *stG10* isoladas em Portugal estavam associadas a um único clone de PFGE (29) e, no presente estudo, foram encontradas exclusivamente no CC15. Apesar de não ser possível fazer a correspondência direta entre os resultados obtidos por PFGE e MLST nos dois estudos, a clonalidade das estirpes com este tipo *emm* evidenciada pelos dois métodos de tipagem sugere que a não deteção de

estirpes *stG10* em anos mais recentes se deve a um decréscimo na prevalência de uma linhagem clonal específica.

Relativamente aos restantes tipos *emm* mais comuns identificados no nosso estudo, embora se assemelhem aos reportados como mais frequentes por outros países, como EUA (4), Japão (66), Finlândia (59) ou França (44), surgem, normalmente, numa distribuição diferente, que varia com a região. O *stG6* pertence aos três tipos de *emm* mais frequentes em estudos de diversos países, sendo o mais frequente nos EUA (4) e em França (44). Os tipos *emm stC74a*, *stG485*, *stG2078*, *stG245* e *stG652* encontram-se habitualmente entre os mais frequentes, mas surgem em posições muito variáveis entre estudos. Os tipos *emm stG643* e *stG480* foram descritos no passado como os mais frequentes num estudo norueguês (68) e finlandês (59), respetivamente. Vários estudos realizados no Japão identificaram o tipo *emm stG6792* como o mais frequente em infeção invasiva no país (66), situação que contrasta com o observado em Portugal e noutros países europeus, dado que apenas num estudo, de França (44), a presença deste tipo *emm* entre os tipos *emm* mais frequentes foi notada.

Apesar da grande utilização da tipagem *emm* na caracterização epidemiológica das estirpes de SDSE, foram relatadas algumas limitações. Desde logo por ser um alvo fraco para se obter a informação evolutiva, pois sendo um gene que codifica um proteína de superfície sofre uma grande pressão seletiva e diversificadora (132). A acrescentar a este aspeto, foi descrita a ocorrência de recombinação genética envolvendo este *locus* (7,120). Assim, a tipagem epidemiológica de SDSE deve ser complementada com outros métodos, como a MLST. A MLST utiliza um esquema de genes “housekeeping” que se consideram evoluir lentamente, sob pouca pressão seletiva e diversificadora, permitindo inferir as relações evolutivas dentro da população de SDSE. Esquemas contendo todos os genes “core” presentes numa coleção têm sido aplicados recentemente a esta metodologia, denominada por cgMLST, aumentando a sua resolução (152). Estudos prévios que utilizaram simultaneamente a tipagem *emm* e a MLST mostraram uma fraca correlação entre os dois métodos de tipagem, na medida que o mesmo tipo *emm* ocorre em linhagens diferentes e uma linhagem definida por MLST pode apresentar múltiplos tipos *emm* (6,136,137). No presente estudo, utilizaram-se os dados obtidos na sequenciação de alto débito para avaliar a diversidade genética e identificar os clones responsáveis por infeção invasiva, mediante a utilização de esquema de MLST e cgMLST.

A tipagem das 316 estirpes por MLST resultou em 76 STs diferentes agrupados em 15 CCs e oito *singletons*, resultando numa elevada diversidade ( $SID > 0,8$ ) tanto a nível de STs como CCs, que se manteve ao longo do estudo e em linha com a diversidade de tipos *emm* detetados. No entanto, a grande maioria das estirpes agrupou num número restrito de CCs, dado que os seis CCs mais frequentes (CC17, CC20, CC29, CC15, CC8 e CC129) perfizeram mais de 90% das estirpes, o que indica que um número limitado de linhagens clonais de SDSE é responsável pela maioria dos casos de infeção invasiva humana em Portugal. Resultados similares tinham sido obtidos num estudo que comparou estirpes de três continentes (137) e num estudo do Japão, naquele que foi o estudo que caracterizou mais estirpes invasivas de SDSE por MLST até à data (6).

O CC mais frequente em Portugal, CC17 (n=101, 32,0%), apesar de incluir apenas estirpes do grupo G de Lancefield, estava associado a diversos tipos *emm* ( $SID=0,874$ ;  $IC95\%: 0,851-0,897$ ), dos quais se destacavam *stG2078* e *stG643*. Este CC foi igualmente identificado como o mais frequente em estirpes invasivas no estudo Japonês acima referido e num outro do Canadá (6,98). No Japão, onde CC17 contabilizou mais de 40% das estirpes invasivas de SDSE isoladas entre 2010 e 2013, as estirpes pertencentes a este CC apresentavam na sua grande maioria o tipo *emm stG6792*, (6), enquanto nenhuma das estirpes isoladas em Portugal tinha este tipo *emm*. Estas observações do mesmo CC associado com diferentes tipos *emm* em diferentes regiões geográficas explicam as diferenças observadas entre os dois países em termos de tipos *emm* prevalentes em infeções invasivas.

O segundo CC mais frequente em Portugal, CC20 (n=78, 24,7%), contrastou com os outros CCs mais frequentes por ser composto por estirpes exclusivamente do grupo C, maioritariamente do tipo *emm stG62647*, o mais frequente (26,6%) do estudo. Os estudos que caracterizaram anteriormente estirpes deste tipo *emm* por MLST – isoladas na Noruega (111), Austrália (137) e Canadá (98) – mostraram que as estirpes com este tipo *emm* ou pertencem a ST20 ou são “singletons” deste ST. Estas observações são congruentes com os resultados do presente estudo, em que as estirpes *stG62647* com o grupo C de Lancefield pertencem ao CC20, embora em Portugal se tenha detetado um número reduzido de estirpes deste tipo *emm* com o grupo G de Lancefield noutra CC.

A grande diversidade de tipos *emm* presentes nas estirpes CC17 e CC15 (e respetivos STs fundadores) sugere que estas linhagens sofreram diversificação considerável desde a sua origem na população. Já a uniformidade de tipo *emm* exibida pelas estirpes CC20 e o

baixo número de STs contidos neste CC está de acordo com a origem recente sugerida para esta linhagem genética pelo estudo norueguês (111). A análise das estirpes por cgMLST, com base em 1481 genes, veio comprovar essa uniformidade das estirpes CC20, na medida em que a distância média entre as estirpes incluídas neste CC foi consideravelmente mais baixa que para os restantes CCs.

Os CCs identificados nas estirpes invasivas de SDSE em Portugal diferiram consideravelmente dos identificados num estudo da China que caracterizou estirpes invasivas e não invasivas (77). De igual modo, CC15 foi o único dos mais frequentes em Portugal com uma representação expressiva no estudo conduzido na Índia que caracterizou estirpes de SDSE isoladas de casos de infecção e colonização da faringe (120).

Apesar destas observações poderem sugerir diferenças na prevalência dos diferentes clones em colonização, infecção não invasiva e infecção invasiva, podem simplesmente ser reflexo de diferenças epidemiológicas entre estirpes isoladas em diferentes regiões geográficas.

## 5.2.Fatores de virulência

Os fatores de virulência e, conseqüentemente, uma virulência aumentada poderão ter um papel importante no aumento da incidência de infeções por SDSE (47,98).

Os genes *fbp54* e *gapC* estavam presentes em todas as estirpes. Além disso, a presença do gene *emm* foi confirmada em todas as estirpes analisadas, em linha com estudos anteriores que apontam para a presença ubíqua deste gene em estirpes de SDSE isoladas de humanos (118). Do operão *hasABC*, necessário para a síntese da cápsula de ácido hialurónico, apenas o *hasC* foi detetado, o que está de acordo com a ausência de cápsula em SDSE (6,33,103,104).

A SLS, codificada pelo gene *sagA*, é a beta-hemolisina das estirpes SDSE, responsável pela atividade beta-hemolítica, idêntica à de *S. pyogenes* (104,105). Esta exotoxina está presente em todas as estirpes SDSE (98,105), como foi verificado neste trabalho. Na estirpe com fenótipo alfa-hemolítico após 24 horas de crescimento não se observaram alterações no operão da SLS. A diferença fenotípica poder-se-á dever a alterações na expressão de um ou mais genes presente no operão.

Os genes *lmb*, *ska*, *slo* e *nga* estavam presentes na grande maioria das estirpes. Estudos do Japão (51,70), China (122,178), Áustria (78) e Canadá (98) também referem a presença destes genes em todas, ou praticamente todas, as estirpes estudadas, sugerindo um possível papel na patogênese das infecções invasivas por SDSE. A observação de que apenas as duas estirpes de origem animal não tinham estes genes sugere que estes possam estar ausentes das estirpes de SDSE habitualmente encontradas nos animais e que sejam particularmente importantes na capacidade das estirpes de SDSE causarem infecções em humanos.

Por outro lado, os genes *gbs0628*, *gb0630*, *gbs0631*, *gbs0632*, *scpA*, *scpB*, *gfbA* e *mf3* surgem num número variável de estirpes. Vários estudos reportam o aparecimento destes num número variável de estirpes, habitualmente em percentagens baixas (40,98,122). Pelo contrário, a presença de *scpA* e *scpB* foram reportados numa percentagem elevada das estirpes em estudos da China (77,122,178), Japão (51,70), Canadá (98), Áustria (78) e Alemanha (40).

O superantigénio *speG*, presente em 56,0% das estirpes, tem sido descrito em percentagens variáveis, mas nunca em todas as estirpes, variando de 35% num estudo proveniente de Itália (88) a 83% reportado pela Noruega (119). A análise do gene *speG* mostrou que este possui uma distribuição clonal, estando uniformemente presente em alguns dos principais CCs. Esta associação com diferentes clones poderá justificar as diferentes percentagens reportadas por estudos de diferentes regiões geográficas, as quais se podem dever às linhagens clonais prevalentes nessas regiões possuírem ou não o gene. O papel deste superantigénio na patogênese da infeção por SDSE permanece dúbio, visto que sucessivos estudos têm falhado na deteção de atividade mitogénica (44,54,119), não estando, possivelmente, associado a infeção invasiva. As estirpes SDSE não possuíam nenhuma outra exotoxina pirogénica além de *speG*. Estes resultados são coerentes com estudos de diversos países que não detetaram outros superantigénios em estirpes SDSE (70,77,88,126,130), apesar de muito esporadicamente terem sido detetadas estirpes com alguns destes genes (40,178). A não deteção destes genes em estirpes SDSE isoladas de doentes com STSS pode indicar a existência de outros fatores com um papel no desenvolvimento desta síndrome ou de superantigénios ainda não identificados.

Os sistemas reguladores *covRS* e *fasCAX* estavam presentes em todas as estirpes. Por outro lado, a presença do sistema regulador *sil* em apenas 34,5% das estirpes não é concordante com estudos que indicam a presença deste regulador em cerca de 80% das

estirpes de SDSE (111). Foi anteriormente proposto que a ausência deste regulador da maioria dos genomas de *S. pyogenes* em oposição com a sua presença em muitas das estirpes de SDSE é uma das razões para uma maior virulência de *S. pyogenes* quando comparado com SDSE (111,112). Esta é o racional para a sugestão de que a ausência de um *silB* funcional nas estirpes CC20 possa originar uma maior virulência das estirpes deste CC e ser responsável pela emergência desta linhagem clonal (111). A caracterização genómica das estirpes invasivas isoladas em Portugal não só sugeriu que o gene *silB* se encontra interrompido em todas as estirpes CC20 analisadas, mas também mostrou que o regulador *sil* se encontra completamente ausente de todas as estirpes CC17, o CC mais prevalente no país. Embora ainda não existam estudos funcionais que comprovem que a ausência do *sil* origina maior virulência destas estirpes, os dados do presente estudo mostram que esta é uma característica dos principais clones de SDSE responsáveis pela infeção invasiva no Homem.

### **5.3.Suscetibilidade a antimicrobianos**

As 316 estirpes deste estudo mostraram-se suscetíveis à penicilina. Em Portugal, à semelhança do que acontece noutros países, a terapêutica de primeira linha para infeções por SDSE é a penicilina, pois, à exceção das descrições de tolerância a este antibiótico e de um único caso de resistência de baixo nível, não foram reportadas resistências significativas (73,74,179). Apesar de a metodologia utilizada no presente estudo não permitir a deteção de estirpes com o nível de resistência à penicilina descrita no estudo dinamarquês (74), o ST de SDSE nas quais foram detetadas as alterações nas PBPs não foi detetado.

À semelhança do que acontece com a penicilina, ainda não foram relatadas resistências à cefotaxima, quinupristina-dalfopristina, vancomicina e linezolid em SDSE (4,44,70,77,79,92,179), o que se verificou nas 316 estirpes desta coleção. A vancomicina é frequentemente utilizada em casos de alergia aos beta-lactâmicos, devido à suscetibilidade uniforme entre estirpes SDSE (3,11,75), tendo sido apenas verificados casos de tolerância no passado (9,76). Apesar de já terem sido reportadas resistências ao cloranfenicol em SDSE no passado (179), todas as estirpes analisadas neste estudo foram igualmente suscetíveis.

A ocorrência de resistência a macrólidos e lincosamidas em SDSE é preocupante pois, à imagem do que sucede com outros estreptococos beta-hemolíticos, são as alternativas recomendadas em casos de doentes com alergia aos beta-lactâmicos. Neste estudo, a taxa de resistência à eritromicina foi de 29,4%, semelhante ao descrito em estudos dos EUA (4), Argentina (92), China (79), Itália (88) e França (44), onde as taxas foram sempre cerca de 20% ou superiores. No entanto, estudos do Canadá (67), da Áustria (78), do Brasil (64) e de Israel (99) reportaram taxas ligeiramente inferiores, abaixo dos 20%, e um da Índia não detetou estirpes resistentes a macrólidos (179). As variações geográficas observadas poder-se-ão dever a diferenças de suscetibilidade nas linhagens clonais presentes nos vários locais, bem como a diferenças no consumo de macrólidos por cada país (180). A elevada taxa de resistência observada em Portugal sugere que o uso de macrólidos como terapêutica alternativa deve ser feito com precaução e controlo.

O fenótipo de resistência a macrólidos e lincosamidas mais frequente foi o iMLS<sub>B</sub> (69,9%), o que difere de vários estudos anteriores. Estudos conduzidos em França (44), na China (77,79) e no Japão (6) detetaram maioritariamente o fenótipo cMSL<sub>B</sub> e um estudo brasileiro (64) identificou o fenótipo M, raro no nosso estudo, como o mais frequente. À semelhança das observações em Portugal, o fenótipo iMLS<sub>B</sub> foi o mais frequente em estudos da Argentina (92) e da Áustria (78). O gene que conferiu resistência à maioria das estirpes com fenótipos MLS<sub>B</sub>, foi o *ermA* (80,65%), resultado congruente com estudos de Espanha (181), França (44), Itália (88), Argentina (92), Brasil (64) e Japão (66), onde a maioria ou todas as estirpes possuíam este gene. No entanto, estudos da China (77,79), Japão (6,182) referem o gene *ermB* como o mais frequente. O gene *ermT*, responsável pela resistência em apenas duas estirpes, aparece como pouco comum também noutros estudos, como na Itália (88). Uma estirpe com fenótipo M tinha o gene *mefA*, semelhante ao descrito em estudos do Japão (6,66,70), China (79) e Argentina (92). A presença de genes de resistência aos macrólidos em estirpes não resistentes já foi reportada por outros estudos, que verificaram os genes *ermA* (183), *ermC* (183–185) e *mefA* (186) em estirpes de cocos de Gram-positivo suscetíveis e intermédias.

As tetraciclinas são a classe de antibióticos à qual têm sido relatadas as maiores taxas de resistência e, por isso, já não são uma opção de tratamento para infeções por SDSE. Em Portugal observou-se uma redução da taxa de resistência a este antimicrobiano ao longo dos anos de estudo. A taxa de resistência encontrada nesta coleção de estirpes (16,1%) foi inferior à reportada pela maioria dos países, como França (44), Itália (88),

Áustria (78), Argentina (92), Brasil (64), Israel (99) e China (77), cujas taxas reportadas foram variaram entre 25 e 60%. Apenas um estudo da Índia (179) descreveu taxas de resistências baixas, 7,55%. A diferença na taxa de resistência é coerente com o menor consumo de tetraciclinas em Portugal em relação aos outros países da Europa (180). Ainda assim, a tendência decrescente observada não é coincidente com nenhuma diminuição no consumo de tetraciclinas em Portugal durante o período do estudo (180).

O gene que mais frequentemente conferiu resistência à tetraciclina foi o *tetM* (66,7%), seguido do *tetO* (19,6%). Este cenário é idêntico ao reportado ao observado em França (44), Itália (88) e Argentina (92). O gene *tetL* surgiu num número diminuto de estirpes que também possuíam o gene *tetM*, situação previamente descrita em estudos que também detetaram *tetL* isoladamente (44). O gene *tetW* apenas foi detetado numa estirpe resistente, não sendo um achado comum em SDSE. Os genes *tetW* e *tetM* foram encontrados em estirpes intermédias e suscetíveis, respetivamente. A presença dos genes *tetM*, *tetO*, *tetS* e *tetK* em estirpes de cocos de Gram-positivo não resistentes já foi notada por vários autores (185–189).

Na maioria dos estudos que reportam taxas de resistência às fluoroquinolonas esta é inferior a 5%, nos EUA (4), China (79) e Japão (6). Neste estudo, a taxa de resistência à levofloxacina foi de 4,1%, que é idêntica às taxas descritas por outros países. Ainda assim, em Portugal, em 2010, foi reportada uma taxa de resistência bastante superior, de 12,1% (83). O aumento da resistência neste período poderá estar associado ao aumento do uso de fluoroquinolonas, ou seja, as diferenças no consumo poderão ser o fator major para as diferenças observadas (83), tanto entre países como entre os dois períodos de estudo em Portugal. Assim, a diminuição da taxa de resistência poderá dever-se a uma diminuição no consumo de fluoroquinolonas, pois Portugal era o país europeu onde esta classe de antibióticos era mais utilizada (83,190,191) e esta utilização tem vindo a diminuir ao longo dos anos, já não estando Portugal entre os países europeus onde o consumo é maior (180,192).

A aquisição de uma mutação em apenas uma das enzimas de ligação das fluoroquinolonas normalmente não confere resistência significativa, sendo as estirpes suscetíveis ou intermédias (80,84). As mutações nas subunidades *gyrA* e *parC*, dos alvos de ligação ADN girase e topoisomerase IV são necessárias e suficientes para conferir resistência (6,80,84). Mutações nas outras subunidades (*gyrB* e *parE*) são menos comuns (80,84). Neste estudo, todas as estirpes resistentes possuíam substituições de aminoácidos

em *gyrA* e *parC* e as estirpes intermédias apenas em *parC*. Substituições na posição 81 da *gyrA* e 79 da *parC* são as mais importantes e frequentes para a aquisição de resistência à levofloxacina, como já tinha sido reportado anteriormente (83). Algumas modificações, nas posições 151 e 80 da *gyrA* e *parC*, respetivamente, ainda não tinham sido descritas em SDSE, bem como uma substituição para um aminoácido diferente na posição 85 da *gyrA* e 83 da *parC*. Certas mutações encontradas já foram previamente identificadas como não associadas a resistência, como é o caso das substituições nas posições 91 e 140 da *parC* (83).

Os aminoglicosídeos, embora possam ser utilizados em combinação com a penicilina para o tratamento de infeções invasivas graves (3,9,11,32,73), raramente são utilizados em estudos epidemiológicos. Apesar de SDSE ser intrinsecamente resistente aos aminoglicosídeos, as resistências de alto nível são pouco comuns (81,84,89,95–97). No presente estudo, as taxas de resistência de alto nível foram de 0,3% e 2,9%, à gentamicina e estreptomicina, respetivamente. Um estudo argentino reportou uma estirpe com resistência de alto nível (81). As baixas taxas de resistência reportadas estão de acordo com a pouca utilização dos aminoglicosídeos, estando Portugal entre os países da Europa em que menos se utilizam e que mostra uma tendência decrescente no consumo desta classe de antimicrobianos (180).

Neste estudo observou-se uma alta resistência a vários antimicrobianos, e algumas estirpes apresentaram multirresistência (n=17, 5,4%). As resistências, embora tenham sido encontradas em várias linhagens, estavam associadas a alguns clones. A resistência à eritromicina e clindamicina encontravam-se associadas aos CC17 e CC8, respetivamente, e a resistência à tetraciclina e levofloxacina ao CC15 e CC17, respetivamente. As resistências aos aminoglicosídeos não se mostraram relacionadas significativamente com nenhum CC. Uma análise mais aprofundada da informação genómica permitirá a identificação dos elementos genéticos móveis associados à resistência.



## 6. CONCLUSÃO

---

SDSE é um microrganismo emergente e com cada vez mais importância como agente patogénico humano, responsável por um largo espectro de manifestações clínicas, desde infeções moderadas e a infeções invasivas graves. O aumento da incidência de infeções invasivas e o escasso conhecimento da epidemiologia e mecanismo de infeção reforçam a importância de estudos epidemiológicos para o estabelecimento de um cenário global e para avaliar tendências.

Neste trabalho caracterizou-se fenotipicamente e genotipicamente uma coleção de 316 estirpes de SDSE responsáveis por infeção invasiva no Homem em Portugal, nos anos de 2011 a 2017, com o objetivo principal de definir as linhagens predominantes e as características e fatores que contribuem para a maior virulência e prevalência dessas.

A utilização da sequenciação de alto débito permitiu elucidar a diversidade genética das estirpes. A comparação entre os resultados obtidos pelos diferentes métodos de tipagem mostrou uma baixa congruência entre o tipo *emm* e os resultados obtidos por MLST, mostrando que os resultados obtidos pela tipagem do gene *emm* são um pobre indicador da proximidade genética das estirpes. A análise por cgMLST sustentou a existência dos CCs definidos por MLST convencional e confirmou a presença de linhagens clonais predominantes, entre as quais se destacam CC17 e CC20, altamente prevalentes em diferentes regiões geográficas.

O aumento da incidência das infeções invasivas por SDSE poderá ser o resultado da presença de diversos fatores de virulência, nem todos identificados, ou, como sugerido, por diferenças na regulação da expressão dos mesmos.

As resistências aos antimicrobianos também foram avaliadas e definidos os determinantes de resistência. A taxa de resistência mais elevada foi verificada para os macrólidos e lincosamidas, o que, dada a importância destes antimicrobianos como opção terapêutica no tratamento das infeções por SDSE, é preocupante.

### 6.1 Perspetivas futuras

Este trabalho é uma contribuição para o conhecimento sobre a infeção invasiva por SDSE em Portugal. Alguns aspetos poderão ser elucidados com trabalho futuro, nomeadamente:

- Estudar estirpes de colonização e infeção não invasiva, utilizando também a sequenciação de alto débito, para avaliar diferenças na distribuição de linhagens clonais entre colonização, infeções não invasivas e infeções invasivas. Pretender-se-á com esta abordagem encontrar diferenças no potencial invasivo de diferentes linhagens, bem como identificar fatores nos genomas dessas linhagens que possam justificar uma distribuição diferencial;
- Expandir a análise dos potenciais fatores de virulência a outros *loci* não analisados neste trabalho, como por exemplo as regiões FCT, tentando avaliar a sua diversidade nos genomas de SDSE;
- Explorar a relação do locus *sil* com a virulência das estirpes de SDSE, nomeadamente através da elucidação de quais os fatores de virulência que possam estar sobre a sua regulação;
- Identificar elementos genéticos móveis, tais como bacteriófagos, e a sua associação com alguns fatores de virulência e determinantes de resistência, para avaliar o seu papel na disseminação dos genes que codificam estas proteínas;
- Continuar a vigilância epidemiológica das infeções por SDSE, a fim de detetar a emergência de novas linhagens clonais e garantir a monitorização das resistências.

## 7. REFERÊNCIAS

---

1. Murray PR, Rosenthal KS, Pfaller MiA. *Medical Microbiology*. 8th ed. Elsevier; 2016.
2. Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, et al., editors. *Manual of Clinical Microbiology*. 11th ed. AMS; 2015.
3. Brandt CM, Spellerberg B. Human infections due to *Streptococcus dysgalactiae* subspecies *equisimilis*. *Clin Infect Dis*. 2009;49(5):766–72.
4. Broyles LN, Van Beneden C, Beall B, Facklam R, Shewmaker PL, Malpiedi P, et al. Population-Based Study of Invasive Disease Due to  $\beta$ -Hemolytic *Streptococci* of Groups Other than A and B. *Clin Infect Dis*. 2009;48(6):706–12.
5. Jensen A, Kilian M. Delineation of *Streptococcus dysgalactiae*, Its Subspecies, and Its Clinical and Phylogenetic Relationship to *Streptococcus pyogenes*. *J Clin Microbiol*. 2012;50(1):113–26.
6. Wajima T, Morozumi M, Hanada S, Sunaoshi K, Chiba N, Iwata S, et al. Molecular characterization of invasive *Streptococcus dysgalactiae* subsp. *Equisimilis*, Japan. *Emerg Infect Dis*. 2016;22(2):247–54.
7. McNeilly CL, McMillan DJ. Horizontal gene transfer and recombination in *Streptococcus dysgalactiae* subsp. *equisimilis*. *Front Microbiol*. 2014;5:676.
8. Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, Siepel A. Replacing and additive horizontal gene transfer in *Streptococcus*. *Mol Biol Evol*. 2012;29(11):3309–20.
9. Efstratiou A. Pyogenic streptococci of Lancefield groups C and G as pathogens in man. *Soc Appl Bacteriol Symp Ser*. 1997;26:72S-79S.
10. Efstratiou A. Outbreaks of human infection caused by pyogenic streptococci of Lancefield groups C and G. *J Med Microbiol*. 1989;29(3):207–19.
11. Bradley SF, Gordon JJ, Baumgartner DD, Marasco WA, Kauffman CA. Group C streptococcal bacteremia: analysis of 88 cases. *Rev Infect Dis*. 1991;13(2):270–80.

12. Farrow JAE, Collins MD. Taxonomic Studies on Streptococci of Serological Groups C, G and L and Possibly Related Taxa. *Syst Appl Microbiol.* 1984;5(4):483–93.
13. Vandamme P, Pot B, Falsen E, Kersters K, Devriese LA. Taxonomic study of lancefield streptococcal groups C, G, and L (*Streptococcus dysgalactiae*) and proposal of *S. dysgalactiae* subsp. *equisimilis* subsp. nov. *Int J Syst Bacteriol.* 1996;46(3):774–81.
14. Vieira V V, Teixeira LM, Zahner V, Momen H, Facklam RR, Steigerwalt AG, et al. Genetic relationships among the different phenotypes of *Streptococcus dysgalactiae* strains. *Int J Syst Bacteriol.* 1998;48 Pt 4:1231–43.
15. Pinho MD, Erol E, Ribeiro-Gonçalves B, Mendes CI, Carriço JA, Matos SC, et al. Beta-hemolytic *Streptococcus dysgalactiae* strains isolated from horses are a genetically distinct population within the *Streptococcus dysgalactiae* taxon. *Sci Rep.* 2016;6:31736.
16. Nishiki I, Yoshida T, Fujiwara A. Complete genome sequence and characterization of virulence genes in Lancefield group C *Streptococcus dysgalactiae* isolated from farmed amberjack (*Seriola dumerili*). *Microbiol Immunol.* 2019;63(7):243–50.
17. Takahashi T, Ubukata K, Watanabe H. Invasive infection caused by *Streptococcus dysgalactiae* subsp. *equisimilis*: characteristics of strains and clinical features. *J Infect Chemother.* 2011;17(1):1–10.
18. Facklam R. What happened to the streptococci: overview of taxonomic and nomenclature changes. *Clin Microbiol Rev.* 2002;15(4):613–30.
19. Sherman JM. The Streptococci. *Bacteriol Rev.* 1937;1(1):3–97.
20. Lancefield RC. A serological differentiation of human and other groups of hemolytic streptococci. *J Exp Med.* 1933;57(4):571–95.
21. Kłós M, Wojkowska-Mach J. Pathogenicity of Virulent Species of Group C Streptococci in Human. *Can J Infect Dis Med Microbiol.* 2017;2017:1–5.
22. Woo PC, Fung AM, Lau SK, Wong SS, Yuen KY. Group G beta-hemolytic streptococcal bacteremia characterized by 16S ribosomal RNA gene sequencing. *J Clin Microbiol.* 2001;39(9):3147–55.

23. Cherkaoui A, Emonet S, Fernandez J, Schorderet D, Schrenzel J. Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for rapid identification of Beta-hemolytic streptococci. *J Clin Microbiol.* 2011;49(8):3004–5.
24. Jensen CS, Dam-Nielsen C, Arpi M. Matrix-assisted laser desorption/ionization-time of flight mass spectrometry identification of large colony beta-hemolytic streptococci containing Lancefield groups A, C, and G. *Infect Dis (London, England).* 2015/05/09. 2015;47(8):575–9.
25. Singhal N, Kumar M, Kanaujia PK, Viridi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis [Internet]. Vol. 6, *Frontiers in Microbiology* . 2015. p. 791.
26. Ciszewski M, Zegarski K, Szewczyk EM. *Streptococcus dysgalactiae* subsp. *equisimilis* Isolated From Infections in Dogs and Humans: Are Current Subspecies Identification Criteria accurate? *Curr Microbiol.* 2016;73(5):684–8.
27. Dawson ED, Taylor AW, Smagala JA, Rowlen KL. Molecular detection of *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subsp. *equisimilis*. *Mol Biotechnol.* 2009;42(1):117–27.
28. McMillan DJ, Vu T, Bramhachari P V, Kaul SY, Bouvet A, Shaila MS, et al. Molecular markers for discriminating *Streptococcus pyogenes* and *S. dysgalactiae* subspecies *equisimilis*. *Eur J Clin Microbiol Infect Dis.* 2010;29(5):585–9.
29. Pinho MD, Melo-Cristino J, Ramirez M. Clonal relationships between invasive and noninvasive Lancefield group C and G Streptococci and emm-specific differences in invasiveness. *J Clin Microbiol.* 2006;44(3):841–6.
30. Hashikawa S, Iinuma Y, Furushita M, Ohkura T, Nada T, Torii K, et al. Characterization of group C and G streptococcal strains that cause streptococcal toxic shock syndrome. *J Clin Microbiol.* 2004;42(1):186–92.
31. Kawata K, Minakami T, Mori Y, Katsumi M, Kataoka Y, Ezawa A, et al. rDNA sequence analyses of *Streptococcus dysgalactiae* subsp. *equisimilis* isolates from pigs. *Int J Syst Evol Microbiol.* 2003;53(Pt 6):1941–6.
32. Rantala S. *Streptococcus dysgalactiae* subsp. *equisimilis* bacteremia: an emerging

- infection. *Eur J Clin Microbiol Infect Dis*. 2014;33(8):1303–10.
33. Watanabe S, Takemoto N, Ogura K, Miyoshi-Akiyama T. Severe invasive streptococcal infection by *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subsp. *equisimilis*. *Microbiol Immunol*. 2016;60(1):1–9.
  34. McDonald M, Towers RJ, Andrews RM, Carapetis JR, Currie BJ. Epidemiology of *Streptococcus dysgalactiae* subsp. *equisimilis* in tropical communities, Northern Australia. *Emerg Infect Dis*. 2007;13(11):1694–700.
  35. Bramhachari P V, Kaul SY, McMillan DJ, Shaila MS, Karmarkar MG, Sriprakash KS. Disease burden due to *Streptococcus dysgalactiae* subsp. *equisimilis* (group G and C streptococcus) is higher than that due to *Streptococcus pyogenes* among Mumbai school children. *J Med Microbiol*. 2010;59(Pt 2):220–3.
  36. Kakuya F, Kinebuchi T, Okubo H, Matsuo K, Kuroda M, Fujiyasu H. Acute Pharyngitis Associated With *Streptococcus dysgalactiae* Subspecies *equisimilis* in Children. *Pediatr Infect Dis J*. 2018;37(6):537–42.
  37. Harrington AT, Clarridge JE. Impact of identification of *Streptococcus dysgalactiae* subspecies *equisimilis* from throat cultures in an adult population. *Diagn Microbiol Infect Dis*. 2013;76(1):20–3.
  38. Zaoutis T, Attia M, Gross R, Klein J. The role of group C and group G streptococci in acute pharyngitis in children. *Clin Microbiol Infect*. 2004;10(1):37–40.
  39. O Luiz FB de, Alves KB, Barros RR. Prevalence and long-term persistence of beta-haemolytic streptococci throat carriage among children and young adults. *J Med Microbiol*. 2019/08/16. 2019;68(10):1526–33.
  40. Jaalama M, Palomaki O, Vuento R, Jokinen A, Uotila J. Prevalence and Clinical Significance of *Streptococcus dysgalactiae* subspecies *equisimilis* (Groups C or G Streptococci) Colonization in Pregnant Women: A Retrospective Cohort Study. *Infect Dis Obstet Gynecol*. 2018;2018:2321046.
  41. Yamaguchi T, Kawahara R, Katsukawa C, Kanki M, Harada T, Yonogi S, et al. Foodborne Outbreak of Group G Streptococcal Pharyngitis in a School Dormitory in Osaka, Japan. *J Clin Microbiol*. 2018;56(5):e01884-17.
  42. Silva LG, Genteluci GL, de Mattos M, Glatthardt T, Sá Figueiredo AM, Ferreira-

- Carvalho BT. Group C *Streptococcus dysgalactiae* subsp. *equisimilis* in south-east Brazil: genetic diversity, resistance profile and the first report of human and equine isolates belonging to the same multilocus sequence typing lineage. *J Med Microbiol.* 2015;64(Pt 5):551—558.
43. Schrieber L, Towers R, Muscatello G, Speare R. Transmission of *Streptococcus dysgalactiae* subsp. *equisimilis* between Child and Dog in an Aboriginal Australian Community. *Zoonoses Public Health.* 2014;61(2):145–8.
44. Loubinoux J, Plainvert C, Collobert G, Touak G, Bouvet A, Poyart C. Adult Invasive and Noninvasive Infections Due to *Streptococcus dysgalactiae* subsp. *equisimilis* in France from 2006 to 2010. *J Clin Microbiol.* 2013;51(8):2724–7.
45. Carmeli Y, Ruoff KL. Report of cases of and taxonomic considerations for large-colony-forming Lancefield group C streptococcal bacteremia. *J Clin Microbiol.* 1995;33(8):2114–7.
46. Barnham M, Kerby J, Chandler RS, Millar MR. Group C streptococci in human infection: a study of 308 isolates with clinical correlations. *Epidemiol Infect.* 1989;102(3):379–90.
47. Rantala S, Vuopio-Varkila J, Vuento R, Huhtala H, Syrjanen J. Clinical presentations and epidemiology of beta-haemolytic streptococcal bacteraemia: a population-based study. *Clin Microbiol Infect.* 2009;15(3):286–8.
48. Tan LKK, Eccersley LRJ, Sriskandan S. Current views of haemolytic streptococcal pathogenesis. *Curr Opin Infect Dis.* 2014;27(2):155–64.
49. Cohen-Poradosu R, Jaffe J, Lavi D, Grisariu-Greenzaid S, Nir-Paz R, Valinsky L, et al. Group G streptococcal bacteremia in Jerusalem. *Emerg Infect Dis.* 2004;10(8):1455–60.
50. Lewthwaite P, Parsons HK, Bates CJ, McKendrick MW, Dockrell DH. Group G streptococcal bacteraemia: an opportunistic infection associated with immune senescence. *Scand J Infect Dis.* 2002;34(2):83–7.
51. Ikebe T, Murayama S, Saitoh K, Yamai S, Suzuki R, Isobe J, et al. Surveillance of severe invasive group-G streptococcal infections and molecular typing of the isolates in Japan. *Epidemiol Infect.* 2004;132(1):145–9.

52. Lindbaek M, Hoiby EA, Lermark G, Steinsholt IM, Hjortdahl P. Clinical symptoms and signs in sore throat patients with large colony variant beta-haemolytic streptococci groups C or G versus group A. *Br J Gen Pract.* 2005;55(517):615–9.
53. Turner JC, Fox A, Fox K, Addy C, Garrison CZ, Herron B, et al. Role of group C beta-hemolytic streptococci in pharyngitis: epidemiologic study of clinical features associated with isolation of group C streptococci. *J Clin Microbiol.* 1993;31(4):808–11.
54. Kittang BR, Bruun T, Langeland N, Mylvaganam H, Glambek M, Skrede S. Invasive group A, C and G streptococcal disease in western Norway: virulence gene profiles, clinical features and outcomes. *Clin Microbiol Infect.* 2011;17(3):358–64.
55. Harris P, Siew D-A, Proud M, Buettner P, Norton R. Bacteraemia caused by beta-haemolytic streptococci in North Queensland: changing trends over a 14-year period. *Clin Microbiol Infect.* 2011;17(8):1216–22.
56. Ekelund K, Skinhoj P, Madsen J, Konradsen HB. Invasive group A, B, C and G streptococcal infections in Denmark 1999-2002: epidemiological and clinical aspects. *Clin Microbiol Infect.* 2005;11(7):569–76.
57. Bert F, Lambert-Zechovsky N. Analysis of a case of recurrent bacteraemia due to group A *Streptococcus equisimilis* by pulsed-field gel electrophoresis. *Infection.* 1997;25(4):250–1.
58. Sylvetsky N, Raveh D, Schlesinger Y, Rudensky B, Yinnon AM. Bacteremia due to beta-hemolytic *Streptococcus* group G: increasing incidence and clinical characteristics of patients. *Am J Med.* 2002;112(8):622–6.
59. Rantala S, Vähäkuopus S, Vuopio-Varkila J, Vuento R, Syrjänen J. *Streptococcus dysgalactiae* subsp. *equisimilis* Bacteremia, Finland, 1995–2004. *Emerg Infect Dis J.* 2010;16(5):843.
60. Hindsholm M, Schonheyder HC. Clinical presentation and outcome of bacteraemia caused by beta-haemolytic streptococci serogroup G. *APMIS.* 2002;110(7–8):554–8.

61. Liao C-H, Liu L-C, Huang Y-T, Teng L-J, Hsueh P-R. Bacteremia Caused by Group G Streptococci, Taiwan. *Emerg Infect Dis*. 2008;14(5):837–40.
62. Wagner JG, Schlievert PM, Assimacopoulos AP, Stoehr JA, Carson PJ, Komadina K. Acute Group G Streptococcal Myositis Associated with Streptococcal Toxic Shock Syndrome: Case Report and Review. *Clin Infect Dis*. 1996;23(5):1159–61.
63. Nei T, Akutsu K, Shima A, Tsuboi I, Suzuki H, Yamamoto T, et al. A case of streptococcal toxic shock syndrome due to Group G streptococci identified as *Streptococcus dysgalactiae* subsp. *equisimilis*. *J Infect Chemother*. 2012;18(6):919–24.
64. de Souza JPJ, Santos AR, de Paula GR, Barros RR. Antimicrobial susceptibility and genetic relationships among *Streptococcus dysgalactiae* subsp. *equisimilis* isolates in Rio de Janeiro. *Infect Dis (London, England)*. 2016;48(9):676–81.
65. Haidan A, Talay SR, Rohde M, Sriprakash KS, Currie BJ, Chhatwal GS. Pharyngeal carriage of group C and group G streptococci and acute rheumatic fever in an Aboriginal population. *Lancet*. 2000;356(9236):1167–9.
66. Takahashi T, Sunaoshi K, Sunakawa K, Fujishima S, Watanabe H, Ubukata K. Clinical aspects of invasive infections with *Streptococcus dysgalactiae* ssp. *equisimilis* in Japan: Differences with respect to *Streptococcus pyogenes* and *Streptococcus agalactiae* infections. *Clin Microbiol Infect*. 2010;16(8):1097–103.
67. Laupland KB, Ross T, Church DL, Gregson DB. Population-based surveillance of invasive pyogenic streptococcal infection in a large Canadian region. *Clin Microbiol Infect*. 2006;12(3):224–30.
68. Oppegaard O, Mylvaganam H, Kittang BR. Beta-haemolytic group A, C and G streptococcal infections in Western Norway: A 15-year retrospective survey. *Clin Microbiol Infect*. 2015;21(2):171–8.
69. Kristensen B, Schønheyder H. A 13-year survey of bacteraemia due to -haemolytic streptococci in a Danish county. *J Med Microbiol*. 1995;43:63–7.
70. Ikebe T, Oguro Y, Ogata K, Katsukawa C, Isobe J, Shima T, et al. Surveillance of severe invasive group G streptococcal infections in Japan during 2002–2008. *Jpn J Infect Dis*. 2010;63(5):372–5.

71. Pillai A, Thomas S, Williams C. Clindamycin in the treatment of group G beta-haemolytic streptococcal infections. *J Infect.* 2005;51(4):e207-11.
72. Noble J, Tyburski M, Berman M, Greenspan J, Tenenbaum M. Antimicrobial tolerance in group G streptococci. *Lancet.* 1980;316(8201):982.
73. Portnoy D, Prentis J, Richards GK. Penicillin tolerance of human isolates of group C streptococci. *Antimicrob Agents Chemother.* 1981;20(2):235–8.
74. Fuursted K, Stegger M, Hoffmann S, Lambertsen L, Andersen PS, Deleuran M, et al. Description and characterization of a penicillin-resistant *Streptococcus dysgalactiae* subsp. *equisimilis* clone isolated from blood in three epidemiologically linked patients. *J Antimicrob Chemother.* 2016;71(12):3376–80.
75. Bonofiglio L, Gaggioli P, Garcia Gabarrot G, Kaufman S, Mollerach M, Toresani I, et al. Susceptibility to beta-lactams in beta-hemolytic streptococci. *Rev Argent Microbiol.* 2018;50(4):431–5.
76. Zaoutis T, Schneider B, Steele Moore L, Klein JD. Antibiotic susceptibilities of group C and group G streptococci isolated from patients with invasive infections: evidence of vancomycin tolerance among group G serotypes. *J Clin Microbiol.* 1999;37(10):3380–3.
77. Lu B, Fang Y, Huang L, Diao B, Du X, Kan B, et al. Molecular characterization and antibiotic resistance of clinical *Streptococcus dysgalactiae* subsp. *equisimilis* in Beijing, China. *Infect Genet Evol.* 2016;40:119–25.
78. Leitner E, Zollner-Schwetz I, Zarfel G, Masoud-Landgraf L, Gehrler M, Wagner-Eibel U, et al. Prevalence of emm types and antimicrobial susceptibility of *Streptococcus dysgalactiae* subsp. *equisimilis* in Austria. *Int J Med Microbiol.* 2015;305(8):918–24.
79. Lo H-H, Nien H-H, Cheng Y-Y, Su F-Y. Antibiotic susceptibility pattern and erythromycin resistance mechanisms in beta-hemolytic group G *Streptococcus dysgalactiae* subspecies *equisimilis* isolates from central Taiwan. *J Microbiol Immunol Infect.* 2015;48(6):613–7.
80. Hooper DC. Fluoroquinolone resistance among Gram-positive cocci. *Lancet Infect*

- Dis. 2002;2(9):530–8.
81. Lopardo HA, Vidal P, Sparo M, Jeric P, Centron D, Facklam RR, et al. Six-month multicenter study on invasive infections due to *Streptococcus pyogenes* and *Streptococcus dysgalactiae* subsp. *equisimilis* in Argentina. *J Clin Microbiol.* 2005;43(2):802–7.
  82. Ikebe T, Okuno R, Sasaki M, Kanda Y, Otsuka H, Kawahara R, et al. Molecular characterization and antibiotic resistance of *Streptococcus dysgalactiae* subspecies *equisimilis* isolated from patients with streptococcal toxic shock syndrome. *J Infect Chemother.* 2018;24(2):117–22.
  83. Pinho MD, Melo-Cristino J, Ramirez M. Fluoroquinolone resistance in *Streptococcus dysgalactiae* subsp. *equisimilis* and evidence for a shared global gene pool with *Streptococcus pyogenes*. *Antimicrob Agents Chemother.* 2010;54(5):1769–77.
  84. Woodford N. Biological counterstrike: antibiotic resistance mechanisms of Gram-positive cocci. *Clin Microbiol Infect.* 2005;11 Suppl 3:2–21.
  85. Roberts MC, Sutcliffe J, Courvalin P, Jensen LB, Rood J, Seppala H. Nomenclature for macrolide and macrolide-lincosamide-streptogramin B resistance determinants. *Antimicrob Agents Chemother.* 1999;43(12):2823–30.
  86. Roberts MC. Update on macrolide-lincosamide-streptogramin, ketolide, and oxazolidinone resistance genes. *FEMS Microbiol Lett.* 2008;282(2):147–59.
  87. Palmieri C, Magi G, Creti R, Baldassarri L, Imperi M, Gherardi G, et al. Interspecies mobilization of an *erm(T)*-carrying plasmid of *Streptococcus dysgalactiae* subsp. *equisimilis* by a coresident ICE of the ICESa2603 family. *J Antimicrob Chemother.* 2012;68(1):23–6.
  88. Gherardi G, Imperi M, Palmieri C, Magi G, Facinelli B, Baldassarri L, et al. Genetic diversity and virulence properties of *Streptococcus dysgalactiae* subsp. *equisimilis* from different sources. *J Med Microbiol.* 2014;63(Pt 1):90–8.
  89. Clinical and Laboratory Standards Institute (CLSI). Performance Standards for Antimicrobial Susceptibility Testing. 28th ed. CLSI supplement M100. 2018.
  90. Chopra I, Roberts M. Tetracycline antibiotics: mode of action, applications,

- molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev.* 2001;65(2):232–60.
91. Roberts MC. Update on acquired tetracycline resistance genes. *FEMS Microbiol Lett.* 2005;245(2):195–203.
  92. Traverso F, Blanco A, Villalon P, Beratz N, Saez Nieto JA, Lopardo H. Molecular characterization of invasive *Streptococcus dysgalactiae* subsp. *equisimilis*. Multicenter study: Argentina 2011-2012. *Rev Argent Microbiol.* 2016;48(4):279–89.
  93. Varaldo PE, Montanari MP, Giovanetti E. Genetic elements responsible for erythromycin resistance in streptococci. *Antimicrob Agents Chemother.* 2008/11/10. 2009;53(2):343–53.
  94. Roberts AP, Mullany P. Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. *FEMS Microbiol Rev.* 2011/07/04. 2011;35(5):856–71.
  95. Vakulenko SB, Mobashery S. Versatility of aminoglycosides and prospects for their future. *Clin Microbiol Rev.* 2003;16(3):430–50.
  96. Ramirez MS, Tolmasky ME. Aminoglycoside modifying enzymes. *Drug Resist Updat.* 2010;13(6):151–71.
  97. Galimand M, Lambert T, Gerbaud G, Courvalin P. High-level aminoglycoside resistance in the beta-hemolytic group G *Streptococcus* isolate BM2721. *Antimicrob Agents Chemother.* 1999;43(12):3008–10.
  98. Lothar SA, Demczuk W, Martin I, Mulvey MR, Dufault B, Lagacé-Wiens P, et al. Clonal Clusters and Virulence Factors of Group C and G *Streptococcus* Causing Severe Infections, Manitoba, Canada, 2012–2014. *Emerg Infect Dis J.* 2017;23(7):1092.
  99. Halperin T, Levine H, Korenman Z, Burstein S, Amber R, Sela T, et al. Molecular characterization and antibiotic resistance of group G streptococci in Israel: comparison of invasive, non-invasive and carriage isolates. *Eur J Clin Microbiol Infect Dis.* 2016;35(10):1649–54.
  100. Sriprakash KS, Hartas J. Lateral genetic transfers between group A and G

- streptococci for M-like genes are ongoing. *Microb Pathog.* 1996;20(5):275–85.
101. Davies MR, Tran TN, McMillan DJ, Gardiner DL, Currie BJ, Sriprakash KS. Inter-species genetic movement may blur the epidemiology of streptococcal diseases in endemic regions. *Microbes Infect.* 2005;7(9–10):1128–38.
  102. Suzuki H, Lefébure T, Hubisz MJ, Pavinski Bitar P, Lang P, Siepel A, et al. Comparative Genomic Analysis of the *Streptococcus dysgalactiae* Species Group: Gene Content, Molecular Adaptation, and Promoter Evolution. *Genome Biol Evol.* 2011;3:168–85.
  103. Watanabe S, Kirikae T, Miyoshi-Akiyama T. Complete genome sequence of *Streptococcus dysgalactiae* subsp. *equisimilis* 167 carrying Lancefield group C antigen and comparative genomics of *S. dysgalactiae* subsp. *equisimilis* strains. *Genome Biol Evol.* 2013;5(9):1644–51.
  104. Shimomura Y, Okumura K, Murayama SY, Yagi J, Ubukata K, Kirikae T, et al. Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS). *BMC Genomics.* 2011;12:17.
  105. Humar D, Datta V, Bast DJ, Beall B, De Azavedo JCS, Nizet V. Streptolysin S and necrotising infections produced by group G streptococcus. *Lancet.* 2002;359(9301):124–9.
  106. Siemens N, Kittang BR, Chakrakodi B, Oppegaard O, Johansson L, Bruun T, et al. Increased cytotoxicity and streptolysin O activity in group G streptococcal strains causing invasive tissue infections. *Sci Rep.* 2015;5:16945.
  107. Kimoto H, Fujii Y, Yokota Y, Taketo A. Molecular characterization of NADase-streptolysin O operon of hemolytic streptococci. *Biochim Biophys Acta.* 2004/11/30. 2005;1681(2–3):134–49.
  108. Bricker AL, Carey VJ, Wessels MR. Role of NADase in virulence in experimental invasive group A streptococcal infection. *Infect Immun.* 2005;73(10):6562–6.
  109. Steiner K, Malke H. Dual control of streptokinase and streptolysin S production by the *covRS* and *fasCAX* two-component regulators in *Streptococcus dysgalactiae* subsp. *equisimilis*. *Infect Immun.* 2002;70(7):3627–36.

110. Hasegawa T, Matsumoto M, Hata N, Yano H, Isaka M, Tatsuno I. Homologous role of CovRS two-component regulatory system in NAD<sup>+</sup>-glycohydrolase activity in *Streptococcus dysgalactiae* subsp. *equisimilis* as in *Streptococcus pyogenes*. *APMIS*. 2019;127(2):87–92.
111. Oppegaard O, Mylvaganam H, Skrede S, Lindemann PC, Kittang BR. Emergence of a *Streptococcus dysgalactiae* subspecies *equisimilis* stG62647-lineage associated with severe clinical manifestations. *Sci Rep*. 2017;7(1):7589.
112. Michael-Gayego A, Dan-Goor M, Jaffe J, Hidalgo-Grass C, Moses AE. Characterization of sil in invasive group A and G streptococci: antibodies against bacterial pheromone peptide SilCR result in severe infection. *Infect Immun*. 2013;81(11):4121–7.
113. Davies MR, McMillan DJ, Beiko RG, Barroso V, Geffers R, Sriprakash KS, et al. Virulence Profiling of *Streptococcus dysgalactiae* Subspecies *equisimilis* Isolated from Infected Humans Reveals 2 Distinct Genetic Lineages That Do Not Segregate with Their Phenotypes or Propensity to Cause Diseases. *Clin Infect Dis*. 2007;44(11):1442–54.
114. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. *Clin Microbiol Rev*. 1989;2(3):285–314.
115. Cunningham MW. Pathogenesis of group A streptococcal infections. *Clin Microbiol Rev*. 2000;13(3):470–511.
116. Bisno AL, Craven DE, McCabe WR. M proteins of group G streptococci isolated from bacteremic human infections. *Infect Immun*. 1987;55(3):753–7.
117. Jones KF, Fischetti VA. Biological and immunochemical identity of M protein on group G streptococci with M protein on group A streptococci. *Infect Immun*. 1987;55(3):502–6.
118. Schnitzler N, Podbielski A, Baumgarten G, Mignon M, Kaufhold A. M or M-like protein gene polymorphisms in human group G streptococci. *J Clin Microbiol*. 1995;33(2):356–63.
119. Kittang BR, Skrede S, Langeland N, Haanshuus CG, Mylvaganam H. emm gene diversity, superantigen gene profiles and presence of SlaA among clinical isolates

- of group A, C and G streptococci from western Norway. *Eur J Clin Microbiol Infect Dis*. 2011;30(3):423–33.
120. McMillan DJ, Kaul SY, Bramhachari P V, Smeesters PR, Vu T, Karmarkar MG, et al. Recombination drives genetic diversification of *Streptococcus dysgalactiae* subspecies *equisimilis* in a region of streptococcal endemicity. *PLoS One*. 2011;6(8):e21346.
  121. Smeesters PR, McMillan DJ, Sriprakash KS. The streptococcal M protein: a highly versatile molecule. *Trends Microbiol*. 2010;18(6):275–82.
  122. Lo H-H, Cheng W-S. Distribution of virulence factors and association with emm polymorphism or isolation site among beta-hemolytic group G *Streptococcus dysgalactiae* subspecies *equisimilis*. *APMIS*. 2014;123(1):45–52.
  123. McMillan DJ, Dreze P-A, Vu T, Bessen DE, Guglielmini J, Steer AC, et al. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. *Clin Microbiol Infect*. 2013;19(5):E222-9.
  124. Collins CM, Kimura A, Bisno AL. Group G streptococcal M protein exhibits structural features analogous to those of class I M protein of group A streptococci. *Infect Immun*. 1992;60(9):3689–96.
  125. Kittang BR, Langeland N, Mylvaganam H. Distribution of emm types and subtypes among noninvasive group A, C and G streptococcal isolates in western Norway. *APMIS*. 2008;116(6):457–64.
  126. Brandt CM, Schweizer KG, Holland R, Lütticken R, Freyaldenhoven BS. Lack of mitogenic activity of speG- and speG<sub>dys</sub>-positive *Streptococcus dysgalactiae* subspecies *equisimilis* isolates from patients with invasive infections. *Int J Med Microbiol*. 2005;295(8):539–46.
  127. Zhao J, Hayashi T, Saarinen S, Papageorgiou AC, Kato H, Imanishi K, et al. Cloning, Expression, and Characterization of the Superantigen Streptococcal Pyrogenic Exotoxin G from *Streptococcus dysgalactiae*. *Infect Immun*. 2007;75(4):1721–9.
  128. Bisno AL, Brito MO, Collins CM. Molecular basis of group A streptococcal virulence. *Lancet Infect Dis*. 2003;3(4):191–200.

129. Korem M, Hidalgo-Grass C, Michael-Gayego A, Nir-Paz R, Salameh S, Moses AE. Streptococcal pyrogenic exotoxin G gene in blood and pharyngeal isolates of *Streptococcus dysgalactiae* subspecies *equisimilis* has a limited role in pathogenesis. *J Microbiol Immunol Infect.* 2014;47(4):292–6.
130. Sachse S, Seidel P, Gerlach D, Günther E, Rödel J, Straube E, et al. Superantigen-like gene(s) in human pathogenic *Streptococcus dysgalactiae*, subsp. *equisimilis*: genomic localisation of the gene encoding streptococcal pyrogenic exotoxin G (*speGdys*). *Pathog Dis.* 2002;34(2):159–67.
131. Okumura K, Shimomura Y, Murayama SY, Yagi J, Ubukata K, Kirikae T, et al. Evolutionary paths of streptococcal and staphylococcal superantigens. *BMC Genomics.* 2012;13:404.
132. van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect.* 2007;13 Suppl 3:1–46.
133. Efstratiou A, Colman G, Hahn G, Timoney JF, Boeufgras JM, Monget D. Biochemical differences among human and animal streptococci of Lancefield group C or group G. *J Med Microbiol.* 1994;41(2):145–8.
134. Efstratiou A. The serotyping of hospital strains of streptococci belonging to Lancefield group C and group G. *J Hyg (Lond).* 1983;90(1):71–80.
135. Bert F, Branger C, Poutrel B, Lambert-Zechovsky N. Differentiation of human and animal strains of *Streptococcus dysgalactiae* by pulsed-field gel electrophoresis. *FEMS Microbiol Lett.* 1997;150(1):107–12.
136. Ahmad Y, Gertz REJ, Li Z, Sakota V, Broyles LN, Van Beneden C, et al. Genetic relationships deduced from *emm* and multilocus sequence typing of invasive *Streptococcus dysgalactiae* subsp. *equisimilis* and *S. canis* recovered from isolates collected in the United States. *J Clin Microbiol.* 2009;47(7):2046–54.
137. McMillan DJ, Bessen DE, Pinho M, Ford C, Hall GS, Melo-Cristino J, et al. Population genetics of *Streptococcus dysgalactiae* subspecies *equisimilis* reveals widely dispersed clones and extensive recombination. *PLoS One.* 2010;5(7):e11741.

- 
138. Olive DM, Bean P. Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol.* 1999;37(6):1661–9.
  139. Bert F, Branger C, Lambert-Zechovsky N. Pulsed-field gel electrophoresis is more discriminating than multilocus enzyme electrophoresis and random amplified polymorphic DNA analysis for typing pyogenic streptococci. *Curr Microbiol.* 1997;34(4):226–9.
  140. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, et al. Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. *PLOS Pathog.* 2012;8(8):e1002824.
  141. Beall B, Facklam R, Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. *J Clin Microbiol.* 1996;34(4):953–8.
  142. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *PANS.* 1998;95(6):3140–5.
  143. Francisco AP, Bugalho M, Ramirez M, Carriço JA. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics.* 2009;10(1):152.
  144. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol.* 2004;186(5):1518–30.
  145. Feil EJ, Enright MC. Analyses of clonality and the evolution of bacterial pathogens. *Curr Opin Microbiol.* 2004;7(3):308–13.
  146. Tagini F, Aubert B, Troillet N, Pillonel T, Praz G, Crisinel PA, et al. Importance of whole genome sequencing for the assessment of outbreaks in diagnostic laboratories: analysis of a case series of invasive *Streptococcus pyogenes* infections. *Eur J Clin Microbiol Infect Dis.* 2017;36(7):1173–80.
  147. Rossen JWA, Friedrich AW, Moran-Gilad J. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin Microbiol Infect.* 2018;24(4):355–60.

148. Goldberg B, Sichtig H, Geyer C, Ledebner N, Weinstock GM. Making the Leap from Research Laboratory to Clinic: Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease Diagnostics. *MBio*. 2015;6(6):e01888-15.
149. Gwinn M, MacCannell D, Armstrong GL. Next-Generation Sequencing of Infectious Pathogens. *JAMA*. 2019;321(9):893–4.
150. Athey TBT, Teatero S, Li A, Marchand-Austin A, Beall BW, Fittipaldi N. Deriving group A Streptococcus typing information from short-read whole-genome sequencing data. *J Clin Microbiol*. 2014;52(6):1871–6.
151. Dunne WMJ, Westblade LF, Ford B. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis*. 2012;31(8):1719–26.
152. Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol*. 2013/09/02. 2013;11(10):728–36.
153. Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. *Nat Rev Microbiol*. 2015/11/09. 2015;13(12):787–94.
154. Hudzicki J. Kirby-Bauer Disk Diffusion Susceptibility Test Protocol. 2009.
155. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
156. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77.
157. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9(11):e112963.
158. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9.
159. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating

- signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785–6.
160. Lagesen K, Hallin P, Rodland EA, Staerfeldt H-H, Rognes T, Ussery DW. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35(9):3100–8.
161. Nascimento M, Sousa A, Ramirez M, Francisco AP, Carriço JA, Vaz C. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics*. 2016/09/06. 2017;33(1):128–9.
162. Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, Santos S, et al. chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. *Microb genomics*. 2018;4(3).
163. Ribeiro-Goncalves B, Francisco AP, Vaz C, Ramirez M, Carrico JA. PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res*. 2016;44(W1):W246-51.
164. Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carriço JA. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics*. 2012;13:87.
165. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res*. 2019;47(D1):D687–92.
166. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother*. 2014;58(1):212–20.
167. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012;67(11):2640–4.
168. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2020;48(D1):D517–25.
169. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol*.

- 1988;26(11):2465–6.
170. Severiano A, Pinto FR, Ramirez M, Carrico JA. Adjusted Wallace coefficient as a measure of congruence between typing methods. *J Clin Microbiol.* 2011;49(11):3997–4000.
171. Laupland KB, Pasquill K, Parfitt EC, Steele L. Bloodstream infection due to  $\beta$ -hemolytic streptococci: a population-based comparative analysis. *Infection.* 2019;
172. Katsukawa C, Tamaru A, Morikawa Y. *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *Kansenshogaku Zasshi.* 2002;76(3):155–60.
173. Tanaka D, Isobe J, Watahiki M, Nagai Y, Katsukawa C, Kawahara R, et al. Genetic features of clinical isolates of *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *J Clin Microbiol.* 2008;46(4):1526–9.
174. Brandt CM, Haase G, Schnitzler N, Zbinden R, Lutticken R. Characterization of blood culture isolates of *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *J Clin Microbiol.* 1999;37(12):4194–7.
175. Tseng S-P, Lin Y-Y, Tsai J-C, Hsueh P-R, Chen H-J, Hung W-C, et al. Distribution of emm types and genetic characterization of the *mgc* locus in group G *Streptococcus dysgalactiae* subsp. *equisimilis* from a hospital in northern Taiwan. *J Clin Microbiol.* 2010;48(8):2975–7.
176. Pinho MD. Clonal diversity and antibiotic susceptibility of invasive *Streptococcus dysgalactiae* subsp. *equisimilis* isolated in Portugal: O553. In: *Clinical Microbiology & Infection* 18. London, UK; 2012.
177. Trell K, Nilson B, Rasmussen M. Species and emm-type distribution of group C and G streptococci from different sites of isolation. *Diagn Microbiol Infect Dis.* 2016;86(4):467–9.
178. Tsai C-T, Chi C-Y, Ho C-M, Lin P-C, Chou C-H, Wang J-H, et al. Correlation of virulence genes to clinical manifestations and outcome in patients with *Streptococcus dysgalactiae* subspecies *equisimilis* bacteremia. *J Microbiol Immunol Infect.* 2014;47(6):462–8.
179. Dhanda V, Chaudhary P, Toor D, Kumar R, Chakraborti A. Antimicrobial

- susceptibility pattern of beta-haemolytic group A, C and G streptococci isolated from North India. *J Med Microbiol.* 2013;62(Pt 3):386–93.
180. European Centre for Disease Prevention and Control. Antimicrobial consumption database (ESAC-Net) [Internet]. 2018.
181. Merino Diaz L, Torres Sanchez MJ, Aznar Martin J. Prevalence and mechanisms of erythromycin and clindamycin resistance in clinical isolates of beta-haemolytic streptococci of Lancefield groups A, B, C and G in Seville, Spain. *Clin Microbiol Infect.* 2008;14(1):85–7.
182. Sunaoshi K, Murayama SY, Adachi K, Yagoshi M, Okuzumi K, Chiba N, et al. Molecular emm genotyping and antibiotic susceptibility of *Streptococcus dysgalactiae* subsp. *equisimilis* isolated from invasive and non-invasive infections. *J Med Microbiol.* 2010;59(Pt 1):82–8.
183. Coutinho V de LS, Paiva RM, Reiter KC, de-Paris F, Barth AL, Machado ABMP. Distribution of *erm* genes and low prevalence of inducible resistance to clindamycin among staphylococci isolates. *Braz J Infect Dis.* 2010;14(6):564–8.
184. Martineau F, Picard FJ, Lansac N, Menard C, Roy PH, Ouellette M, et al. Correlation between the resistance genotype determined by multiplex PCR assays and the antibiotic susceptibility patterns of *Staphylococcus aureus* and *Staphylococcus epidermidis*. *Antimicrob Agents Chemother.* 2000;44(2):231–8.
185. Ruegg PL, Oliveira L, Jin W, Okwumabua O. Phenotypic antimicrobial susceptibility and occurrence of selected resistance genes in gram-positive mastitis pathogens isolated from Wisconsin dairy cows. *J Dairy Sci.* 2015;98(7):4521–34.
186. Haenni M, Saras E, Bertin S, Leblond P, Madec J-Y, Payot S. Diversity and mobility of integrative and conjugative elements in bovine isolates of *Streptococcus agalactiae*, *S. dysgalactiae* subsp. *dysgalactiae*, and *S. uberis*. *Appl Environ Microbiol.* 2010;76(24):7957–65.
187. Brenciani A, Bacciaglia A, Vecchi M, Vitali LA, Varaldo PE, Giovanetti E. Genetic elements carrying *erm(B)* in *Streptococcus pyogenes* and association with *tet(M)* tetracycline resistance gene. *Antimicrob Agents Chemother.* 2007;51(4):1209–16.

188. Metcalf BJ, Chochua S, Gertz REJ, Hawkins PA, Ricaldi J, Li Z, et al. Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive *Streptococcus agalactiae* recovered in the USA. *Clin Microbiol Infect.* 2017;23(8):574.e7-574.e14.
189. Malhotra-Kumar S, Lammens C, Martel A, Mallentjer C, Chapelle S, Verhoeven J, et al. Oropharyngeal carriage of macrolide-resistant viridans group streptococci: a prevalence study among healthy adults in Belgium. *J Antimicrob Chemother.* 2004;53(2):271–6.
190. Ferech M, Coenen S, Malhotra-Kumar S, Dvorakova K, Hendrickx E, Suetens C, et al. European Surveillance of Antimicrobial Consumption (ESAC): outpatient quinolone use in Europe. *J Antimicrob Chemother.* 2006;58(2):423–7.
191. Adriaenssens N, Coenen S, Versporten A, Muller A, Minalu G, Faes C, et al. European Surveillance of Antimicrobial Consumption (ESAC): outpatient quinolone use in Europe (1997–2009). *J Antimicrob Chemother.* 2011;66(suppl\_6):vi47–56.
192. European Centre for Disease Prevention and Control. Antimicrobial consumption - Annual epidemiological report 2017. Stockholm: ECDC; 2018.

**Tabela A1:** Lista de laboratórios hospitalares.

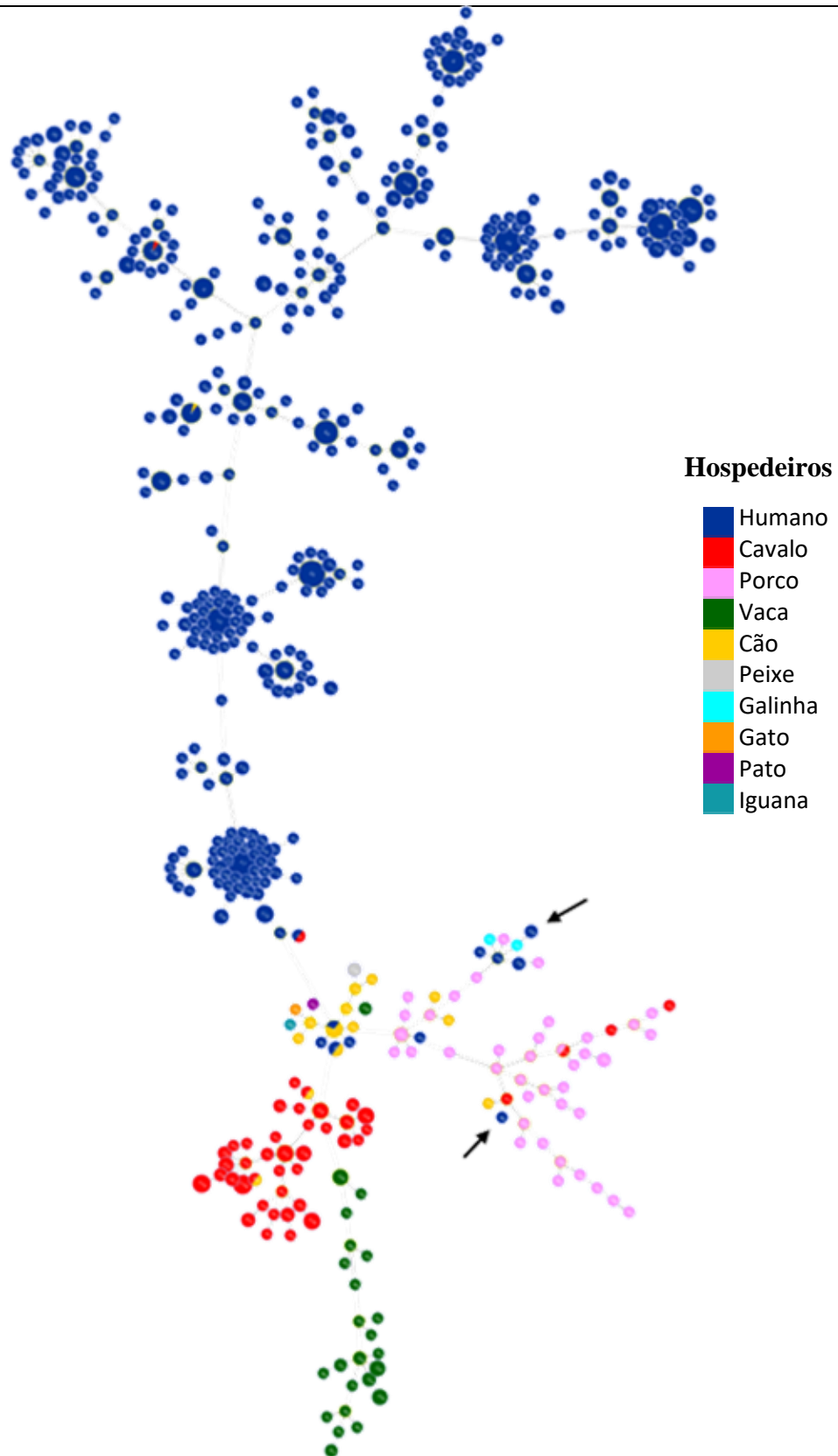
Centro Hospitalar de Leiria - Hospital de Santo André
Centro Hospitalar de Lisboa Ocidental
Centro Hospitalar de Vila Nova de Gaia/Espinho
Centro Hospitalar do Baixo Vouga - Hospital Infante D. Pedro
Centro Hospitalar do Funchal - Hospital Dr Nélio Mendonça
Centro Hospitalar do Médio Ave
Centro Hospitalar e Universitário de Coimbra
Centro Hospitalar Entre Douro e Vouga - Hospital São Sebastião
Centro Hospitalar Póvoa de Varzim/Vila do Conde
Centro Hospitalar Tondela Viseu
Centro Hospitalar Universitário de Lisboa Central
Centro Hospitalar Universitário de Lisboa Norte
Centro Hospitalar Universitário de S. João
Centro Hospitalar Universitário do Algarve - Hospital de Portimão
Centro Hospitalar Universitário do Porto - Hospital de Santo António
Hospital Beatriz Ângelo - Loures
Hospital da Senhora da Oliveira Guimarães
Hospital de Braga
Hospital de Cascais Dr. José de Almeida
Hospital Espírito Santo - Évora
Hospital Garcia de Orta - Almada
SAMS Lisboa
Unidade Local de Saúde de Matosinhos - Hospital Pedro Hispano
Unidade Local de Saúde do Alto Minho - Hospital Santa Luzia de Viana do Castelo

---

**Tabela A2:** Diferenças entre CCs.

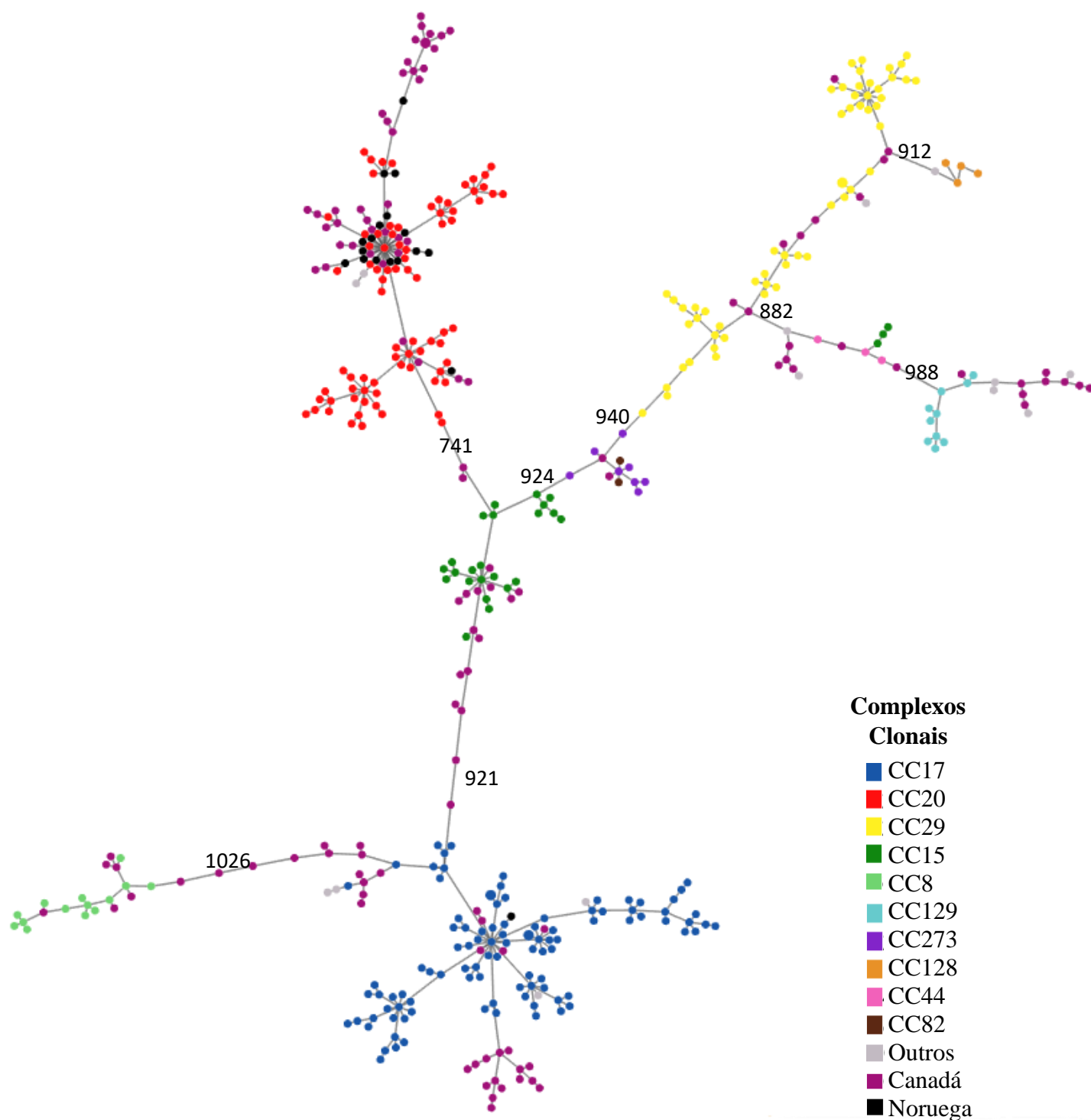
	CC17	CC20	CC29	CC15	CC8	CC129	CC273	CC128	CC44	CC82	CC3	CC34	CC63	CC74	CC138	CC248	S*
CC17			948		1093											1447	149
CC20				748													33
CC29	948														253		891
CC15		748					924		980								914
CC8	1093																
CC129														943			1074
CC273			940	924						854							
CC128										796							259
CC44				980													812
CC82							854			796							
CC3																	1164
CC34														1102			
CC63																	1057
CC74						943						1102					455
CC138			253														
CC248	1447																1415
S*	149	668	891			1074		259	812		1164		1057	455		1415	
	33		914														

\*S inclui oito estirpes pertencentes a STs singletons.



**Figura A1:** Árvore de extensão mínima de 1775 estirpes de *S. dysgalactiae* isoladas de diferentes hospedeiros com base no esquema MLST.

Inclui todas as estirpes de *S. dysgalactiae* presentes na base de dados (<https://pubmlst.org/sdysgalactiae/>) a 1 de dezembro de 2019. As estirpes ST249 e ST483 encontradas em infecção invasiva em Portugal estão indicadas com setas. 1554 estirpes isoladas de humanos, 109 de cavalos, 43 de porcos, 42 de vacas, 19 de cães, três de peixes, duas de galinhas, uma de gato, uma de pato e uma de iguana.



**Figura A2:** Árvore de extensão mínima de 456 estirpes invasivas isoladas de diferentes países com base no esquema cgMLST.

Representa a distribuição dos CCs definidos por MLST.

316 estirpes invasivas isoladas em Portugal, 122 estirpes de infecção invasiva isoladas no Canadá e 18 estirpes *stG62647* isoladas na Noruega.

Outros inclui seis CCs (CCs 3, 34, 63, 75, 138 e 248) com uma única estirpe identificada no conjunto das 316 estirpes analisadas e oito estirpes pertencentes a STs singletons.