UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA

NATURALLY OCCURRING MUTATIONS IN REGULATORY PROTEINS AMONG *STREPTOCOCCUS PYOGENES* ISOLATES FROM DISTINCT HUMAN INFECTIONS

CATARINA TERESA CONDINHO PATO

Orientador: Professor Doutor Mário Nuno Ramos de Almeida Ramirez

Tese especialmente elaborada para a obtenção do grau de Doutor em Ciências e Tecnologias da Saúde, especialidade em Microbiologia

2018

UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA

NATURALLY OCCURRING MUTATIONS IN REGULATORY PROTEINS AMONG *STREPTOCOCCUS PYOGENES* ISOLATES FROM DISTINCT HUMAN INFECTIONS

CATARINA TERESA CONDINHO PATO

Orientador: Professor Doutor Mário Nuno Ramos de Almeida Ramirez

Tese especialmente elaborada para a obtenção do grau de Doutor em Ciências e Tecnologias da Saúde, especialidade em Microbiologia

Júri:

Presidente:

Doutor José Augusto Gamito Melo Cristino, Professor Catedrático e Presidente do Conselho Científico da Faculdade de Medicina da Universidade de Lisboa

Vogais:

Doctor Shiranee Sriskandan, *Professor of Faculty of Medicine, Imperial College London*;

Doutor Paulo Jorge Pereira Cruz Paixão, Professor Auxiliar da Faculdade de Ciências Médicas - Universidade NOVA de Lisboa;

Doutora Isabel Maria dos Santos Leitão Couto, Professora Auxiliar do Instituto de Higiene e Medicina Tropical - Universidade NOVA de Lisboa;

Doutora Maria Constança Matias Ferreira Pomba, Professora Associada da Faculdade de Veterinária da Universidade de Lisboa;

Doutor Mário Nuno Ramos de Almeida Ramirez, Professor Associado com Agregação da Faculdade de Medicina da Universidade de Lisboa (*Orientador*);

Doutora Ana Isabel de Aquino Friães, Professora Auxiliar da Faculdade de Medicina da Universidade de Lisboa.

Fundação para a Ciência e Tecnologia (SFRH/BD/81329/2011)

2018

AKNOWLEGMENTS

In first place I would like to express my sincere gratitude to my PhD supervisor, Professor Mário Ramirez. First, for having accepted me as PhD student but mostly for his will and availability to share his immense knowledge and offering suggestions during the innumerous work discussions and revisions, that were definitely essential for the development of my critical way of thinking and were the basis of all my work decisions.

My sincere thanks to Professor José Melo Cristino, whom I deeply respect and admire, for having provided me with the opportunity to join his team and whose professional life is an inspiration.

I am also deeply grateful to Ana Friães, with whom I worked more closely, for having introduced me to the majority of the methods used in this thesis, for all the work revisions and of course for all the guidance and support.

A special thank to my PhD partners Andreia Horácio, Jorge Miranda, (MRamirez Lab), and Sofia Cerqueira (PSimas Lab), who shared this journey with me. I want to thank them for their friendship, support and for all the incentive.

I also wish to thank all my other colleagues at MRamirez Lab, for all the technical support and mostly for the friendly environment and team spirit. These include Sandra Aguiar, Maria João Frias, Elizabete Martins, João Carriço, Margarida Carrolo, Catarina Costa, Marcos Pinho, Filomena Costa, Alice Faria, Joana Lopes, Raquel Garcia, Mickael Silva and Miguel Machado.

For the financial support, I want to thank Fundação para a Ciência e Tecnologia for having granted me with a Ph.D. scholarship (SFRH/BD/81329/2011).

My final thanks are to my friends and family, especially to my parents, for always believing in me, for keeping me motivated and mostly for all their love. I could not have done it without them.

SUMMARY

Keywords: *Streptococcus pyogenes*, CovRS, SpeB, skin and soft tissue infections

Streptococcus pyogenes (Group A *Streptococcus*, GAS) is among the most prevalent bacterial pathogens of humans and it is responsible for a wide range of infections, from pharyngitis and impetigo to life-threatening conditions such as necrotizing fasciitis and streptococcal toxic shock syndrome. However, until today, despite many years of research, there is still no consensus regarding which are the genotypic or phenotypic characteristics that confer to an isolate a certain tissue preference or a more invasive potential. The nasopharyngeal mucosa and skin are considered the primary sources of isolates responsible for invasive infections. This suggests a further ability of these isolates to invade and survive in deeper tissues. In the last years, spontaneous mutations occurring in the *covRS* two-component regulatory system have been considered a possible explanation for the transition from localized to systemic infection. It is estimated that this system controls directly or indirectly the expression of 10-15% of the GAS genome. As a consequence of these mutations, among the distinct patterns of expression of several virulence factors, the downregulation of the extracellular cysteine protease SpeB, has been considered crucial for the switch to a hypervirulent phenotype. The downregulation of SpeB has also been described as a consequence of mutations in a stand-alone transcriptional regulator named RopB. However, contrasting results were reported whether *covRS* mutations were more prone to occur in certain lineages, namely those more frequently associated with invasive infections or if they occur in a similar proportion among invasive and non-invasive isolates. To address these questions, in the present thesis, we determined the sequence of the *covRS* and *ropB* of 191 isolates from invasive infection and pharyngitis and evaluated the production of SpeB, as well of NAD glycohydrolase (NADase) and streptolysin S (SLS), which are two virulence factors supposed to be under the influence of CovRS. Moreover, skin and soft tissue infections (SSTI) are frequently considered the focal points for the development of invasive disease. However, most of the knowledge about GAS skin isolates is from studies that intended to find differences between isolates recovered from invasive and non-invasive infections generically. The majority of these studies use isolates recovered from SSTI but mostly from pharyngitis and compared those together against the isolates recovered from

invasive infections, resulting in few data regarding isolates recovered from SSTI. Therefore, we characterized by multiple typing methods a total of 320 isolates from SSTI recovered in Portugal and performed the comparison with invasive isolates recovered during the same period which were previously characterized. All SSTI isolates were also tested for SpeB activity and for those without detectable SpeB activity we determined the sequence of *covRS* and *ropB* genes.

Overall, we found that isolates with null *covS* alleles, which are predicted to eliminate the protein function have a significant association with invasive infections comparative with isolates from pharyngitis and SSTI. Additionally, none of these isolates, as expected, had SpeB activity, and, with few exceptions, they showed an increased activity of both NADase and SLS that could explain their potentially higher invasiveness. Even so, this mechanism was found to be uncommon, corresponding to only 10% of invasive isolates, which could be due to an overall fitness cost of these mutations. Moreover, null *covS* alleles were not more prevalent among isolates from clones frequently associated with invasive infection such as *emm*1 and *emm*64 and instead they were distributed throughout diverse genetic backgrounds. The few exceptions regarded the levels of NADase and SLS points to the complexity of the regulatory networks among distinct GAS lineages. Additionally, no null *covR* alleles were detected in our isolates and *ropB* null alleles were found in a low fraction of GAS isolates and were not associated with any infection type. Regarding SpeB activity, it was detected in a similar proportion in isolates recovered from the different sources, and therefore its absence was not associated to any type of infection, suggesting that its abrogation cannot by itself explain the higher ability of certain clones to cause invasive disease. Among SSTI isolates, we found that *emm*89 type isolates were the most prevalent and were significantly associated with these infections when compared with invasive isolates. In contrast, *emm*1, *emm*3, and *emm*64 isolates were associated with invasive infections. Within *emm89* isolates, SSTI were only associated to those that lack the *hasABC* locus, corresponding to a recently emerged acapsular clade (clade 3) that also carries a variant of the *ngs-ifs-slo* locus. These results suggest that for some unknown reason these isolates may have an increased potential to cause SSTI. As a consequence of known differences in the *emm*-type between isolates causing these two types of infections, we also found significant associations between the ability to bind to different host proteins. This ability was presumed by inferring the *emm*-cluster through the *emm*-type results. The *emm*-cluster is a recent classification based on the entire

sequence of the *emm* gene (M protein) where each cluster shares binding motifs to host proteins and other structural properties. Therefore, the ability to bind fibrinogen and albumin were significantly associated with invasive isolates, whereas the ability to bind to C4BP and IgG were associated with SSTI isolates. Differences in the presence of superantigen (SAg) genes, SAg profiles and in the distribution of sequence types (ST) determined by Multilocus Sequence Typing (MLST) were also noted. The possible impact of these differences in the ability of the isolates to cause these distinct infections remains to be clarified. Moreover, within each *emm* type the same MLST defined lineages and SAg profiles could be found in both types of infection, questioning the possibility that these characteristics dictate the tissue tropism of each isolate.

In summary, the results described in this thesis indicate that isolates responsible for SSTI are genetically distinct from those recovered from normally sterile sites and while some GAS clones have more capacity to invade deeper tissues, others are more prone to cause SSTI. Moreover, the significant presence of null *covS* mutations among invasive isolates and the fact that no association was observed regarding the absence of SpeB activity in isolates from different types of infections, suggests that the role of spontaneous mutations impairing the CovRS activity is probably related with the regulation of others virulence factors under its control in addition to SpeB.

RESUMO

Palavras-Chave: *Streptococcus pyogenes*, CovRS, SpeB, infeções da pele e tecidos moles

As bactérias da espécie *Streptococcus pyogenes* estão entre os principais agentes bacterianos responsáveis por infeções no Homem. São também frequentemente denominadas por estreptococos pertencendo ao Grupo A de Lancefield, por aglutinarem com o soro A da classificação de Lancefield, uma investigadora cujos trabalhos foram cruciais para o atual conhecimento sobre esta espécie. Uma das características mais particulares desde agente é a capacidade de causar um grande espectro de infeções variando relativamente ao local e à gravidade. Pode causar desde doenças autolimitadas da faringe e pele, como faringo-amigdalites e lesões de impetigo, a infeções invasivas frequentemente associadas a elevada morbilidade e mortalidade, como são o caso da fasceíte necrosante e a síndrome do choque tóxico. Também de referir que o seu principal impacto mundial advém das complicações autoimunes, hoje em dia ainda com elevada prevalência nos países em desenvolvimento, onde estão incluídas a febre reumática e a glomerulonefrite. Assim sendo, para causar esta variedade de infeções, este microrganismo tem que se adaptar a diferentes ambientes. A orofaringe e a pele são consideradas como os principais focos de estirpes[1] invasivas, o que implica que estas estirpes têm que invadir e sobreviver nos tecidos celulares subcutâneos. Porém, mesmo depois de décadas de investigação, ainda não há consenso relativamente a quais são as características moleculares ou fenotípicas que conferem a uma estirpe um maior potencial invasivo ou uma determinada preferência por um local anatómico.

Em 2006, um estudo com um modelo murganho de infeção da pele e tecidos moles sugeriu a ocorrência de mutações espontâneas num sistema regulador durante o processo de infeção como uma explicação para a transição da infeção localizada para sistémica. Vários estudos foram publicados de seguida a apoiar esta hipótese, incluindo alguns que descreveram a presença destas mutações em estirpes recolhidas de infeções no Homem. O sistema regulador em causa é constituído por dois componentes e denomina-se de CovRS. Estima-se que este sistema controle direta ou indiretamente cerca de 10 a 15% de todo o genoma de *S. pyogenes*. Entre as alterações verificadas na

[1] Neste resumo, a palavra estirpe é utilizada para referir um microrganismo isolado de um determinado produto biológico

expressão de vários fatores de virulência, a abolição da atividade de uma protease extracelular denominada SpeB tem sido considerada como um evento crucial para uma maior virulência. Teoricamente, na ausência de SpeB, são preservados vários fatores de virulência presentes na superfície bacteriana que contribuem para o processo invasivo. Adicionalmente, a diminuição da atividade de SpeB tem também sido descrita como resultado de mutações num outro regulador, designado RopB. Porém, enquanto alguns estudos defendem que as mutações nestes reguladores ocorrem mais frequentemente em estirpes recolhidas de infeções invasivas, outros defendem que estas ocorrem na mesma proporção em estipes de infeções não invasivas. Alguns defendem ainda que estas mutações estão limitadas às linhagens frequentemente associadas a infeções invasivas. De modo a abordar estas questões, determinámos as sequências dos genes *covRS* e *ropB* de 191 estirpes recolhidas de infeções invasivas e de amigdalites. Determinámos também, a atividade de SpeB de cada uma destas estirpes e os respetivos valores de NAD glicohidrolase e de estreptolisina S que são dois fatores de virulência aparentemente também sob o controlo do regulador CovRS. Por sua vez, as infeções da pele e tecidos moles são frequentemente consideradas como um foco para o desenvolvimento de infeções invasivas. Porém, a informação disponível sobre estirpes da pele é escassa porque a maioria do conhecimento atual advém de estudos cujo objetivo é a identificação de diferenças entre estirpes recolhidas de infeções invasivas e não invasivas. Os estudos em questão tendem a analisar as características das estirpes isoladas da pele em conjunto com estirpes recolhidas de exsudados faríngeos, o que resulta numa difícil avaliação das características específicas desta população. Desta forma, caracterizámos 320 estirpes recolhidas de infeções da pele e tecidos moles em Portugal através de vários métodos de tipagem e procedemos à sua comparação com estirpes invasivas recolhidas durante o mesmo período previamente trabalhadas. Avaliámos ainda, para as 320 estipes, a presença de atividade de SpeB e naquelas que não apresentavam atividade, determinámos as respetivas sequências dos genes *covRS* e do *ropB*.

Verificámos existir uma associação significativa entre estirpes cujas mutações se preveem eliminar a função proteica do CovS e as infeções invasivas, comparativamente com estirpes recolhidas de faringo-amigdalites e de infeções da pele e tecidos moles. Adicionalmente, assim como esperado, nenhuma destas estirpes apresentava atividade SpeB e, salvo algumas exceções, apresentavam um aumento dos níveis de atividade de NAD glicohidrolase e de estreptolisina S. Embora estas diferenças sejam concordantes

com uma maior capacidade invasiva, este mecanismo apenas foi detetado em 10% das estirpes invasivas analisadas. Pensa-se que estas mutações possam estar associadas a um elevado "fitness cost", dificultando por exemplo a capacidade de colonização e de transmissão, o que pode explicar a sua reduzida prevalência. Verificou-se ainda que estas mutações se encontram dispersas por estirpes de diversas linhagens e não estão apenas restritas às mais frequentemente associadas a infeções invasivas, como estirpes do tipo *emm*1 e *emm*64. As exceções detetadas nos valores de atividade de NAD glicohidrolase e estreptolisina S realçam a complexidade já conhecida dos mecanismos de regulação de *S. pyogenes* e as potenciais diferenças entre estirpes de linhagens distintas. Relativamente ao *covR,* não foi detetada nenhuma mutação que se preveja eliminar a função proteica. Por sua vez, para o gene *ropB,* este tipo de mutações foram detetadas num reduzido número de estirpes e não foram associadas a nenhum tipo de infeção. No que se refere à atividade de SpeB, esta foi detetada na maioria das estirpes e numa proporção semelhante em estirpes de diferentes origens, pelo que a sua ausência não foi associada a nenhum tipo de infeção. Estes resultados sugerem que a ausência de SpeB por si só não explica a maior capacidade de certas estirpes para causar infeção invasiva.

Estirpes do tipo *emm*89 eram as mais prevalentes e estavam significativamente associadas às infeções da pele e tecidos moles, quando comparadas com as estirpes invasivas. Por outro lado, estirpes do tipo *emm*1, *emm*3 e *emm*64 foram significativamente associadas a infeções invasivas. Porém, de entre as estirpes do tipo *emm*89, as infeções da pele e tecidos moles apenas foram associadas às estirpes que não têm o locus *hasABC*, responsável pela síntese da cápsula de ácido hialurónico. Estas estipes correspondem a uma linhagem recente sem cápsula que possui uma variante do locus *ngs-ifs-slo*. Estes resultados sugerem que, por alguma razão ainda desconhecida, estas estirpes podem apresentar uma maior apetência para causar infeções da pele e tecidos moles. Refletindo os tipos de *emm* associados a cada tipo de infeção, outras associações significativas foram detetadas, nomeadamente a presumida capacidade de ligação a diferentes proteínas do Homem durante o processo de infeção. Esta capacidade foi presumida deduzindo os "clusters" de *emm* através dos tipos de *emm.* Os "clusters" correspondem a uma classificação mais recente que se baseia na totalidade da sequência do gene *emm*, ao contrário dos tipos de *emm* que se baseiam apenas numa porção deste. Segundos os autores, as estirpes de cada "cluster" partilham zonas de ligação a proteínas do hospedeiro. Assim sendo, enquanto a capacidade de ligação ao

fibrinogénio e à albumina foram significativamente associadas às estirpes invasivas, a capacidade de ligação ao C4BP e IgG foram associadas às estirpes da pele e tecidos moles. Também foram notadas diferenças na presença de genes codificantes de superantigénios, nos perfis de superantigénios e nos "sequence types" determinados através de "Multilocus Sequence Typing". Embora se tenham verificado diversas diferenças entre estirpes com diferentes origens, permanece por esclarecer se estas influenciam a apresentação da doença e qual o mecanismo subjacente para que isto ocorra. Por sua vez, ao considerar estirpes do mesmo tipo de *emm*, verificaram-se os mesmos perfis de superantigénios e os mesmos "sequence types" em estirpes de ambos os tipos de infeção, o que sugere que é pouco provável que estes indicadores estejam relacionados com a preferência de cada estipe para um determinado local anatómico ou a um maior potencial invasivo.

Resumidamente, os resultados descritos nesta tese indicam que as estirpes responsáveis por infeções da pele e tecidos moles são uma população geneticamente distinta das estirpes recolhidas de locais normalmente estéreis e enquanto algumas linhagens têm uma maior capacidade invasiva outras são mais propensas a causar infeções da pele e tecidos moles. Por fim, o facto de se ter verificado uma presença significativa de mutações no *covS* que se preveem que eliminem a função proteica entre as estirpes invasivas associado ao facto de não se identificarem diferenças na atividade SpeB entre estirpes de infeções distintas, sugere que o mecanismo subjacente às mutações espontâneas no CovRS como causa de maior virulência, pode estar relacionado com outros fatores de virulência sob o controlo deste sistema para além do SpeB.

THESIS OUTLINE

The work described in the present thesis intended to evaluate the importance of CovRS and RopB regulators among *Streptococcus pyogenes* isolates recovered from distinct types of infections including in isolates from skin and soft tissue infections that were fully characterized through several typing methods and compared with contemporary invasive isolates.

The thesis comprises 4 chapters, organized as follows:

Chapter 1 corresponds to the general introduction that highlights the importance of *Streptococcus pyogenes* including a historical overview, the infections for which it is responsible and their burden, a description of the best known virulence factors, their mechanisms of action, their contribution to virulence and some of their regulators. It also includes a brief description of typing methods, therapeutic management of infections and mechanisms of antimicrobial resistance.

Chapter 2 is dedicated to the study of CovRS and RopB regulators and the consequences of their variability in several virulence factors, for which we specifically screened the strains. In this study, we used a collection of isolates from pharyngitis and from invasive infections to evaluate potential differences.

Chapter 3 consists in a detailed characterization of a collection of isolates recovered from skin and soft tissues infections using several typing methods and their comparison with a collection of invasive isolates recovered during the same period that were previously characterized. We also screened skin and soft tissue isolates for the presence of protease activity and sequenced *covRS* and *ropB* in a subset of isolates in order to understand their importance in this type of infections and to compare it with the results described in chapter 2.

Chapter 4 corresponds to the general discussion. In this chapter is provided a summary of the main results obtained in this thesis and its integrated discussion. It also includes perspectives for future work.

ABBREVIATIONS

ADP	Adenosine diphosphate
APSGN	Acute poststreptococcal glomerulonephritis
ARF	Acute Rheumatic Fever
ASO	Anti-Streptolysin O
cADPR	cyclic ADP-ribose
C4BP	C4-binding protein
CDC	Centers for disease control and prevention
CI	Confidence interval
CLSI	Clinical and Laboratory Standards Institute
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
ECM	Extracellular matrix
FCT	Fibronectin-binding, collagen-binding, T antigen
FDR	False discovery rate
FHL-1	Factor H-like protein 1
GAS	Group A streptococci
HBP	Heparin binding protein
HLA	Human leukocyte antigen
IFN	Interferon
Ig	Immunoglobulin
iGAS	Invasive GAS infections
IL	Interleukin
ITP	Invasive transcriptome profile
LTA	Lipoteichoic acid
M	Macrolides (resistance phenotype)
MGE	Mobile genetic element
MHC	Major histocompatibility complex
MIC	Minimum inhibitory concentration
MLS_B	Macrolides, lincosamides, streptogramins B (resistance phenotype) $cMLS_B$: constitutive MLSB $iMLS_B$: inducible MLSB

MLST	Multilocus Sequence Typing
mRNA	Messenger ribonucleic acid
MRSA	Methicillin resistant *Staphylococcus aureus*
NAD	Nicotinamide adenine dinucleotide
NADase	NAD glycohydrolase
NET	Neutrophil extracellular trap
PBPs	Penicillin Binding Proteins
PCR	Polymerase chain reaction
PMN	Polymorphonuclear leukocyte
PTP	Pharyngeal transcriptome profile
PYR	Pyrrolidonylarylamidase
QRDR	Quinolone resistance-determining region
RADT	Rapid antigen detection test
RALP	RofA-like protein
RHD	Rheumatic Heart Disease
rRNA	ribosomal ribonucleic acid
SAg	Superantigen
Sic	Streptococcal inhibitor of complement
SID	Simpson's index of diversity
SLO	Streptolysin O
SLS	Streptolysin S
SMEZ	Streptococcal mitogenic exotoxin Z
SOF	Serum opacity factor
Spe	Streptococcal pyrogenic exotoxin
SSA	Streptococcal superantigen
SSTI	Skin and Soft tissue infection
ST	Sequence type
STSS	Streptococcal toxic shock syndrome
TCR	T cell receptor
TCS	Two-component system regulator
TLR	Toll-like receptor
TNF	Tumor necrosis factor
WB	Western Blot

WGS	Whole-genome sequencing
WHO	World Health Organization
WT	Wild type

TABLE OF CONTENTS

CHAPTER IV

CHAPTER I

GENERAL INTRODUCTION

1. HISTORICAL BACKGROUND

According to a review about history of streptococcal research (Ferretti and Kohler, 2016), the first potential description of streptococcal infections dates to the 4th century BC, in the original writings of Hippocrates where are described the symptoms of childbed fever and erysipelas. In 1874, an Austrian surgeon named Theodor Billroth, described "small organisms as found in either isolated or arranged in pairs, sometimes in chains of four to twenty or more links (Streptococcus; Gr. Strepto, a chain and coccus, a berry)" in cases of erysipelas and wound infections. In 1879, Louis Pasteur established the real importance of streptococci when he isolated the microorganism from the uteruses and blood of women with puerperal fever and demonstrated that *Streptococcus* was the etiological agent responsible for the disease. At that time, puerperal fever was the cause of high mortality among women and newborns. The name "*Streptococcus pyogenes"* (Gr., pyo, pus, and genes, forming) came from Julius Rosenbach in 1884, due to his observations of bacteria recovered from suppurative lesions. Meanwhile, other names also emerged such as *eryespaltis*, *scarlatinae*, and *puerperalis*, mostly according to the disease associated. However, in 1932, due to the lack of unique characteristics of the microorganisms isolated from specific diseases, Andrews & Christie suggested that all previous species names be included in the single name *Streptococcus pyogenes*. Since 1903, with the introduction of blood agar plates by Hugo Schottmüller that is possible to differentiate streptococci based on their type of hemolysis. Those with a clear zone surrounding a colony were termed *Streptococcus haemolyticus*. Then, in 1933, Rebecca Lancefield developed the classification system, that is currently used, which allowed the differentiation of *Streptococcus haemolyticus* isolates in distinct groups (Lancefield, 1933). *Streptococcus pyogenes* corresponds to Lancefield group A antigen and therefore the name Group A *Streptococcus* (GAS) is commonly used as an alternative.

2. GENERAL FEATURES

Streptococcus pyogenes is a Gram positive bacterium. Individual cells are presented as spherical cocci with 1 to 2µm in diameter arranged in short chains in clinical specimens. When grown in liquid media longer chains are observed (Figure 1). It is a facultative anaerobe, catalase negative, with optimal growth on blood-enriched

agar media (Murray *et al*., 2013). After overnight incubation at 35-37°C on blood agar plates, characteristic white colonies with 1 to 2 mm surrounded by large clear zones of β-hemolysis are observed (Figure 1). Among the species belonging to the *Streptococcus* genus, three distinct patterns of hemolysis are observed: α-hemolysis (partial or green hemolysis), β-hemolysis (complete hemolysis), and γ-hemolysis (absence of hemolysis).

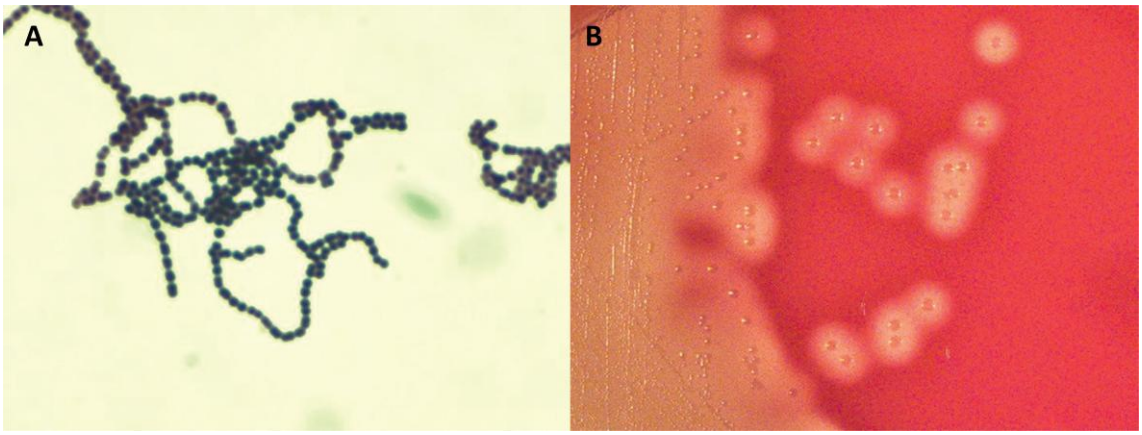


Figure 1. *Streptococcus pyogenes*. A- Chains with Gram stain. B- Colonies in blood agar with a surrounding area of β-hemolysis. Both images were reproduced from (Murray *et al*., 2013).

A minority of *Streptococcus pyogenes* strains exhibiting α-hemolysis and γ-hemolysis on blood agar were also reported, which were associated with the lack of ability to produce streptolysin S (SLS) (discussed later) (Yoshino *et al.*, 2010). Some strains are covered by a hyaluronic acid capsule that is antigenically indistinguishable from the hyaluronic acid in mammalian connective tissues (discussed later) (Murray *et al*., 2013). Additionally, within the cell wall of *Streptococcus pyogenes* is the group-specific carbohydrate, which is a dimmer of N-acetylglucosamine and rhamnose. This carbohydrate comprises about 10% of the dry weight of the cell and is the base of the serological classification developed by Rebecca Lancefield in 1933, which identifies *Streptococcus pyogenes* as Lancefield's group A streptococci (Lancefield, 1933). After the enzymatic extraction of the carbohydrate, a latex agglutination test with each specific group serum allows the identification of *Streptococcus pyogenes* (group A), *Streptococcus agalactiae* (group B), and other species (groups C, F and G) (Murray *et al*., 2013). There are some strains of *Streptococcus dysgalactiae* subsp. *equisimilis* and *Streptococcus anginosus* group that share the same group-specific carbohydrate with *Streptococcus pyogenes* reacting positively with group A serum, however these are uncommonly recovered from human infections (Facklam *et al.*, 2002).

Susceptibility to bacitracin was another test used for the identification of *Streptococcus pyogenes*. However, since the report of bacitracin resistant *Streptococcus pyogenes* isolates, including in Portugal, this test is no longer reliable (Malhotra-Kumar *et al.*, 2003; Pires *et al.*, 2009; Silva-Costa *et al.*, 2008). Another biochemical feature associated with *Streptococcus pyogenes* is its pyrrolidonylarylamidase (PYR) activity that can be tested by the PYR test. Among β-hemolytic streptococci, there are strains from other species susceptible to bacitracin or positive for the PYR test. However, *Streptococcus pyogenes* is the only species reported of having these both characteristics (Facklam, 2002).

3. INFECTIONS AND THEIR CONSEQUENCES

GAS is responsible for a variety of pyogenic infections ranging from mild superficial infections of the respiratory tract and skin, such as pharyngitis and impetigo, to extremely severe invasive infections, such as necrotizing fasciitis and streptococcal toxic shock syndrome (STSS) (Cunningham, 2000; Walker *et al.*, 2014). GAS infections have also been related to the development of postinfection autoimmune sequelae, which are responsible for high morbidity and mortality worldwide. These conditions include, mainly, acute rheumatic fever (ARF) and rheumatic heart disease (RHD), as well as acute poststreptococcal glomerulonephritis (APSGN) (Walker *et al.*, 2014; Cunningham, 2000). Before the advent of antibiotics, streptococcal diseases such as scarlet fever, erysipelas, and puerperal fever were considered major health problems for centuries, however their incidence and mortality rates started to fall just before penicillin's widespread use after the Second World War, suggesting that other factors, such as the host, the pathogen and the environment contributed to the decrease of the impact of these diseases (Efstratiou and Lamagni, 2016). Since the 1980s, it was observed a resurgence of severe invasive GAS infections with several reports of both suppurative and non-suppurative *S. pyogenes* sequelae. This re-emergence in the incidence of invasive GAS infections was associated mainly with the rise of the M1 clone, which is dominant among invasive *S. pyogenes* isolates in most developed countries (Efstratiou and Lamagni, 2016; Cunningham, 2000).

In addition to the wide range of suppurative infections and non-suppurative complications, GAS can also asymptomatically colonize the nasopharyngeal mucosa

and skin. Each of these sites represents the primary reservoirs of GAS (Cunningham, 2000). The ability of GAS to survive in human saliva is thought to allow the transmission from infected persons or asymptomatic carriers via respiratory droplets (Shelburne *et al.*, 2005). Additionally, the ability of GAS to colonize and persist in the skin allows transmission from person-to-person through skin contact. Furthermore, there are reports of GAS disease by food-borne outbreaks (Walker *et al.*, 2014).

According to the World Health Organization (WHO), GAS was considered the ninth leading infectious cause of human mortality, mainly attributed to invasive infections and RHD in developing countries where transmission is promoted by poor living conditions and where there is no access to a prompt and adequate antimicrobial treatment (Carapetis *et al.*, 2005).

3.1. Pharyngitis and Scarlet fever

GAS is the most common bacterial cause of pharyngitis, accounting for over 600 million cases annually among people aged over 4 years (Bisno, 2001; Carapetis *et al.*, 2005). In temperate climates, the incidence of streptococcal pharyngitis is highest during winter and early spring occurring most commonly among school-age children between 5 and 15 years of age, although all ages groups are susceptible (Wessels, 2011). The clinical symptoms of GAS pharyngitis include a sudden-onset of a sore throat, pain on swallowing, fever, chills, malaise, and headache (Bisno, 2001). Symptoms such as abdominal pain, nausea and vomiting can also be present, particularly in younger children (Bisno, 2001; Wessels, 2011). The signs that are frequently observed are tonsillopharyngeal erythema, patchy exudates, soft palate petechiae, beefy red and swollen uvula, and anterior cervical lymphadenitis (Bisno, 2001; Choby, 2009). Occasionally, GAS pharyngitis is also accompanied by scarlet fever, also known as scarlatina. This syndrome is thought that occur after a pharyngeal infection with a strain that secretes bacteriophage-encoded streptococcal pyrogenic exotoxins (SPE), mainly SpeA, SpeC and SSA (Silva-Costa *et al.*, 2014). Although uncommon, scarlet fever may occur due to GAS infections at other sites (Cunningham, 2000). Scarlet fever signs occur within 1 or 2 days after the manifestations of clinical symptoms of pharyngitis and include a characteristic diffuse erythematous rash that appears first on the upper chest and then spreads to the extremities sparing the area around the mouth, palms and soles, and a yellowish-white coating covering the tongue

that when is shed reveals a red and raw surface, which is commonly named "strawberry tongue". The rash disappears in the next 5 to 7 days followed by the desquamation of the superficial skin layer (Murray *et al*., 2013).

Usually, uncomplicated GAS pharyngitis is self-limiting even without specific treatment. Fever resolves within 3 to 5 days and throat pain resolves within 1 week (Wessels, 2011). However, occasionally untreated GAS pharyngitis can result in suppurative and nonsuppurative complications and therefore an adequate antibiotic therapy is recommended (Shulman *et al.*, 2012). Suppurative complications include bacteremia, cervical lymphadenitis, endocarditis, mastoiditis, meningitis, otitis media, peritonsillar/retropharyngeal abscess and pneumonia (Choby, 2009).

Transmission occurs by person-to-person contact, probably through nasal secretions or saliva droplets from carriers or infected individuals. Therefore, the highest incidence occurs among people attending crowded places such as schools and military training facilities (Cunningham, 2000). It is estimated that in developed countries, approximately 15% of school-age children and 4-10% of adults may suffer from a symptomatic episode of GAS pharyngitis annually (Carapetis *et al.*, 2005). In developing countries, these values are presumably 5 to 10 times higher (Carapetis *et al.*, 2005).

3.2. Impetigo

Impetigo (pyoderma) is a contagious, purulent skin infection of the superficial epidermis that manifests itself as numerous vesicles that rapidly progress into pustules, which then enlarge and rupture to form the characteristic thick, honey-colored crusts (Bisno and Stevens, 1996). Generally, pyoderma lesions occur on exposed areas such as the face and extremities (Bisno and Stevens, 1996). They are commonly accompanied by regional lymphadenopathy (Bisno and Stevens, 1996).

Impetigo usually heals spontaneously within two weeks without scarring. However, treatment relieves the discomfort, improves cosmetic appearance and prevents the spread of the organism (Cole and Gazewood, 2007). It is the third most common skin disease in children and, additionally to GAS, it can also be caused by *Staphylococcus aureus* (Sladden and Johnston, 2004). Transmission occurs through direct skin contact and most often affects children between 2 to 6 years old, particularly those who live in tropical and subtropical climates in areas with poor hygiene and

crowded living conditions (Sladden and Johnston, 2004). Accordingly, the highest prevalence is reported among aboriginal Australians and Pacific Islands nations (Carapetis *et al.*, 2005). Overall, in less developed countries, its estimated burden is over 111 million cases of pyoderma annually among children under 15 years (Carapetis *et al.*, 2005).

3.3. Invasive infections

Additionally to superficial infections, GAS has the ability to penetrate epithelial surfaces and produce several invasive disease manifestations ranging from less severe forms, such cellulitis, to diseases associated with high morbidity and mortality rates, such as necrotizing fasciitis or STSS (Cole *et al.*, 2011; Walker *et al.*, 2014). Others infections include bacteremia, puerperal sepsis, septic arthritis, pneumonia, meningitis, abscesses, osteomyelitis, endocarditis and peritonitis (Walker *et al.*, 2014). It is estimated that there are over 660 000 cases of GAS invasive infections each year, of which 97% occur in less developed countries, and over 160 000 deaths (Carapetis *et al.*, 2005).

Cellulitis and erysipelas are two conditions, both manifested by local signs of inflammation such as warmth, erythema, and pain, usually accompanied by fever and leukocytosis (Bisno and Stevens, 1996). Portals of entry may include local trauma or abrasions and psoriatic, eczematous, or tinea lesions (Bisno and Stevens, 1996). Erysipelas involves the superficial layers of the skin, while cellulitis affects the subcutaneous tissues. In contrast to cellulitis, in erysipelas, there is commonly a distinct demarcation of the area of inflammation, which rises above the surrounding normal skin. Although the two conditions are readily distinguishable when they occur in their typical form, there is a range of possible tissue involvement which can make this distinction difficult (Bisno and Stevens, 1996). Erysipelas on the legs and feet account for up to 85% of cases and in only about 5% of cases are the blood cultures positive (Chartier and Grosshans, 1990; Bisno and Stevens, 1996). Facial erysipelas may occur after GAS pharyngitis (Bisno and Stevens, 1996).

Necrotizing fasciitis, also popularly known as “flesh-eating disease”, is a severe GAS infection involving the skin, subcutaneous and deep soft tissue, and muscle. Within hours to a few days, the infection can progress for a small skin lesion, usually mistaken as an insect bite, to a highly lethal condition. This occurs as a result of rapid

bacterial growth and spread along the fascial sheaths that separate adjacent muscle groups, which are then breached, with the progression of infection, resulting in severe necrosis of adjacent tissues (Olsen and Musser, 2010). Rapid surgical debridement of infected tissue in the initial stages (12-24h) of the disease has been demonstrated to be essential for patient survival (Olsen and Musser, 2010). However, initially, patients most often present nonspecific signs such as fever, vomiting, and diarrhea which delays the correct diagnosis and contributes to the high morbidity and mortality associated to necrotizing fasciitis (Olsen and Musser, 2010). Mortality rates are reported to be 24%, 32% and 16% among necrotizing fasciitis cases in United States, Europe, and Japan, respectively (Lamagni *et al.*, 2008; O'Loughlin *et al.*, 2007; Lin *et al.*, 2013). The most frequently affected anatomical sites are the lower and upper extremities. However, cases in the head, neck, upper torso, and abdominal wall were also reported (Olsen and Musser, 2010). The progression of the disease can lead to a widespread muscle cell death named necrotizing myositis which is associated with a poor outcome (Olsen and Musser, 2010).

Several GAS infections may be further complicated by the development of STSS. However, there is a closer association with necrotizing fasciitis, which is present in 50% of the patients with STSS (Bisno and Stevens, 1996). The diagnostic criteria for STSS requires the isolation of GAS from a normally sterile site and hypotension that is refractory to adequate volume resuscitation, plus at least two of the following conditions: renal dysfunction, respiratory distress, hepatic dysfunction, coagulopathy, erythroderma, soft tissue necrosis with pain, tissue destruction and skin discoloration (Reglinski and Sriskandan, 2014). The development of STSS has been associated with the host response to superantigen (SAg) production by GAS (discussed later). STSS is responsible for high rates of mortality with 44% of patients dying within a week of developing the disease (Lamagni *et al.*, 2008). Blood cultures are positive in about 60% to 100% of patients with STSS (McCormick *et al.*, 2001).

Major risk factors for invasive GAS infections and STSS include disease or injury compromising the mucosal surface or skin barrier such as chicken pox, decubitus ulcers, penetrating injuries, minor cuts, burns, splinters and surgical procedures, childbirth, animal bites, and intravenous drug abuse (Bisno and Stevens, 1996; Olsen and Musser, 2010). However, in about one-half of STSS cases, there is no obvious point of entry (McCormick *et al.*, 2001). It is hypothesized that these cases are often preceded by blunt trauma, muscle strain, hematoma or joint effusion at the infection site what could

promote the seeding of bacteria at the site of injury in a patient with a transient bacteremia as a consequence of a superficial infection or colonization of the throat or skin (McCormick *et al.*, 2001; Johansson *et al.*, 2010; Stevens, 2000). According to this hypothesis, symptoms of a sore throat may occasionally precede STSS (Bisno and Stevens, 1996).

3.4. Acute rheumatic fever and rheumatic heart disease

ARF is a delayed systemic disorder that can occur within 1 to 5 weeks following an untreated GAS pharyngitis (Cunningham, 2000). Major clinical manifestations based on the updated Jones criteria include inflammation of the joints (migratory polyarthritis), of the heart (carditis), of the central nervous system (Sydenham´s chorea), and of the skin (erythema marginatum and/or subcutaneous nodules) (Cunningham, 2008). This disease is an autoimmune complication and results partially from the molecular mimicry between the M protein on GAS surface (discussed later) and components of human tissues, such as host cardiac myosin, which leads to the production of cross-reactive antibodies and T cells responsible for tissue destruction (Cunningham, 2008). Although arthritis is the most frequent manifestation, present in 60 to 80% of the cases, carditis, which is present in 30% to 45%, is the most serious because it affects the myocardium and the heart valves, and frequently results in mitral and/or aortic regurgitation, manifested in a heart murmur upon auscultation, that ultimately leads to valve replacement or death (Cunningham, 2000; Lee *et al.*, 2009). Sydenham's chorea is a neurological disorder, present in 10% of ARF cases manifested as involuntary movements, muscle weakness, and emotional disturbances. Erythema marginatum and subcutaneous nodules are less frequently observed. Erythema marginatum has been described in 2% of the cases and is characterized as a distinct red circinate rash (Cunningham, 2000; Lee *et al.*, 2009). The high rates of morbidity and mortality associated with ARF are mostly due to the long-term damage of the heart, resulting in RHD. It is estimated that RHD affects over 2.4 million children aged 5 to 14 and a total of at least 15.6 million people worldwide, with 282000 new cases and 233000 deaths annually (Carapetis *et al.*, 2005). The highest prevalence is found in indigenous populations and developing countries (Carapetis *et al.*, 2005). Streptococcal strains which commonly cause pyoderma do not lead to rheumatic fever (Bisno and Stevens, 1996).

3.5. Acute poststreptococcal glomerulonephritis

APSGN is an immune-mediated disorder mainly affecting the kidneys. Clinical presentation includes edema, hypertension, hematuria, urinary sediment abnormalities, and decreased serum complement levels, with little fever (Cunningham, 2000). Although some GAS antigens have been proposed to be related with APSGN, the exact nephritogenic components underlining the disease remain to be elucidated (Walker *et al.*, 2014). In fact, the nephritogenicity of GAS appears to be partially related with specific M protein serotypes, however not all strains the same M serotype are nephritogenic. APSGN is particularly associated with pyoderma and skin infections which occur in the summer of southern and temperate climates, however, both pharyngeal and skin infections can trigger the disease. The latency periods differ according to the type of preceding infection. After a skin infection, it may be 3 to 6 weeks, while after a throat infection it may be 1 to 2 weeks (Cunningham, 2000). Recurrences are rare. If supportive care is provided, long term renal damage due to APSGN is uncommon, with a mortality of 1% of the total cases. Even so, it is estimated that there are over 470 000 cases of APSGN each year, affecting mostly children and young adults, with approximately 5000 deaths, 97% of which in less developed countries (Carapetis *et al.*, 2005; Cunningham, 2000).

3.6. Pediatric autoimmune neuropsychiatric disorders

Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections, commonly named PANDAS, are very controversial. PANDAS include rare obsessive-compulsive disorders and Tourette´s syndrome in children and have been hypothesized to be a consequence of GAS infection. While some studies suggest that there may be an overlap in the immunopathogenic mechanisms of PANDAS and Sydenham chorea, others support that these identities are not related (Kirvan *et al.*, 2006; Murphy *et al.*, 2010; Walker *et al.*, 2014).

3.7. Poststreptococcal reactive arthritis

Other reported immune sequel of GAS infections is referred to as poststreptococcal reactive arthritis, which is distinct from the migratory polyarthritis observed in ARF (Jansen *et al.*, 1999). Usually, no cardiac involvement is observed in

patients with reactive arthritis (Iglesias-Gamarra *et al.*, 2001). The molecular mechanism is also not understood, although apparently it is also related to a possible molecular mimicry between the M protein and proteins in cartilage and synovium (Baird *et al.*, 1991).

4. CARRIAGE

Although GAS is not considered part of the normal microbiota, pharyngeal carriage can occur without clinical symptoms of disease (Cunningham, 2000). While pharyngitis can resolve in 7-10 days, asymptomatic carriage can persist for weeks or even months (Martin *et al.*, 2004). This can occur following resolution of clinical disease or with no antecedent history of clinical symptoms (Martin *et al.*, 2004). The definition of a GAS carrier is an asymptomatic individual with a positive throat culture and no serologic response (Martin *et al.*, 2004). However, most of the carriage studies do not look for serological response due to the technical issues related with the screening. The prevalence of carriage is variable according with the type of the study, reaching higher rates among longitudinal studies. According to data from a meta-analysis that highlight the substantial differences among studies, the prevalence of GAS carriage among children older than 5 years with no signs or symptoms of pharyngitis was 12% (Shaikh *et al.*, 2010). In the Lisbon area, it was reported an asymptomatically colonization rate of 10.7% in children and 3.3% in adults between 2000-2006 (Pires *et al.*, 2011).

The underlining mechanisms why some GAS isolates cause pharyngitis while others persist asymptomatically in the upper respiratory tract remain unclear. The correction of a mutation in a regulator responsible for a hyperencapsulated phenotype in M18 isolates resulted in reduced *in vivo* pharyngeal carriage duration in a murine model and was associated with a drop in bacterial airborne transmission during infection, supporting the potential role of capsule in the establishment of the carrier state (Lynskey *et al.*, 2013). However, it was also reported that the asymptomatic carriage of GAS is associated with the elimination of capsule production due to spontaneous mutations in the capsule encoding *has* operon (Flores *et al.*, 2014). The authors suggested that increased capsule production could contribute to acute symptomatic infection, but when

the infection resolves, loss of capsule allows for greater adherence and internalization resulting in a better adaptation to the colonization of the human host and persistence.

Skin carriage, has been reported in endemic conditions, however, there is very little information about it (Bessen *et al.*, 2014).

5. VIRULENCE FACTORS

The first step of bacterial infection is adhesion to the host. Numerous cell-surface proteins have been reported in GAS that work as adhesins promoting the interaction with multiple host components, such as extracellular matrix proteins (ECM) including fibronectin, collagen, and laminin (Walker *et al.*, 2014). Among the best studied are fibronectin binding proteins, such as M and M-like proteins, PrtF1, PrtF2, SOF and FbaA (Yamaguchi *et al.*, 2013). Moreover, the presence of lipoteichoic acid (LTA) on the cell surface has been proposed to contribute to the initial adherence of GAS to cell surfaces through the establishment of weak hydrophobic interactions that promote long distance attachment through long surface appendages, such as pili. This weak interaction is then replaced by stronger binding events after closer contact between GAS and the host cell (Hasty *et al.*, 1992; Walker *et al.*, 2014).

After adherence, in order to overcome host immune defenses, GAS have a large number of virulence factors, including some that also work as adhesins, with a wide variety of roles, including resistance to opsonophagocytosis by inhibiting complement deposition and activation, resistance to antimicrobial peptides and impairment of neutrophil killing mechanisms (Walker *et al.*, 2014). For several virulence factors distinct roles have been reported, which are often associated with distinct mechanisms of action and that in some cases are apparently contradictory.

In the next section is the description of some of the best known GAS virulence factors as well of others that are referred throughout the thesis.

5.1. M protein

The M protein is a major surface protein encoded by the *emm* gene and a major virulence factor of GAS. It is α-helical coiled-coil dimer composed of two polypeptide chains anchored in the cell membrane, traversing the cell wall and appearing as fibrils on the cell surface (Bisno *et al.*, 2003). The C (carboxy) terminal end of the molecule,

which is highly conserved, is attached to the cell wall by a LXPTG motif (Fischetti *et al.*, 1990). Upstream of the LPXTG motif is the wall spanning domain, which is rich in proline and glycine (Smeesters *et al.*, 2010a). The N (amino) terminal part of the molecule extends into the environment, and the tip is constituted by a hypervariable region which varies among different clinical isolates (Smeesters *et al.*, 2010b; Bisno *et al.*, 2003). Antigenic differences in this region constitute the basis for the Lancefield serological typing scheme developed 50 years ago to identify different GAS strains (Lancefield, 1962). More than 80 distinct serotypes were identified. Currently, with molecular technologies, this classification has been replaced by a sequence-based *emm* typing (discussed later). Briefly, it is performed through the PCR amplification of the 5' end of the *emm* gene that corresponds to the hypervariable region of the protein.

The M protein 6 (M6) was the first to be studied and provided an archetypal example of an M protein. Virtually, all M proteins are composed of a common framework, including a conserved signal peptide, the hypervariable amino terminus, a less variable central domain, and the highly conserved carboxy-terminus. Each chain of M6 protein comprises four repeat blocks (A, B, C and D repeats), each differing in size and amino acids sequence (Smeesters *et al.*, 2010b). Sequence conservation increases from A repeats to D repeats. The hypervariable region of the M6 protein is constituted by the non-helical part with 11 amino acids, along with a segment of the adjacent region A (Smeesters *et al.*, 2010b; Bisno *et al.*, 2003).

Lying immediately upstream and downstream of the *emm* gene, which is present in all GAS isolates, are two paralogous *emm*-like genes referred to as *mrp* and *enn*, respectively. Both these genes encode proteins similar to the M protein (M-like proteins), which form fibrils on the cell surface and have also a cell-wall spanning domain. However, this peptidoglycan-spanning domain is not identical for all M and M-like proteins. Therefore, based on the nucleotide sequences encoding the peptidoglycan-spanning domains it was possible to identify four major distinct forms or subfamilies (SF) of *emm* and *emm*-like genes: *emm* belong either to SF-1 or SF-2, *mrp* is always SF-4, and *enn* is SF-1 or SF-3. Furthermore, these subfamilies were found mostly (>99%) distributed into five distinct possible chromosomal arrangements, named *emm* patterns A through E (Bessen *et al.*, 2014; Hollingshead *et al.*, 1994; Hollingshead *et al.*, 1993; Bessen *et al.*, 1997). Due to the structural similarity between pattern A, B, and C, they were grouped together and are commonly referred to as pattern A-C. All isolates of this pattern have an SF-1 *emm* gene and lack *mrp*. Both patterns D and E have *mrp* and an

SF-3 form of *enn*. However, while pattern D has and SF-1 form of the *emm* gene, pattern E has a SF-2 form. The SF-1 form of the *emm* gene corresponds to a longer peptidoglycan-spanning domain than the SF-2, suggesting that isolates of pattern A-C and D have ticker cell walls than those belonging to *emm*-pattern E (Bessen *et al.*, 2014). Furthermore, a significant correlation between *emm*-pattern and preferred tissue site of infection was observed (Bessen *et al.*, 1996). While GAS isolates of the *emm* pattern A-C showed a statistically significant association with pharyngitis (throat specialists), isolates of the *emm* pattern D were associated with impetigo (skin specialists). GAS isolates of the *emm* pattern E were associated to both sites and therefore referred to as generalists (Bessen *et al.*, 1996). These observations led to the theory that there must be intrinsic properties related to *emm* patterns that are responsible for the tissue site preference (GAS tissue tropism). Moreover, isolates recovered from normally sterile sites belong mainly to *emm* pattern A-C, suggesting that these isolates are probably transmitted to new hosts by respiratory droplets and that the nasopharyngeal mucosa rather than impetiginous lesions, is the main reservoir of group A streptococci causing severe invasive disease in the population (Bessen *et al.*, 1997; Fiorentino *et al.*, 1997). Currently, the *emm* pattern was determinate for isolates of 184 different *emm* types and since a given *emm* type usually has the same *emm* pattern it is possible to predict the *emm* pattern of an isolate based on its *emm* type (McMillan *et al.*, 2013a; McGregor *et al.*, 2004). Overall, 21% of *emm* types are pattern A-C, 38%, pattern D and 37%, pattern E (McMillan *et al.*, 2013a). The association of *emm*-patterns with site of infection was reinforced in a meta-analysis of 5439 isolates whereby *emm* pattern group was inferred from the *emm* type (Bessen *et al.*, 2014). The results showed that overall pattern A-C strains represented 46.6% of pharyngitis isolates but only 8.2% of impetigo isolates. In contrast, *emm* pattern D strains represent 49.8% of impetigo isolates, but only 1.7% of pharyngitis isolates. Pattern E isolates account for almost equal fractions of throat and skin infections (51.7% and 42%, respectively).

The study of other M proteins along the years has identified significant differences comparing with M6 structural protein model (pattern A-C) and currently, it was updated with three other representative M proteins, which were selected as prototypes for the structural characteristics of each *emm* pattern (M5, M80, and M77) (Figure 2) (McMillan *et al.*, 2013a). In general, M proteins belonging to pattern A-C were the longest, followed by pattern D, and those of pattern E represented the shortest. The A repeats are more frequent amongst the pattern A-C and absent from the vast

majority of M proteins belonging to the pattern D and E groups. The B repeats are present in the majority of the pattern A-C and D but absent from most M proteins of the pattern E group. On the contrary, C repeated regions were present in all M proteins. Additionally, the non-helicoidal region in the amino-terminus was absent in 20% of the M proteins, including the M80 protein.

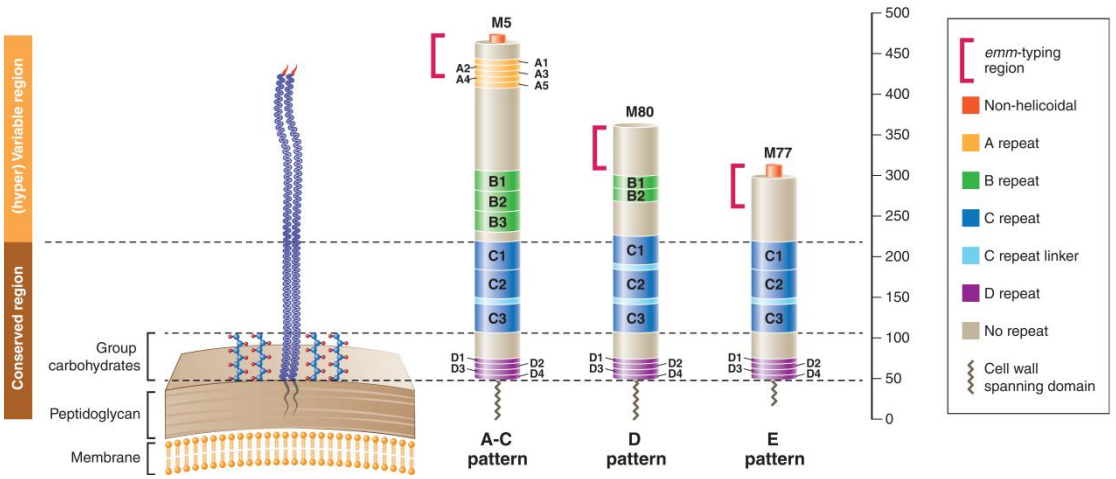


Figure 2. Three representative M proteins (M5, M80, and M77) were selected as prototypes for the structural characteristics within each *emm* pattern group. M protein length and the size of the repeat and non-repeat regions are drawn to scale. Reproduced from (McMillan *et al.*, 2013a).

Distinct M proteins can be further differentiated into two groups (class I and class II) depending on their immunodeterminants in the conserved C repeat domain. While isolates with M proteins of class I expose a surface domain that reacts with antibodies against the C repeated region, M proteins of class II lack this domain and therefore no reaction is observed (Bessen *et al.*, 1989; Bessen and Fischetti, 1990). This method of differentiation into the two classes is correlated with the expression of a streptococcal apoproteinase, named the serum opacity factor (SOF). SOF, is a surface protein that also has a secreted form that binds to fibronectin and enzymatically disrupts the structure of high-density lipoproteins in the blood (Courtney and Pownall, 2010). It is also the basis of a serological typing scheme that was used previously (discussed later). Thus, in general, streptococcal serotypes expressing class II M proteins are considered SOF positive and those expressing class I M proteins are considered SOF negative (Cunningham, 2000). Generally, class I/*sof* negative M proteins belong to the *emm* pattern A-C and D groups and class II/*sof* positive M proteins belong predominantly to *emm* pattern E group, although some exceptions have been noted (Bessen *et al.*, 2014).

The *emm* type of an M protein is considered largely predictive of the structure of the full-length protein (McMillan *et al.*, 2013a). Even so, there is size variation among M proteins of isolates of the same M/*emm* type, which was first reported in M6 isolates (Fischetti, 1989). It was reported that 81% of 80 different *emm* types showed intra-*emm*-type differences in the size of M proteins. Nonetheless, within each one of these *emm* types, an average of 69% of isolates belonged to the most common size variant (McMillan *et al.*, 2013a). Moreover, that intra-*emm*-type size variation is evenly distributed across the three *emm* pattern groups (McMillan *et al.*, 2013a). Size variation has been attributed to intragenic recombination because it has been shown that spontaneous M6 protein size variants can occur *in vitro* by deletions resulting from homologous recombination events between intragenic tandem repeats (Smeesters *et al.*, 2010b; Fischetti, 1989). Normally, opsonic antibodies directed against the variable portion of the M protein will confer type-specific protective immunity (Bisno *et al.*, 2003). Therefore, this size variation could confer a selective advantage, because it could result in the change of their antigenic profile and therefore protect against host antibodies (Oehmcke *et al.*, 2010). Thus, for certain M types, antibodies against the variable region confer strain-specific immunity but not necessarily serotype-specific immunity. In this way, protective immunity could be strain-specific rather type-specific (de Malmanche and Martin, 1994).

In general, M proteins interact with numerous host ligands, which include fibronectin, albumin, plasminogen, fibrinogen, kininogen, factor H, factor H-like protein 1 (FHL-1), C4b binding protein (C4BP), IgA, IgG (1,2,3,4), and the keratinocyte membrane cofactor CD46 (Smeesters *et al.*, 2010b). However, the specific interaction with the host ligands apparently depends on the M/*emm* type (Smeesters *et al.*, 2010a). Probably, this specificity is even strain-specific due to the structural diversity within M proteins of the same M/*emm* serotype/type. In addition, it was observed that an M1 protein at low temperature (20°C) present a coiled-coil dimeric configuration, with high-affinity binding to fibrinogen in the B repeat domain, and an unfolded state, with no binding activity at 37°C (Nilson *et al.*, 1995). Therefore, the environmental conditions could also influence the interactions between M protein and the host, preventing or promoting the infection (Smeesters *et al.*, 2010b). In temperate countries the incidence of clinical pharyngitis increases during winter and it has been shown that pharyngeal temperatures could drop to as low as 26°C, a fact that could be

partly mediated by changes in the M protein (Rouadi *et al.*, 1999; Smeesters *et al.*, 2010b).

It is well established that the M protein plays a critical role in GAS resistance to phagocytic killing by inhibiting complement deposition on the bacterial surface (Bisno, 1979; Courtney *et al.*, 1997; Moses *et al.*, 1997). This mechanism results from the M protein binding to complement-inhibitory proteins, including C4BP, factor H, and FHL-1. C4BP is an inhibitor of the classical complement pathway and it binds to the amino-terminal sequence of several M proteins (Johnsson *et al.*, 1996; Johnsson *et al.*, 1998). It was shown that C4BP retains its complement regulatory function when bound to M protein (Carlsson *et al.*, 2003). Furthermore, the binding of C4BP at the amino terminus competes with serotype-specific antibodies for target sites (Berggard *et al.*, 2001). Contrasting results are reported regarding the percentage of isolates that bind C4BP, which is probably due to the distinct composition of M serotypes in the collections tested. In one study, it was observed that most of the analyzed isolates were able to bind C4BP, accounting all *sof* positive isolates and 80% of *sof* negative isolates. Among the C4BP non-binding *sof* negative isolates, the majority belong to A-C pattern (Smeesters *et al.*, 2010b). In another study, only 33% of the isolates analyzed were able to bind C4BP (Perez-Caballero *et al.*, 2000).

Factor H and FHL-1, which is a splice variant of factor H, are both inhibitors of the alternative pathway of complement. While factor H binds to C repeat region of M protein, FHL-1 binds to the amino terminus (Johnsson *et al.*, 1998). Unlike FHL-1 and C4BP, factor H does not bind to the M protein under physiologic conditions (Perez-Caballero *et al.*, 2000), so the significance of this interaction for pathogenesis remains unclear. Considering C4BP, factor H, and FHL-1 together, it was reported that only 43.5% out of 69 isolates have the capacity to bind at least one of the three complement regulators and that 16% were able to bind all three (Perez-Caballero *et al.*, 2000).

Another antiphagocytic mechanism is the binding of fibrinogen to the M protein which decreases complement deposition resulting from the classical complement pathway (Carlsson *et al.*, 2005). The amino acid sequence of the fibrinogen binding motifs differs among M1 and M5 proteins, suggesting that these domains have evolved independently in different M-proteins (Ringdahl *et al.*, 2000). Moreover, binding to fibrinogen does not always inhibit the activation of the classical pathway. While fibrinogen binding to M5 reduced the amount of classical pathway C3 convertase on the bacterial surface (Carlsson *et al.*, 2005), fibrinogen binding to the M6 protein had little

or no effect on complement deposition or phagocytosis resistance (Horstmann *et al.*, 1992). Moreover, opsonization may not result in phagocytic ingestion. Additionally to the M protein, the hyaluronate capsule is also important in conferring resistance to opsonization and phagocytic killing. However, the importance of these two virulence factors is variably dependent of the strain (Dale *et al.*, 1996).

It was also observed that isolates with the affinity for fibrinogen do not bind C4BP and vice versa (Carlsson *et al.*, 2005). Therefore it was proposed that all M proteins may share ability to recruit either fibrinogen or C4BP, which inhibit complement deposition via the classical pathway (Carlsson *et al.*, 2005). Moreover, Mrp, which is a M-like protein, can also bind fibrinogen. In *sof* positive isolates, it is the Mrp that binds fibrinogen instead of the M protein (Courtney *et al.*, 2006). Additionally, upon expression of either an M5 or an M22 protein in a heterologous background, although both proteins have a normal structure and a normal ability to bind complement inhibitors, neither confers resistance to phagocytosis (Kotarsky *et al.*, 2000). These data reinforce the idea that the surface expression of the M protein alone may not be sufficient for GAS isolates to resist phagocytosis, highlighting the importance of other strain-specific factors (Kotarsky *et al.*, 2000).

The M protein can also be released from the cell surface possibly through the action of a secreted bacterial cysteine protease (discussed later) (Berge and Bjorck, 1995). The released M1 protein-fibrinogen complexes interact with β_2 integrins on the surface of neutrophils. As a result, these cells produce heparin-binding protein (HBP), which is a potent inflammatory mediator that induces vascular leakage (Herwald *et al.*, 2004). Streptococcal M1 protein can also interact with Toll-like receptor 2 on human peripheral monocytes, resulting in the expression of cytokines: interleukin (IL)-6, IL-1β and tumor necrosis factor α (Pahlman *et al.*, 2006). This response is significantly increased in the presence of HBP (Pahlman *et al.*, 2006). These M protein/fibrinogen complexes were identified in tissue biopsies from a patient with necrotizing fasciitis and STSS, supporting the importance of these complexes in severe streptococcal infections (Herwald *et al.*, 2004). In addition, IL-6 is significantly upregulated in severe streptococcal infections such as sepsis (Holub *et al.*, 2004). It was also shown that the soluble form of the M1 protein is a potent inducer of T cell proliferation and release of T-helper-1 type cytokines (Pahlman *et al.*, 2008). These observations suggest that the M protein has a superantigen activity. However, this theory remains controversial with several studies reporting conflicting results (Oehmcke *et al.*, 2010).

Human serum albumin binds to the C repeated domain of several distinct M proteins (Smeesters *et al.*, 2010b). However, this interaction with the M1 protein is dependent on the presence of the flanking B and D repeats regions strengthening the coiled-coil structure of the C repeated region (Gubbe *et al.*, 1997). It was reported that the binding of fibrinogen and albumin to B- and C-repeats, respectively, inhibits antibody binding under physiological conditions (Sandin *et al.*, 2006).

Several M proteins have also the ability to bind host plasminogen through their sub-N-terminal regions (Smeesters *et al.*, 2010b). Upon cleavage, plasminogen in the plasma turns into plasmin. Plasmin has the ability to degrade fibrin clots, connective tissue, ECM components, and adhesion proteins (Walker *et al.*, 2014). Human plasminogen can be activated by host factors as well by a GAS secreted virulence factor designated streptokinase (discussed later) (Ringdahl *et al.*, 2000). In addition to the M protein, other GAS surface proteins also bind plasminogen, including several M-like proteins. It was observed an enhanced mortality in "humanized" transgenic mouse expressing human plasminogen infected with GAS. Moreover, the susceptibility of these mice to GAS infection was dependent on the expression of bacterial streptokinase (Sun *et al.*, 2004). The host inflammatory response to bacterial infections possibly results in local thrombosis and microvascular occlusion. The formation of active plasmin on the bacterial surface may break the fibrinogen layer deposited by the host around the site of local infection allowing the spread of bacteria to deeper tissues (Sun *et al.*, 2004). In agreement, for GAS strains that express a plasminogen binding M protein, targeted elimination of the plasminogen binding motif reduces the capacity of GAS to cause lethal infection in the humanized plasminogen mouse model (Walker *et al.*, 2014). In addition, the plasminogen accumulation on the GAS surface prevents deposition of the opsonin C3b, impairing phagocytic uptake of GAS by neutrophils (Ly *et al.*, 2014). The presence of plasmin activity on the GAS surface was also shown to promote degradation of the human antimicrobial peptide cathelicidin LL-37 (Hollands *et al.*, 2012). Furthermore, plasminogen bound to the cell surface of GAS can act as a bridging molecule for bacterial interactions with host cells, for example contributing to keratinocyte invasion (Siemens *et al.*, 2011)

Several M proteins also have the ability to bind to immunoglobulins IgG and IgA (Smeesters *et al.*, 2010b). IgA is predominantly found in mucosal secretions and is most effective at impairing bacterial adhesion (Walker *et al.*, 2014). IgG is found in the bloodstream and injured tissues, facilitating the contact between the phagocytes and

bacteria via the Fc receptors on phagocytes (Walker *et al.*, 2014). The consequences of these interactions between M protein and immunoglobulins are still not clear but probably contribute to resistance to phagocytosis (Smeesters *et al.*, 2010b).

M protein on the GAS surface can also bind to kininogen, which is a precursor of the vasoactive peptide hormone bradykinin (Ben Nasr *et al.*, 1997). The cleavage of kininogen at the bacterial surface is associated with the release of bradykinin (Ben Nasr *et al.*, 1997). Bradykinin is responsible for several proinflammatory reactions, namely induction of vasodilatation, vascular leakage, and chronic pain (Oehmcke *et al.*, 2010). The role of this mechanism during infection is still not clear.

Furthermore, the M protein impairs phagosome maturation by inhibiting the fusion of azurophilic granules within the phagosome, increasing the survival of certain GAS strains inside human neutrophils (Staali *et al.*, 2006). Protein M1 was also shown to stimulate the formation of neutrophil extracellular traps (NETs). However, the same study showed that it also promotes resistance to the human cathelicidin antimicrobial peptide LL-37, an important effector of bacterial killing within such phagocyte extracellular traps (Lauth *et al.*, 2009)

The M protein is also among the multiple GAS adhesins that are involved in the initial interaction of GAS with the human host. It is suggested that binding of GAS to epithelial cells involves an interaction between M protein and host fibronectin as well with the keratinocyte membrane cofactor CD46 (Okada *et al.*, 1995; Cue *et al.*, 2001). Additionally, was also reported that most of the M proteins bind to glycosaminoglycans promoting the attachment of GAS to human epithelial cells and skin fibroblasts. The binding of M protein to fibronectin is apparently also related with the GAS ability to internalize into human epithelial lung cells (Oehmcke *et al.*, 2010).

5.2. Superantigens

SAgs are major virulence factors that belong to a family of low-molecular-weight extracellular toxins (23.6-27.4 kDa) with a highly potent mitogenic activity due to their ability to trigger excessive stimulation of human T lymphocytes resulting in a massive production of proinflammatory cytokines (Proft and Fraser, 2016).

To date, although there is some confusion with the nomenclature in the literature, eleven distinct GAS superantigens genes have been reported, including the streptococcal pyrogenic exotoxins (SpeA, SpeC, SpeG, SpeH, SpeI, SpeJ, SpeK, SpeL

and SpeM), the streptococcal superantigen (SSA) and streptococcal mitogenic exotoxin Z (SmeZ) (Proft and Fraser, 2016; Commons *et al.*, 2014). Of these, three are chromosomally encoded (SpeG, SpeJ, and SmeZ), while the remaining eight are encoded in temperate phages (SpeA, SpeC, SpeG, SpeH, SpeI, SpeJ, SpeK, SpeL; SpeM, and SSA) (Proft and Fraser, 2016). All superantigens have a signal N-terminal sequence which is cleaved at the bacterial cell surface to allow secretion (Sriskandan *et al.*, 2007). Originally, two additional superantigens were described. However, later it was noted that these two chromosomally encoded toxins named SpeB and SpeF were, in fact, a cysteine protease and a DNase, respectively (Gerlach *et al.*, 2001; Kapur *et al.*, 1993).

With the exception of *smeZ*, SAgs genes are well conserved and show only minor allelic variations although these may have important functional consequences. Currently, the known SAg variants are six *speA* alleles, three *speC*, six *speG*, two *speH*, two *speI*, three *speJ*, one *speK*, three *speL*, four *speM*, three *ssa*, and 56 *smeZ* (Proft and Fraser, 2016). This diversity could result in differences regarding the superantigenic activity and may also contribute to antigenic variation between alleles of the same toxin (Sriskandan *et al.*, 2007). Curiously several *smeZ* genes contain nonsense mutations resulting in the expression of truncated and inactive forms of this toxin (Turner *et al.*, 2012).

The mitogenic activity results from the ability of SAgs to simultaneously bind to major histocompatibility (MHC) class II molecules and to the variable β-chain of the T cell receptor (TcRVβ), thereby bringing MHC class II molecules and TCR together. In contrast to conventional peptides, the binding of SAgs to MHC class II occur outside the peptide-binding groove and does not require previous processing by antigen presenting cells. Thus, SAgs are able to bypass the conventional antigen presentation process, also resulting in the activation of T cells. However, the magnitude of T cell activation induced by superantigens exceeds by several logs that which can be obtained by conventional antigens (Sriskandan *et al.*, 2007).

SAgs bind to the variable region of the TcRVβ. Within the human T cell repertoire, there are less than 50 different TcR β-chains which are distributed among 25 major Vβ types. Even so, each SAg binds to more than one specific Vβ region and therefore is estimated that a single SAg is capable of activating up to 25% of an individual´s T cell repertoire (Proft and Fraser, 2016; Reglinski and Sriskandan, 2014). The oligoclonal activation of T cells and antigen presenting cells results in a massive

release of proinflammatory cytokines, including IL-1β, tumor necrosis factor alpha (TNF-α), and T cell mediators, such as IL-2 (Proft and Fraser, 2016). The unrestrained release of these cytokines could result in the activation of complement, coagulation, and fibrinolytic cascades ending in hypotension and the widespread organ damage characteristic of STSS (Bisno *et al.*, 2003).

Regarding MHC class II, SAgs can bind to the invariant α-chain or the polymorphic β-chain. With the exception of SpeA and SSA, all other nine Streptococcal SAgs binds the polymorphic β-chain (Proft and Fraser, 2016).

Recently, was been shown that SAgs can also bind CD28, which is a co-stimulatory receptor constitutively expressed on T cells, which apparently are crucial for the superantigenic activity of at least SAgs that bind to the MHC class II α-chain (Proft and Fraser, 2016). In addition, SpeA and SpeC, SpeJ and SSA are capable of forming monomeric dimers. However, the physiological relevance of this dimerization remains to be elucidated (Proft and Fraser, 2016; Commons *et al.*, 2014).

Little is known about the regulation of *Streptococcus pyogenes* SAg production, but under typical laboratory growth conditions, only small amounts of SAgs are secreted. A significant upregulation is detected after infection, which is apparently related to host factors that are still unknown. The expression of *speA* increases 7 days after the subcutaneous inoculation into BALB/c mice and it remains high after 21 in vitro passages (Kazmi *et al.*, 2001). The *in vitro* expression of *speC* is increased in culture with the presence of human pharyngeal cells (Broudy *et al.*, 2001). Analysis using DNA microarrays showed a significant increase in *speA*, *speG*, *speJ* and *smeZ* transcripts of an M1 strain growing in human blood, compared to growth in medium (Graham *et al.*, 2005). Additionally, microarrays showed that *speA*, *speJ*, and *smeZ* were highly expressed in distinct phases during infection of cynomolgus macaques. Of all GAS SAgs, SmeZ was reported to be at least 10 times more potent than any other in the ability to stimulate cytokine release by human T cells (Proft and Fraser, 2016).

The presence of SAgs has been associated with the severity of GAS infections. The predominant serotypes/*emm* types recovered from STSS, frequently produce SpeA and SpeC (Talkington *et al.*, 1993; Yu and Ferretti, 1989). It was also reported a strong correlation of M1 isolates carrying *speA* and *speC* with SSTS (Vlaminckx *et al.*, 2003). Consistently, several studies observed that *speA* was more frequent in isolates recovered from invasive disease rather than in isolates from non-invasive disease (Cleary *et al.*, 1992; Kittang *et al.*, 2011; Hauser *et al.*, 1991; Friães *et al.*, 2012). In contrast, there are

also some reports that showed no significant differences in the frequency of *speA* between invasive isolates and isolates from carriage or pharyngitis (Mylvaganam *et al.*, 2000; Hsueh *et al.*, 1998; Descheemaeker *et al.*, 2000). More rarely, SAg genes *speK*, *ssa*, *speJ* and *speM* have also been associated to invasive isolates (Ikebe *et al.*, 2002; Murakami *et al.*, 2002; Friães *et al.*, 2012; Lintges *et al.*, 2010).

5.3. Hyaluronic acid capsule

The majority of GAS isolates produce a capsule that consists of a hyaluronic acid polysaccharide, which is a linear, high-molecular-mass polymer of glucuronic-β-1,3-*N*-acetylglucosamine. The capsule is produced by enzymes encoded in the highly conserved *hasABC* synthase operon and is upregulated in the presence of human blood. The first gene of the operon (*hasA*) encodes a membrane-associated hyaluronate synthase which is essential for capsule biosynthesis. On another hand, *hasC* is not essential for the capsule synthesis. In M1T1 isolates a second UDP-glucose 6-dehydrogenase encoded by *hasB2* is located outside the *hasABC* operon and can replace that encoded by *hasB* (Walker *et al.*, 2014). The glucuronic-β-1,3-*N*-acetylglucosamine is structurally identical to the hyaluronic acid expressed on human cell surfaces and connective tissue, which can contribute for evasion of the host immune response through molecular mimicry (Cunningham, 2000). It was shown that the hyaluronic acid capsule facilitates adherence to the oropharyngeal epithelium by binding human CD44 (Cywes *et al.*, 2000). However, it was also reported that the upregulation of capsule synthesis reduces the capacity to bind to epithelial cell layers (Hollands *et al.*, 2010). Although the exact role of capsule in adherence, colonization, and persistence remain controversial, there is much more consensus about the role of capsule in the following steps of infection. The mucoid phenotype of GAS colonies is attributed to the higher expression of capsule and strains with this phenotype are frequently isolated from patients with GAS invasive disease (Cole *et al.*, 2011). Consistently, the capsule was shown to inhibit complement deposition, to promote resistance to opsonophagocytosis and to enhance GAS M1T1 survival within NETs by inhibition of human cathelicidin antimicrobial peptide LL-37 (Cole *et al.*, 2011). It was also shown to be essential for full virulence in mouse subcutaneous and intraperitoneal models as well in a nonhuman primate model of GAS invasive disease (Ashbaugh *et al.*, 1998; Wessels *et al.*, 1991; Moses *et al.*, 1997). In 2012, it was reported the circulation of M4 and M22 GAS strains lacking the *hasABC* operon and consistently none of these strains produced detectable

hyaluronic acid capsule. Even so, these strains were shown to proliferate *ex vivo* in human blood. The authors suggest that these strains may have alternative compensatory mechanisms to promote virulence (Flores *et al.*, 2012). More recently, it was reported of a new highly successful acapsular *emm*89 clade that became rapidly dominant in multiple geographic regions including Portugal (Turner *et al.*, 2015; Friães *et al.*, 2015). These isolates present a variant of the *nga* promoter region, which is associated with increased production of the secreted cytolytic toxins NADase and Streptolysin O (SLO). Due to the high virulence of these isolates in a mouse model of soft-tissue infection, it was suggested that the capsule is dispensable if toxin levels are high (Zhu *et al.*, 2015). On the other hand, *in vitro* studies indicate that isolates belonging to this clade may be at an advantage for environmental persistence and transmission, resulting in increased numbers rather than severity of infections (Turner *et al.*, 2015).

5.4. Pili

In 2005, Mora *et al*. described four variant of long surface extended pilus-like structures in GAS strains of several serotypes (Mora *et al.*, 2005). They also showed that these structures actually corresponded to the trypsin resistance T antigens described by Lancefield more than 60 years ago (Lancefield and Dole, 1946). The variability of the T antigens constituted the basis for the Lancefield serological T-typing system (discussed later). These pili, are heteropolymeric structures consisting of a backbone protein and either 1 or 2 ancillary proteins covalently joined and linked to the cell wall by a series of sortase mediated transpeptidase reactions (Mora *et al.*, 2005).

The pilus is encoded in a discrete 11-16Kb highly recombinogenic zone of the GAS chromosome named FCT region, based on the presence of fibronectin- and collagen-binding proteins as well as the T antigen (Bessen and Kalia, 2002). Currently, there are 9 different variants described of the pilus gene region, designated FCT types 1-9 (Kratovac *et al.*, 2007; Falugi *et al.*, 2008). There is a hypothesis put forward that the gene products coded in FCT region contribute for tissue site preference for infection (Kratovac *et al.*, 2007). The FCT typing combined with the ability to produce biofilm was also proposed as parameters in GAS epidemiological studies (Koller *et al.*, 2010)

Furthermore, there is evidence that immunization of mice with a combination of recombinant pilus proteins confer protection against mucosal challenge in a mouse model of infection (Mora *et al.*, 2005). According to Falugi *et al.*, a vaccine comprising a combination of 12 backbone variants would protect against more than 90% of the

circulating strains (Falugi *et al.*, 2008). Moreover, some T types share sufficiently high similarity levels, thereby cross-protection could potentially occur, allowing the reduction of the number of backbone proteins required for obtaining a wide coverage (Dale *et al.*, 2016).

Several roles have been attributed to pili which are mainly associated with GAS adherence to host cells, colonization and biofilm formation (Manetti *et al.*, 2007; Abbot *et al.*, 2007)**.** These roles are attributed mainly to the activity of the ancillary protein 1 (Cpa) of a FCT-1 pilus variant (Becherelli *et al.*, 2012). Nagata *et al*., showed that the disruption of pilus assembly by sortase deletion resulted in a more aggressive phenotype in a dermonecrotic mouse infection model, suggesting that pilus expression could result in a less virulent phenotype (Nakata *et al.*, 2009). The pilus-encoding genes in the FCT region are controlled mainly by RofA/Nra and MsmR, which are two stand-alone regulators encoded in the same locus (Nakata *et al.*, 2005; Kreikemeyer *et al.*, 2007). Moreover, it was reported a higher expression of the pilus backbone at 30°C compared with 37°C, which would be consistent with a more important role of pili in superficial skin infections (Nakata *et al.*, 2009; Abbot *et al.*, 2007).

5.5. Streptococcal pyrogenic exotoxin B

Streptococcal pyrogenic exotoxin B (SpeB) is a potent broad-spectrum cysteine protease secreted by GAS (Carroll and Musser, 2011). The *speB* gene is chromosomally located and even though it is found in almost all isolates, not all express a detectable SpeB activity (Chaussee *et al.*, 1996; Talkington *et al.*, 1993). The production of fully active, mature SpeB, is a complex multistage process involving a tight regulatory control including post-transcriptional and post-translational mechanisms that are influenced by at least 21 gene products (Carroll and Musser, 2011). Moreover, these mechanisms are also influenced by several environmental signals, including changes in pH and electrolytes (Loughman and Caparon, 2006). The transcription of *speB* occurs from two promoters and is directly regulated by RopB, CcpA, and probably by CovRS. Following transcription and translation, the SpeB protein is secreted as an inactive 42-kDa zymogen, which is then autocatalytically processed through a series of at least 8 intermediates to form the 28-kDa mature active protease (Chen *et al.*, 2003; Carroll and Musser, 2011). Mature SpeB can cleave many GAS proteins, as well host proteins, with apparently opposing impacts in virulence (Table 1). The expression of SpeB has been

shown to contribute for the survival and proliferation at the site of localized skin infections (Svensson *et al.*, 2000; Cole *et al.*, 2006). It was also shown to contribute for persistence and growth of GAS in human saliva *in vitro* (Shelburne *et al.*, 2005). However, the role of SpeB during infection is still not clearly understood. While several studies suggest that proteolytic action of SpeB contributes to virulence, others suggest that its downregulation is a critical step for the transition from localized to invasive disease. In agreement with the first hypothesis, it was reported that among 6775 clinical isolates, 84.3% presented a wild-type SpeB protease phenotype *in vitro* (Olsen *et al.*, 2015). SpeB was also suggested to confer protection against the antibacterial peptide LL-37 (Nyberg *et al.*, 2004a; Johansson *et al.*, 2008). Neutralization of SpeB was significantly lower in acute-phase sera of patients with STSS than in sera from patients with bacteremia or erysipelas (Eriksson *et al.*, 1999). Similarly, it was also reported that the acute-phase serum of patients with GAS bacteremia had significantly lower neutralizing ability against SpeB than did sera from patients with uncomplicated tonsillitis (Norrby-Teglund *et al.*, 1994). In addition, immunization of mice with SpeB or treatment with a protease inhibitor confers protection against lethal challenge (Bjorck *et al.*, 1989; Kapur *et al.*, 1994; Ulrich, 2008). In contrast with the previous results, an inverse relation between disease severity and expression of SpeB among M1T1 isolates recovered from invasive human infections has reported. Among STSS isolates, 41% produced little or no SpeB compared to only 14% of isolates recovered from less severe cases. Moreover, the cysteine protease activity among those isolates that expressed SpeB was significantly lower for STSS isolates than for isolates from less severe cases (Kansal *et al.*, 2000). The downregulation of SpeB production has been considered mainly a consequence of spontaneous *covRS* mutations in M1 isolates (discussed later). In this way, several GAS virulence factors are preserved from proteolytic degradation by SpeB and therefore contribute for the transition from localized to invasive disease (Walker *et al.*, 2007; Cole *et al.*, 2006; Aziz *et al.*, 2004b).

Table1. Host and bacterial proteins cleaved by SpeB. Table reproduced from (Chiang-Ni and Wu, 2008)

Before SpeB cleavage	**After SpeB cleavage**	**Potential effects**	**Reference:**
Host proteins			
Interleukin-1β precursor	Active interleukin-1β	Induces inflammation	(Kapur *et al.*, 1993)
Pro-matrix metalloprotease	Active MMP (2,9)	Enhances tissue damage and bacterial invasion	(Burns *et al.*, 1996; Tamura *et al.*, 2004)

(pro-MM2, 9)			
Fibronectin	Fragmented	Participates in bacterial colonization and invasive infection	(Kapur *et al.*, 1993)
Vitronectin	Degraded	Enhances tissue damage	(Kapur *et al.*, 1993)
Kininogen	Bradykinin	Increases vascular permeability: induce fever and pain	(Herwald *et al.*, 1996)
Immunoglobulin (IgA, IgM, IgD, IgE, IgG)	Cleavage into Fc and FAb fragments	Inhibits immunoglobulin-mediated opsonophagocytosis	(Eriksson and Norgren, 2003; Collin and Olsen, 2001; Collin *et al.*, 2002)
C3b	Degraded	Escapes phagocytosis	(Terao *et al.*, 2008)
Plasminogen	Degraded	Reduces plasmin activity on GAS surface	(Cole *et al.*, 2006)
Bacterial Proteins			
Zymogen form of SpeB	Active SpeB	Degrades or cleaves of bacterial and host proteins	(Liu and Elliott, 1965; Chen *et al.*, 2003; Hauser and Schlievert, 1990; Nomizu *et al.*, 2001)
M protein	Remove 24 amino acids from N-terminus; released from bacterial surface	Alters immunoglobulin binding properties; promote bacterial dissemination	(Raeder *et al.*, 1998; Berge and Bjorck, 1995; Kansal *et al.*, 2003)
Protein F1	Degraded	Reduced bacterial internalization	(Chaussee *et al.*, 2000; Nyberg *et al.*, 2004b)
Endo S	Degraded	Loses IgG glycan-hydrolyzing activity	(Allhorn *et al.*, 2008)
SmeZ	Degraded	Abolishes immune stimulatory activity	(Nooh *et al.*, 2006)
Fba	Degraded	Inhibit binding of FH and FHL-1	(Wei *et al.*, 2005)
C5a peptidase	Released from bacterial surface	Degrades chemotactic complement factor C5a	(Berge and Bjorck, 1995; Ji *et al.*, 1997; Ji *et al.*, 1996)
Streptokinase	Degraded	Unknown	(Johnston and Zabriskie, 1986; Svensson *et al.*, 2002)
Protein H	Released from bacterial surface	Promotes bacterial dissemination	(Berge and Bjorck, 1995)
Sda1	Degraded	Decreases neutrophil extracellular trap clearance	(Walker *et al.*, 2007)

5.6. Streptolysin O

Streptolysin O (SLO) is a 69-kDa oxygen-labile secreted cytolysin which is inhibited by small amounts of cholesterol. Its activity is responsible for the characteristic GAS β-hemolysis under the surface of blood agar medium (Molloy *et al.*,

2011). The screening of anti-streptolysin O (ASO) antibodies in serum of patients is one of the most common serological tests used to confirm an antecedent streptococcal infection. Curiously, infection of the skin does not always elicit a strong ASO response (Cunningham, 2000). SLO oligomerizes to form large pores in host cell membranes, therefore promoting resistance to phagocytic killing by compromising the integrity of host cell membranes and cause caspase-dependent apoptosis in neutrophils, macrophages, and epithelial cells (Walker *et al.*, 2014). In addition to protecting GAS from phagocytic killing, expression of SLO also reduces the host inflammatory responses by reducing the production of TNFα and IL-β. Moreover, in M1 isolates, SLO has been shown to induce the aggregation between platelet and neutrophils, contributing to vascular occlusion and tissue damage. Several studies reported attenuated virulence of SLO mutants in mouse models of invasive disease, supporting the key role of this virulence factor during infection (Walker *et al.*, 2014). In agreement, the expression of SLO is higher in GAS isolates recovered from severe invasive clinical cases rather than isolates of the same serotype recovered from non-invasive infections (Ato *et al.*, 2008).

5.7. Nicotinamide glycohydrolase

Nicotinamide glycohydrolase (NADase) is a secreted protein that catalyses the hydrolysis of the nicotinamide-ribose bond of NAD^+ resulting in nicotinamide and adenosine diphosphoribose (ADP-ribose). NADase is coexpressed with SLO, and penetrates into the cytoplasm of epithelial cells through SLO pores depleting intracellular energy stores (Madden *et al.*, 2001). The toxic effect of NADase on keratinocytes *in vitro* is associated with cytotoxic effects and induction of apoptosis (Bricker *et al.*, 2002). Due to a side reaction, NADase activity can also results in an alternative enzymatic product, named cyclic ADP-ribose (cADPR), which has been shown to act as a signaling molecule in eukaryotic cells, triggering Ca^{2+} entry or release from intracellular stores and may contribute to cytotoxic effects of SLO and NADase (Bricker *et al.*, 2002). Furthermore, in a mouse model of GAS invasive disease, it was reported that isogenic mutants lacking NADse and SLO, were significantly attenuated in virulence compared with the WT parent, confirming that both proteins enhance GAS virulence *in vivo* (Bricker *et al.*, 2005).

5.8. Streptolysin S

Streptolysin S (SLS) is an oxygen-stable cytolytic toxin responsible for the characteristic zone of β-hemolysis surrounding the GAS colonies on the surface of blood agar medium (Molloy *et al.*, 2011). SLS is encoded by the *sag* operon comprising nine genes from *sagA* to *sagI* and is secreted by virtually all GAS isolates. SLS expression enhances GAS resistance to phagocytosis through the formation of hydrophilic pores in the membrane of neutrophils. In addition to erythrocytes and neutrophils, SLS is also able to lyse lymphocytes, platelets and several other host cells (Walker *et al.*, 2014). It was proposed that in invasive soft tissue infection SLS may significantly impair inflammatory cells locally and facilitate the spread of streptococci in tissue (Heath *et al.*, 1999). Using mouse models of invasive GAS infection several studies reported that SLS-deficient mutants, belonging to different M serotypes are markedly less virulent than their isogenic parents (Betschel *et al.*, 1998; Miyoshi-Akiyama *et al.*, 2005; Engleberg *et al.*, 2004; Molloy *et al.*, 2011).

5.9. Streptokinase

Streptokinase (Ska) is a secreted plasminogen activator encoded by the *ska* gene. Ska is highly specific for human plasminogen, which upon cleavage originates activated plasmin, that has the ability to degrade fibrin clots, connective tissue, ECM components, and adhesion proteins, possibly contributing to the spread and dissemination of GAS as discussed before. Phylogenetic studies have shown that *ska* variants are distributed among two 2 main sequence clusters. Furthermore, while strains containing cluster 1 *ska* alleles display soluble plasminogen activation, in strains containing cluster 2 *ska* alleles, this feature could only be detected when plasminogen was prebound with fibrinogen (McArthur *et al.*, 2008). The requirement of human fibrinogen additionally to Ska and plasminogen by some GAS isolates for the acquisition of plasmin activity is related with an indirect plasminogen binding pathway. This alternative pathway involves the formation of a trimolecular complex comprising streptokinase, plasminogen, and fibrinogen (Wang *et al.*, 1995b; Wang *et al.*, 1995a). Therefore the expression of fibrinogen binding proteins, such as several M proteins, PrtF1 and PrtF2 can also contribute to the accumulation of plasminogen on the GAS surface. Supporting the importance of the indirect pathway, it was reported, for a representative isolate, that the presence of fibrinogen and streptokinase resulted in a

significant increase in plasminogen binding. Moreover, this enhancement of plasminogen binding did not occur when streptokinase alone or fibrinogen alone was added (McKay *et al.*, 2004). The same study reported that in the presence of streptokinase and fibrinogen, isolates from invasive infections bound more plasminogen than isolates from uncomplicated infections.

5.10. Extracellular streptodornase D

Streptodornase D, also named Sda1 or SdaD2 is a secreted DNase encoded in a prophage commonly found among M1T1 isolates (Aziz *et al.*, 2004a). The expression of Sda1 was shown to protect GAS against neutrophil killing by degrading the DNA scaffold of NETs (Buchanan *et al.*, 2006; Sumby *et al.*, 2005a). In 2004, NETs were reported for the first time as DNA-structures secreted by host neutrophils structures containing microbicidal effectors such as histones, the granule proteases elastase and myeloperoxidase, and cathelicidin antimicrobial peptide LL-37 (Brinkmann *et al.*, 2004). Sda1 was also shown to suppressed both the TLR-9 mediated immune response and macrophage bactericidal activity (Uchiyama *et al.*, 2012).

5.11. Streptococcal inhibitor of complement

Streptococcal inhibitor of complement (SIC), is a highly polymorphic 31kDa secreted protein that prevents the formation of the membrane attack complex through the binding and inactivation of complement C5b67 complex. Several roles were reported for SIC, including that it contributes to epithelial cell adherence and mucosal colonization. It was also shown to bind innate immune factors such as secretory leukocyte protease inhibitor, lysozyme, human α-defensin-1, and human cathelicidin antimicrobial peptide LL-37 (Walker *et al.*, 2014).

5.12. C5a peptidase

C5a peptidase, also named SCPA, is a protease located on the cell surface that cleaves the polymorphonuclear leukocyte binding site of chemotactic peptide C5a. The cleavage of C5a abolishes its chemoattractant activity, impairing the recruitment of phagocytes to the site of GAS infection. Interestingly, another GAS surface protein (GAPH/PLR/SDH) binds to human C5a, which apparently facilitates the action of SCPA (Walker *et al.*, 2014).

5.13. SpyCEP

SpyCEP is cell-wall protease that cleaves and inactivates the host CXC chemokine IL-8 impairing neutrophil recruitment to the infection site. Is has been shown to increase GAS resistance to neutrophil killing and promote virulence in a murine model of systemic infection (Walker *et al.*, 2014). Additionally, SpyCEP activity is associated with the severity of invasive disease (Turner *et al.*, 2009).

5.14. GRAB

GRAB is a surface protein that binds to the α2-macroglobulin, which is a major proteinase inhibitor of human plasma. In the GAS surface, α2-macroglobulin entraps and inhibits GAS and host proteinases, protecting bacterial surface proteins and virulence factors from degradation. However, some proteinases, such as SpeB, remain proteolytically active against small peptides that can enter the α2-macroglobulin complex. Thereby, it was reported that α2-macroglobulin-SpeB complexes formed on the surface of GAS through protein GRAB confer protection from killing by LL-37 (Nyberg *et al.*, 2004a).

6. HOST FACTORS AND SUSCEPTIBIBLITY TO INFECTION

Several studies suggest that the susceptibility to GAS invasive disease is related to the genetic background of the host, in particular with HLA polymorphisms. These could explain why distinct individuals infected with high related isolates can develop very different manifestations (Chatellier *et al.*, 2000). It was reported that specific HLA class II haplotypes conferred strong protection from severe systemic disease, whereas others increased the risk of severe disease. For example, it was observed that patients with the HLA-DRB1*1501/DQB1*0602 haplotype presented significantly reduced responses to streptococcal SAgs and therefore were less likely to develop severe systemic disease compared to individuals with higher risk or neutral haplotypes (Kotb *et al.*, 2002). Additionally, other study showed that alleles, such as HLA-DR14/DR7/DQ5 are considered of high-risk, while alleles such as HLA-DR15/DQ6 not only strongly protected against severe invasive *S. pyogenes* disease but also resulted in significantly

higher amounts of anti-inflammatory cytokines, such as IL-10, compared to pro-inflammatory cytokines, like IFN-γ (Nooh *et al.*, 2011).

7. VIRULENCE REGULATION MECHANISMS

Currently, 13 two-component regulatory systems are described and more than 100 putative stand-alone transcriptional regulators were found in the available GAS genomes (Walker *et al.*, 2014). Among these, the most studied include the two-component regulatory systems CovRS and Ihr-Irr, and the stand-alone regulators RopB and Mga.

The Ihk-Irr two-component regulatory system has been suggested to control genes involved in cell wall synthesis and resistance to oxidative stress. Its inactivation significantly attenuates streptococcal virulence in mouse models of soft tissue infection and bacteremia (Voyich *et al.*, 2004).

The Mga regulator activates the transcription of multiple virulence genes such as the M protein family (*emm, mrp, arp, and enn*), *sof*, and *sic* (Walker *et al.*, 2014). Two divergent alleles were described (*mga-1,mga-2*) which were correlated with different tissues sites (Bessen *et al.*, 2005). Given its role in controlling the expression of M protein, it was reported to be critical for multiple roles, such as biofilm formation, growth in human blood and resistance to phagocytosis (Hondorp *et al.*, 2013). Moreover, its function is altered depending on sugar availability, by modulating the expression of several colonization factors (Hondorp *et al.*, 2013). Furthermore, Mga mutants were attenuated in a model of GAS invasive skin disease (Hondorp *et al.*, 2013).

Regarding the CovRS regulatory system and RopB, they will be described in more detail in the next sections.

7.1. CovRS two component regulatory system

The CovRS system is a two component regulatory system that belongs to the OmpR family and was initially named CsrRS because was first described as a regulator of capsule synthesis. When inactivated it results in a mucoid colony phenotype associated with the overexpression of hyaluronic acid capsule (Levin and Wessels, 1998). Using DNA microarrays and quantitative RT-PCR, it was observed that CovR

influences directly or indirectly the transcription of between 10 to 15% of all chromosomal genes during *in vitro* growth, including several virulence factors (Sumby *et al.*, 2006; Graham *et al.*, 2002). Therefore, these proteins were renamed as control of virulence regulator (CovRS). It is presumed that the protein CovS acts as a membrane located histidine kinase, whose phosphorylation status is influenced by environmental signals, while CovR is the cytoplasmic DNA-binding response regulator that binds to specific promoters (Cole *et al.*, 2011). Under normal growth conditions, CovS probably acts as a kinase phosphorylating CovR, primarily repressing the expression of a subset of genes encoding virulence factors. In contrast, under stress conditions, such as elevated temperature, acidic pH, high osmolarity, presence of LL-37 and iron starvation, CovS alters the phosphorylation state of CovR, resulting in the derepression of genes encoding virulence factors as well as genes encoding for proteins essential for growing under stress conditions (Dalton and Scott, 2004; Froehlich *et al.*, 2009). CovRS regulation may also be influenced by environmental Mg^{2+} (Gryllos *et al.*, 2003) and could also be related to the response to antibiotic stress (Sawai *et al.*, 2007). Contrasting results have been reported regarding which are the exact virulence factors responsive to the effect of *covS* and *covR* mutations and their impact on virulence in mouse models. This could be partially explained by the interaction of CovRS with other transcriptional regulators, which creates complex regulation patterns that can vary between different strains (Aziz *et al.*, 2010; Sugareva *et al.*, 2010; Horstmann *et al.*, 2014). Even so, the overall virulence factors reported by most studies that are negatively controlled by CovRS include the hyaluronic acid capsule, SLS, SLO, SIC, NADase, SpyCEP, DNase Sda1, C5a peptidase, and Ska, among others. In contrast, mutations in CovRS result in decreased expression of GRAB, and SpeB (Aziz *et al.*, 2010; Hasegawa *et al.*, 2010). It has been suggested that phosphorylated CovR represses the first genes while non-phosphorylated CovR represses the second group (Trevino *et al.*, 2009; Churchward, 2007).

In 2006, Sumby and co-workers using microarray analysis of nine isolates recovered from human infections observed two different transcriptomes, named by them as pharyngeal transcriptome profile (PTP) and invasive transcriptome profile (ITP). They also observed that ITP isolates could be recovered from mice infected subcutaneously with PTP. Complete genome analysis revealed that the only difference between both isolates was a single 7bp insertion within the *covS* gene (Sumby *et al.*, 2006). This frameshift mutation resulted in the production of a 202-amino acid

truncated CovS compared to the 500 amino acid from the WT. They also observed that isolates with an ITP were significantly more virulent than PTP in a mouse bacteremia model, whereas in a mouse model of soft-tissue infection PTP GAS were more virulent than ITP GAS as assessed by lesion volume. Furthermore, they observed that ITP isolates were better able to survive phagocytosis and killing by human polymorphonuclear leukocytes comparing with PTP isolates. Taken together, these results suggest that the mutation in *covS* was the trigger responsible for the transition from PTP to ITP transcriptome during invasive infection. Consistently, in 2001, Engleberg *et al*. had described the recovery of several mucoid isolates from experimental murine skin infections with spontaneous mutations in *covRS* and that these isolates yielded larger and more necrotic lesions (Engleberg *et al.*, 2001). They also found these mutations among clinical isolates from invasive disease (Engleberg *et al.*, 2001). Previously, it had already been observed that the deletion of *covR* enhanced GAS virulence in mouse infection models (Levin and Wessels, 1998; Heath *et al.*, 1999). Meanwhile, several studies have reported the presence of *covRS* mutations among human clinical isolates, and most of them have suggested that these mutations occur more frequently among isolates recovered from invasive infections (Lin *et al.*, 2014; Hasegawa *et al.*, 2010; Ikebe *et al.*, 2010; Shea *et al.*, 2011). Other studies, mostly based on murine infections models, also suggest that isolates of the highly invasive M1T1 clone have a higher ability to acquire *covRS* mutations than strains of other *emm* types (Mayfield *et al.*, 2014; Maamary *et al.*, 2010)

Currently, the most accepted model to explain the presence of these mutations is that, after GAS adhesion and invasion of epithelial cells, the neutrophils recruited to the site of infection exert a strong selection for those isolates having spontaneous mutations within the *covRS* genes. The enhanced ability to resist neutrophil killing allows the dissemination of GAS *covRS* mutants, potentiating the transition from localized to systemic infection (Cole *et al.*, 2011).

The downregulation of SpeB expression has been considered, by the majority of studies with *covRS* mutants, a critical step for the transition from localized to invasive infection. Even though SpeB is required for the establishment of localized GAS infection, it was suggested that the loss of SpeB activity prevents the degradation of key host proteins and GAS virulence factors, such as Ska and M1 protein and allowing the sequestration of human plasmin protease activity on the cell surface (Cole *et al.*, 2006).

The resulting accumulation of plasmin activity on the GAS surface may contribute to GAS systemic dissemination by degrading host tissue barriers (Cole *et al.*, 2006).

Additionally, it was reported that, while the loss of several individual virulence factors did not prevent GAS *covRS* switching *in vivo*, M1 protein, hyaluronic acid capsule, and DNase Sda1 are indispensable for the switching of phenotype (Cole *et al.*, 2010; Walker *et al.*, 2007). The loss of SpeB also spares Sda1 from degradation improving GAS resistance against neutrophil extracellular traps (Sumby *et al.*, 2005b; Aziz *et al.*, 2004b; Buchanan *et al.*, 2006; Walker *et al.*, 2007). Taken together, capsule, M protein, and Sda1 enhance survival of M1T1 GAS within NETs, which could promote the *in vivo* selection of hypervirulent speB-negative *covRS* mutants (Cole *et al.*, 2010).

7.2. Regulator of proteinase B

Regulator of proteinase B **(**RopB), also named Rgg, controls directly the expression of SpeB by binding to *speB* promoter (Carroll and Musser, 2011; Shelburne *et al.*, 2011). RopB is encoded by the *ropB* gene, which is located 940bp away from *speB*, in the opposite DNA strand (Neely *et al.*, 2003). Additionally to SpeB, RopB significantly influences the expression of ≈25% of the GAS genome during the stationary growth phase, including genes that are involved in virulence, stress responses and interactions with the host (Carroll *et al.*, 2011). Naturally occurring mutations in *ropB* have been shown to impair SpeB production (Hollands *et al.*, 2008). In agreement with the controversial role of SpeB, contrasting results have been reported regarding the RopB role during the infectious process. While some studies reported that *ropB* mutations result in decreased GAS virulence in animal models of infection (Carroll *et al.*, 2011; Olsen *et al.*, 2012), others found that clinical isolates with *ropB* mutations recovered from SSTS have an enhanced lethality in mouse models (Ikebe *et al.*, 2010). With the exception of the members of the *speB* operon, a strain-specific variation, inter and intra-serotypic was reported in the transcriptome of *ropB* mutants, which may explain the contrasting results (Dmitriev *et al.*, 2008).

8. TYPING METHODS

Typing methods are the basis of epidemiological studies with bacterial pathogens. Their use is intended to distinguish isolates, which may allow the identification of outbreaks, the determination of the infection source, monitoring vaccination programs, trace cross-transmission of nosocomial pathogens, and evaluate the effectiveness of control measures (Ranjbar *et al.*, 2014). Additionally, typing methods can be used to control the geographical and temporal dissemination of clones with specific features. Currently, there are a vast set of GAS typing methods, which can be divided into phenotypic and genotypic methods (van Belkum *et al.*, 2007). Several phenotypic methods are being replaced by genotypic methods, and some are no longer used.

In the next section, a description of some typing methods is presented, including those with more relevance nowadays.

8.1. Serotyping

While attempting to understand the basis for protective immunity to GAS infection, serotypic diversity in GAS was first reported by Dr. Rebecca Lancefield, who developed the serotyping scheme based on the variability of the N-terminal region of the surface exposed M protein (Lancefield, 1962; Lancefield, 1928; Cunningham, 2000). More than 80 different serotypes were identified, each requiring a specific serum. However, due to several technical difficulties, such as ambiguous results, the discovery of new M types, the existence of non-typeable isolates and the difficultly and the high cost of preparing high-tittered sera, this method was replaced by the currently widely used DNA sequencing method *emm*-typing (Cunningham, 2000; Beall *et al.*, 1996).

Meanwhile, alternative serotyping methods were developed. One of them, that is no longer used, was based on a lipoproteinase named serum opacity factor (SOF). Its presence can be screened by the increased opacification of mammalian serum. Approximately half of GAS isolates produce SOF and there is a correlation between the M type and the antibodies against SOF, which are also type-specific (Cunningham, 2000; Maxted *et al.*, 1973). Genotypic *sof* sequence types were also defined based on the 5’ end of the *sof* gene and, although the *sof* sequence is a useful predictor of *emm*-type, some exceptions have been noticed (Beall *et al.*, 2000).

A third serological typing scheme, that was widely used for GAS, was the T-typing, which is based in the variability of the trypsin-resistant T antigens described by Lancefield more than 60 years ago (Lancefield and Dole, 1946). In 2005, these antigens were described as the backbone protein of long surface extended pilus structures in several GAS strains (Mora *et al.*, 2005; Falugi *et al.*, 2008). More specifically, immunoblot and agglutination assays revealed, with few exceptions, that the T type is determined by the variant of backbone protein (Mora *et al.*, 2005; Falugi *et al.*, 2008). In fact, it was proposed the use of backbone gene sequence typing (*tee* gene), analogous to *emm* typing, as a potential molecular tool that could substitute the serological T classification of GAS strains (Falugi *et al.*, 2008).

8.2. *emm* typing

This method consists in the PCR amplification of the 5' end of the *emm* gene corresponding to the hypervariable amino-terminal region of the M-protein. Then, the sequence of the PCR product is submitted to the CDC database [https://www2a.cdc.gov/ncidod/biotech/strepblast.asp (27/07/2018)], and an *emm* type number is attributed. For those that were M typeable, a strong correlation was observed between M and *emm* type. Currently, more than 250 *emm*-types have been identified. A unique *emm*-type is defined as having <92% sequence identity over the nucleotide sequence corresponding to the first 30 codons of the M protein (Beall *et al.*, 1996). After the implementation of the sequence-based *emm* typing in 1996, this scheme has become the dominant typing scheme for GAS, and is used in virtually all contemporary epidemiological studies. Even so, epidemiological data from Portugal, showed that *emm*-typing should be complemented with other typing methods, for a more accurate identification of GAS clones (Carrico *et al.*, 2006).

8.3. *emm*-clusters

Recently, a new typing method was described based on the portion of the *emm* gene encoding the entire surface-exposed region of the M proteins. This initial study analyzed 1086 GAS isolates collected from 31 countries representing 175 *emm*-types (Sanderson-Smith *et al.*, 2014). The phylogenetic analysis identified two well-supported clades (X and Y) that were further subdivided into 48 *emm*-clusters. Thirty-two *emm*-clusters contained a single M protein, while the remaining 16 *emm*-clusters possessed

multiple M proteins. Of the pattern E *emm*-types, 98% belong to clade X, whereas 92% of pattern A-C *emm*-types belong to clade Y. The *emm*-types belonging to pattern D were found in 3 different portions of the tree. The same study also assessed the binding activity of representative M proteins from each of the dominant *emm*-clusters to key host proteins known to interact with the M protein. These proteins included plasminogen, IgA, IgG, fibrinogen, albumin, and C4BP. It was found that M proteins assigned to an *emm*-cluster, in addition to a high sequence similarity also share functional properties (Sanderson-Smith *et al.*, 2014). Similarly to *emm*-patterns, the *emm*-clusters can also be directly inferred from *emm*-typing information, given that the authors considered the *emm*-type is predictive of the whole M protein sequence (Table 2) (McMillan *et al.*, 2013a).

Table 2. Distribution of *emm*-types per *emm*-cluster. Reproduced from (Sanderson-Smith *et al.*, 2014).

emm-types	*emm*-clusters
4, 60, 78, 165 (st11014), 176 (st213)	E1
13, 27, 50 (50/62), 66, 68, 76, 90, 92, 96, 104, 106, 110, 117, 166 (st1207), 168 (st1389)	E2
9, 15, 25, 44 (44/61), 49, 58, 79, 82, 87, 103, 107, 113, 118, 144 (stknb1), 180 (st2460), 183 (st2904), 209 (st6735), 219 (st9505), 231 (stNS292)	E3
2, 8, 22, 28, 73, 77, 84, 88, 89, 102, 109, 112, 114, 124, 169 (st1731), 175 (st212), 232 (stN554)	E4
34, 51, 134 (st2105), 137 (st465), 170 (st1815), 174 (st211), 205 (st5282)	E5
11, 42, 48, 59, 63, 65 (65/69), 67, 75, 81, 85, 94, 99, 139 (st7323), 158 (stxh1), 172 (st2037), 177 (st2147), 182 (st2861UK), 191 (st369)	E6
164 (st106M), 185 (st2917), 211 (st7406), 236 (sts104)	Single protein *emm*--cluster clade X
36, 54, 207 (st6030)	D1
32, 71, 100, 115, 213 (st7700)	D2
123, 217 (st809)	D3
33, 41, 43, 52, 53, 56, 56.2 (st3850), 64, 70, 72, 80, 83, 86, 91, 93, 98, 101, 108, 116, 119, 120, 121, 178 (st22), 186 (st2940), 192 (st3757), 194 (st38), 208 (st62), 223 (stD432), 224 (stD631), 225 (stD633), 230 (stNS1033), 242 (st2926)	D4
97, 157 (stn165), 184 (st2911)	D5
46, 142 (st818)	A-C1
30, 197 (st4119)	A-C2
1, 163 (st413) 227 (stil103), 238 (1-2), 239 (1-4)	A-C3
12, 39, 193 (st3765), 228 (stil62), 229 (stmd216)	A-C4
3, 31, 133 (st1692)	A-C5
5, 6, 14, 17, 18, 19, 23, 24, 26, 29, 37, 38 (38/40), 47, 57, 74, 105, 122, 140 (st7395), 179 (st221), 218 (st854), 233 (stNS90), 234 (stpa57)	Single protein *emm* - cluster clade Y
55, 95, 111, 215 (st804), 221 (stCK249), 222 (stCK401)	Single protein *emm* cluster outlier

8.4. SAG profiling

Eight of the eleven SAgs reported to date are located in temperate phages (McMillan *et al.*, 2013b). The exchange of mobile genetic elements (MGE), such as phages, has been considered a crucial contribution to the genetic diversity observed among GAS strains. Moreover, it has been suggested to be involved in the emergence of highly successful virulent clones (Lintges *et al.*, 2010; Maamary *et al.*, 2012; Maripuu *et al.*, 2008). Due to the high polymorphism of some SAGs and the use of different primer sets by different groups, it has been difficult to compare the results from distinct studies (Commons *et al.*, 2014). In 2012, a study with 480 isolates recovered throughout Portugal, was the first to report the use of a defined primer set that enabled the amplification of all known superantigen alleles at that time (Friães *et al.*, 2013b). This study identified 49 different SAg profiles and found that SAg profile diversifies faster than other properties also used for molecular typing, namely *emm* type and MLST. Therefore, SAg profiling can be used to complement GAS characterization in epidemiology studies. Even so, it was observed a strong association between SAg profile and *emm* type, where the SAG profile is a better predictor of *emm* type than the reverse. This link between *emm*-type and SAg profile was previously been noted in several studies comprising isolates from different countries (Commons *et al.*, 2008; Le Hello *et al.*, 2010; Vlaminckx *et al.*, 2003; Schmitz *et al.*, 2003).

8.5. Multilocus Sequence Typing

MLST is a typing method based is the sequencing of 7 house-keeping genes used for several bacterial species. In GAS, these genes encode a glucose kinase (*gki*), a glutamine transport ATP-binding protein (*gtr*), a glutamate racemase (*murI*), a DNA mismatch repair (*mutS*), a transketolase (*recP*), a xanthine phosphoribosyltransferase (*xpt*), and a acetoacetyl-CoA thiolase (*yqil*) (Enright *et al.*, 2001). Each variant of each gene corresponds to a specific allele number, and the combination of all seven to a specific sequence type (ST). The MLST sequence data are posted at https://pubmlst.org/spyogenes/ (27/07/2018) and currently, more than 900 ST are described. It has been considered a better method to infer evolutionary relationships between GAS isolates comparing to SAg profiling, because of the strong selection pressure to which virulence factors are subjected (McMillan *et al.*, 2013b).

9. ANTIMICROBIAL THERAPY AND RESISTANCE MECANISMS

Currently, GAS remains universally susceptible to penicillin with no confirmed reports of naturally occurring resistant isolates anywhere in the world. Therefore, penicillin is the antimicrobial of choice for the treatment of most GAS infections. Penicillin is a β-lactam antibiotic that binds to specific penicillin-binding proteins (PBPs) in the bacterial cell wall inhibiting the assembly of the peptidoglycan chains. Even so, therapeutic failures have been reported, which can lead to recurrent infections and persistent throat carriage. Internalization of GAS in epithelial cells has been suggested as one possible mechanism for these failures, given the poor ability of penicillin to penetrate epithelial cells. Biofilm production by GAS was also suggested to contribute for the escape to antimicrobials. Additionally, the production of β-lactamases by other species during infection may also protect GAS against penicillin treatment (Walker *et al.*, 2014).

9.1. Treatment of pharyngitis

Although GAS pharyngitis is typically a self-limiting disease, treatment is recommended in symptomatic patients with a positive rapid antigen detection test (RADT) or a positive culture (Shulman *et al.*, 2012). The treatment is intended to prevent progression to ARF, but also to thwart the symptoms, to decrease the transmission period and also to prevent the potential suppurative complications. Therefore, according to the clinical practise guidelines, penicillin or amoxicillin are the recommended drugs of choice. These choices are based on their narrow spectrum of activity, rare adverse reactions and modest cost (Shulman *et al.*, 2012). Amoxicillin is frequently used in children, more than penicillin, due to the existence of suspension formulations with an acceptable taste. In penicillin-allergic individuals, the treatment should include a first generation cephalosporin, clindamycin, clarithromycin, or azithromycin. The use of a cephalosporin is only considered in patients that are not anaphylactically sensitive (Shulman *et al.*, 2012). The treatment of GAS carriers is not generally recommended because they are unlikely to spread GAS pharyngitis to their close contacts and are at low risk for developing suppurative or non-suppurative complications (Shulman *et al.*, 2012).

9.2. Treatment of impetigo

The treatment of impetigo depends on the number of lesions, their location and the need to limit the spread of infection to others. Topical agents are preferred in an initial approach. Mupirocin is considered the best topical agent, although resistance has been described. Bacitracin and neomycin are less effective alternative choices (Stevens *et al.*, 2005). Oral treatment is recommended in patients who have numerous lesions or who are not responding to topical agents. Nevertheless, it has to be taken into account that *Staphylococcus aureus* is also a common cause of several skin and soft tissue infections, ranging from impetigo to necrotizing fasciitis. Therefore, empirical choices of antimicrobials must include agents with activity against resistant strains including methicillin-resistant *S. aureus* (MRSA) and erythromycin-resistant GAS. Vancomycin, linezolid, clindamycin, daptomycin, doxycycline, minocycline, and trimethoprim-sulfamethoxazole are possible choices (Stevens *et al.*, 2005).

9.3. Treatment of necrotizing fasciitis

When the suspected diagnosis is necrotizing fasciitis, surgical exploration must be performed and debridement or fasciotomy is often needed. After GAS identification, the therapeutic scheme should be adjusted for intravenously administered penicillin G plus clindamycin for 10 to 14 days (Allen and Moore, 2010; Stevens *et al.*, 2005). In two observational studies, adding clindamycin was demonstrated to be more effective that β-lactam antibiotics alone (Zimbelman *et al.*, 1999; Mulla *et al.*, 2003). In a murine model of myositis, a similar observation was attributed to the "Eagle effect" characterized by the reduction in the expression of PBPs during the stationary growth phase (Stevens *et al.*, 1988). However, the most relevant reason for using clindamycin has been its mechanism of action that by inhibiting protein synthesis may limit the production of several virulence factors, including SAg (Walker *et al.*, 2014). Despite all these benefits, clindamycin should not be used as monotherapy because of the existence of clindamycin-resistant isolates (Allen and Moore, 2010).

9.4. Mechanisms of action and antimicrobial resistance

Macrolides (erythromycin, azithromycin, and clarithromycin) inhibit bacterial protein synthesis by binding to the large subunit of the bacterial ribosome (50S) in the

upper portion of the peptide exit channel, which prevents translation, and also by blocking the assembly of the 50S subunit due to the interaction with its precursors (5S and 23S) (Douthwaite and Champney, 2001). Macrolide resistance in GAS comprises two major mechanisms: target site modification and efflux (Silva-Costa *et al.*, 2015a). The modification of the target site is performed by a methyltransferase encoded by the *erm* genes (erythromycin ribosomal methylase) that methylates a single adenine in the 23S rRNA preventing binding of the antimicrobial to the ribosome. This mechanism also confers resistance to lincosamides (clindamycin) and streptogramins B, and results in the characteristic MLS_B phenotype, which can be expressed constitutively ($cMLS_B$) or in an inducible manner ($iMLS_B$). In $iMLS_B$, the methyltransferase is translated only in the presence of macrolides. Currently, more that 20 classes of *erm* genes are known. However, *erm*(B) and *erm*(TR) are the most common among GAS with the MLS_B phenotype. The second mechanism of efflux relies on membrane-spanning proteins encoded by the *mef* genes, of which *mef*(A) is the most common, which work as efflux pumps of 14- and 15- carbon ring macrolides to the outside the cell (Silva-Costa *et al.*, 2015a). Isolates with this mechanism alone presented the M phenotype, remaining sensitive to 16-member macrolides, lincosamides and streptogramins B.

Contrasting results have been reported regarding the rates and trends of macrolide resistance worldwide. While some countries still report high resistance rates, such as China, Japan, and Korea, an overall decreasing trend has been noted in recent years, mostly in Europe. In some cases, no relation to macrolide consumption was been observed (Silva-Costa *et al.*, 2015a). In agreement, a decreasing rate of erythromycin resistance was reported in Portugal among isolates recovered from pharyngitis from 10% in 2007 to 1% in 2013, consistent with a continuous decline observed since 2000 (Silva-Costa *et al.*, 2015b). Among isolates recovered between 2006-2009 from invasive infections in Portugal, a resistance rate of 8% was reported (Friães *et al.*, 2013a).

Tetracyclines inhibit bacterial protein synthesis by blocking the association of aminoacyl-tRNA with the bacterial ribosome (Chopra and Roberts, 2001). Tetracycline resistance has also been reported among GAS, mostly due to enzymes encoded by *tet(M)* and *tet(O)* genes that confer ribosome protection. They share homology with the elongation factors EF-Tu and EF-G, including their GTPase activity. By competing to bind ribosomes, these proteins can lead to the displacement of tetracycline from the ribosome (Chopra and Roberts, 2001). Another resistance mechanism to tetracycline is

the presence of efflux pumps encoded by *tet*(K) and *tet*(L), although these mechanisms are much more uncommon (Roberts, 2005).

Most of the genes conferring resistance to macrolides and tetracyclines are present in mobile genetic elements. Moreover, *erm* and *mef* genes are often collocated together. A well know linkage between *erm*(B) and *tet*(M) results from the insertion of *ermB* into conjugative transposons of the Tn*916* family, which typically carry *tet*(M) (Brenciani *et al.*, 2007). Numerous other combinations were also reported (Varaldo *et al.*, 2009). Consistently, several studies found higher tetracycline resistance values among macrolide-resistant isolates, including in Portugal (Silva-Costa *et al.*, 2015b). This linkage is the reason why both macrolide and tetracycline resistance are often screened together in several studies, even though tetracycline is not included in the therapeutic choices for the treatment of GAS infections.

Fluoroquinolone resistance has also reported in GAS. While reduced susceptibility to fluoroquinolones is associated to point mutations in the quinolone resistance-determination region (QRDR) of the topoisomerase IV encoding gene *parC*, high-level resistance results from additional point mutations in *gyrA* that encodes a DNA gyrase (Hooper, 2002). These point mutations may be acquired by spontaneous mutations or via horizontal gene transfer (Pletz *et al.*, 2006). In Portugal, a 4.9% of ciprofloxacin nonsusceptibility rate was reported in a sample comprising isolates from pharyngitis, skin and soft tissue and invasive infections and also from carriage (Pires *et al.*, 2010). Moreover, this study also reported an association with *emm*6 isolates. In Spain, was reported an increasing trend in fluoroquinolone nonsusceptibility rates reaching 30.7% in 2007 (Montes *et al.*, 2010).

10. VACCINE CANDIDATES

The history of *S. pyogenes* vaccine development dates back more than 90 years, however, currently there is no safe and effective vaccine to prevent *S. pyogenes* infections (Dale *et al.*, 2016). This delay could be attributed the complexity of *S. pyogenes* infections, including the high number of isolates with different *emm* types, the variability of infections in distinct anatomic sites, and the geographic differences in the prevalence and burden of diseases. Besides all this, the economic hurdle is probably a major problem, considering that 95% of all serious *S. pyogenes* disease occurs in low and middle-income countries (Carapetis *et al.*, 2005).

The natural infection with *S .pyogenes* results in protective immunity, explaining the higher incidence of infections among schoolchildren comparing with the incidence in adulthood. In 1959, Lancefield demonstrated that bactericidal type-specific M protein antibodies persisted in humans for years after infections. However, the antibody response to any given streptococcal infection was completely unpredictable because long persisting antibodies in some instances followed clinically mild infections, whereas severe infections might produce no response, or only one of low titer (Lancefield, 1959). Years later, Fox *et al*., showed that the administration of purified M1 protein preparations subcutaneously to adult volunteers conferred protection against challenge with a virulent M1 strain when inoculated in the pharynx (Fox *et al.*, 1973). Likewise, Poly *et al.*, reach the same conclusion after administration by aerosol spray into the nares and oropharynx (Polly *et al.*, 1975)

The major concern in the development of *S. pyogenes* vaccines is related to the potential ability to trigger nonsuppurative autoimmune sequelae such as ARF, due to the presence of autoimmune epitopes. The studies of Massell *et al*. reported a link between a relatively crude M protein vaccine derived from hot acid extracts of type 3 streptococci and at least two cases of acute rheumatic fever (Massell *et al.*, 1968; Massell *et al.*, 1969). Following studies revealed that intact M proteins not only contain protective epitopes but also contain cross-reactive epitopes with human tissue which occasionally resulted in the production of antibodies cross-reacting with human tissue (Cunningham, 2000). Additional studies have shown that the epitopes contained in the hypervariable, type-specific N-terminus of the M protein, evoke antibodies with the greatest bactericidal activity and that they are less likely to cross-react with host tissues (Dale *et al.*, 2016). These observations led to the current and mostly accepted approach of including only N-terminal M peptides. Taking advantage of recombinant techniques, specific 5´ regions of the *emm* genes were linked together and the development of several vaccines containing distinct amounts of peptides was reported. Furthermore, these vaccines have been shown to evoke opsonic antibodies in animals, without evoking tissue cross-reactive antibodies. Two of these, the hexavalent and the 26-valent, reached clinical trials and were proven to be safe, well-tolerated, and to evoke bactericidal antibodies against the vaccine serotypes (Dale *et al.*, 2016). However, due to the increasing data available, a new 30-valent vaccine was formulated based on epidemiological data from a North American pharyngitis study, the CDC´s ABC surveillance in the US, and the StrepEuro study of invasive *S. pyogenes* strains

(Shulman *et al.*, 2009; Dale *et al.*, 2011; Shulman *et al.*, 2004; O'Loughlin *et al.*, 2007; Luca-Harari *et al.*, 2009). The serotypes included in this vaccine account for 98% of all cases of pharyngitis in the US and Canada, 90% of invasive disease in the US and 78% of invasive in Europe (Dale *et al.*, 2011). In animal studies, the 30-valent vaccine evoked bactericidal antibodies against all vaccine serotypes of GAS (Dale *et al.*, 2011). Curiously, the immunization with 30-valent vaccine also evoked bactericidal antibodies against several non-vaccine serotypes of *S. pyogenes* suggesting that the potential efficacy of this vaccine may extend beyond those serotypes represented in the vaccine (Dale *et al.*, 2011; Dale *et al.*, 2013). Overall, 83 different *emm* types have been tested in bactericidal assays, of which significant killing was observed in 73 (88%) of the isolates. Considering the non-vaccine types tested, 43 out of 53 (81%) were killed, with an average killing of 80% (Dale *et al.*, 2016).

Despite the potential presence of cross-reactive epitopes in the C-terminal region of M protein, to date, three main approaches based in this region were described for the development of a safe vaccine. However, no data from human trials were reported yet. These include the use of the entire C-terminal region of the M6 strain as a recombinant protein, the use of a 12 amino-acid minimal B-cell epitope from the C-repeated region (J8) as a synthetic peptide, and the use of B and T cell epitopes from the C-repeated region from an M5 strain as a synthetic peptide or recombinant protein (Dale *et al.*, 2016).

Meanwhile, other protective antigens, alternatives to M protein-based vaccines, have been suggested, such as GAS carbohydrate, C5a peptidase, SOF, SpeB, SpeC, pili, SpyCEP, GRAB, among others (Dale *et al.*, 2016). However, these approaches are still in early stages of development.

REFERENCES

Abbot, E. L., W. D. Smith, G. P. Siou, C. Chiriboga, R. J. Smith, J. A. Wilson, B. H. Hirst and M. A. Kehoe. 2007. Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. Cell Microbiol. 9: 1822-1833.

Allen, U. and D. Moore. 2010. Invasive group A streptococcal disease: Management and chemoprophylaxis. Can J Infect Dis Med Microbiol. 21: 115-118.

Allhorn, M., A. Olsen and M. Collin. 2008. EndoS from *Streptococcus pyogenes* is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity. BMC Microbiol. 8: 3.

Ashbaugh, C. D., S. Alberti and M. R. Wessels. 1998. Molecular analysis of the capsule gene region of group A Streptococcus: the hasAB genes are sufficient for capsule expression. J Bacteriol. 180: 4955-4959.

Ato, M., T. Ikebe, H. Kawabata, T. Takemori and H. Watanabe. 2008. Incompetence of neutrophils to invasive group A streptococcus is attributed to induction of plural virulence factors by dysfunction of a regulator. PLoS One. 3: e3455.

Aziz, R. K., S. A. Ismail, H. W. Park and M. Kotb. 2004a. Post-proteomic identification of a novel phage-encoded streptodornase, Sda1, in invasive M1T1 *Streptococcus pyogenes*. Mol Microbiol. 54: 184-197.

Aziz, R. K., R. Kansal, B. J. Aronow, W. L. Taylor, S. L. Rowe, M. Kubal, G. S. Chhatwal, M. J. Walker and M. Kotb. 2010. Microevolution of group A streptococci in vivo: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. PLoS One. 5: e9798.

Aziz, R. K., M. J. Pabst, A. Jeng, R. Kansal, D. E. Low, V. Nizet and M. Kotb. 2004b. Invasive M1T1 group A Streptococcus undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. Mol Microbiol. 51: 123-134.

Baird, R. W., M. S. Bronze, W. Kraus, H. R. Hill, L. G. Veasey and J. B. Dale. 1991. Epitopes of group A streptococcal M protein shared with antigens of articular cartilage and synovium. J Immunol. 146: 3132-3137.

Beall, B., R. Facklam and T. Thompson. 1996. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. J Clin Microbiol. 34: 953-958.

Beall, B., G. Gherardi, M. Lovgren, R. R. Facklam, B. A. Forwick and G. J. Tyrrell. 2000. *emm* and *sof* gene sequence variation in relation to serological typing of opacity-factor-positive group A streptococci. Microbiology. 146 (Pt 5): 1195-1209.

Becherelli, M., A. G. Manetti, S. Buccato, E. Viciani, L. Ciucchi, G. Mollica, G. Grandi and I. Margarit. 2012. The ancillary protein 1 of *Streptococcus pyogenes* FCT-1 pili mediates cell adhesion and biofilm formation through heterophilic as well as homophilic interactions. Mol Microbiol. 83: 1035-1047.

Ben Nasr, A., H. Herwald, U. Sjobring, T. Renne, W. Muller-Esterl and L. Bjorck. 1997. Absorption of kininogen from human plasma by *Streptococcus pyogenes* is followed by the release of bradykinin. Biochem J. 326 (Pt 3): 657-660.

Berge, A. and L. Bjorck. 1995. Streptococcal cysteine proteinase releases biologically active fragments of streptococcal surface proteins. J Biol Chem. 270: 9862-9867.

Berggard, K., E. Johnsson, E. Morfeldt, J. Persson, M. Stalhammar-Carlemalm and G. Lindahl. 2001. Binding of human C4BP to the hypervariable region of M protein: a molecular mechanism of phagocytosis resistance in *Streptococcus pyogenes*. Mol Microbiol. 42: 539-551.

Bessen, D., K. F. Jones and V. A. Fischetti. 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. J Exp Med. 169: 269-283.

Bessen, D. E., T. R. Fiorentino and S. K. Hollingshead. 1997. Molecular markers for throat and skin isolates of group A streptococci. Adv Exp Med Biol. 418: 537-543.

Bessen, D. E. and V. A. Fischetti. 1990. Differentiation between two biologically distinct classes of group A streptococci by limited substitutions of amino acids within the shared region of M protein-like molecules. J Exp Med. 172: 1757-1764.

Bessen, D. E. and A. Kalia. 2002. Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in Streptococcus pyogenes. Infect Immun. 70: 1159-1167.

Bessen, D. E., A. Manoharan, F. Luo, J. E. Wertz and D. A. Robinson. 2005. Evolution of transcription regulatory genes is linked to niche specialization in the bacterial pathogen *Streptococcus pyogenes*. J Bacteriol. 187: 4163-4172.

Bessen, D. E., W. Michael McShan, S. V. Nguyen, A. Shetty, S. Agrawal and H. Tettelin. 2014. Molecular epidemiology and genomics of group A Streptococcus. Infect Genet Evol.

Bessen, D. E., C. M. Sotir, T. L. Readdy and S. K. Hollingshead. 1996. Genetic correlates of throat and skin isolates of group A streptococci. J Infect Dis. 173: 896-900.

Betschel, S. D., S. M. Borgia, N. L. Barg, D. E. Low and J. C. De Azavedo. 1998. Reduced virulence of group A streptococcal Tn916 mutants that do not produce streptolysin S. Infect Immun. 66: 1671-1679.

Bisno, A. L. 1979. Alternate complement pathway activation by group A streptococci: role of M-protein. Infect Immun. 26: 1172-1176.

Bisno, A. L. 2001. Acute pharyngitis. N Engl J Med. 344: 205-211.

Bisno, A. L., M. O. Brito and C. M. Collins. 2003. Molecular basis of group A streptococcal virulence. Lancet Infect Dis. 3: 191-200.

Bisno, A. L. and D. L. Stevens. 1996. Streptococcal infections of skin and soft tissues. N Engl J Med. 334: 240-245.

Bjorck, L., P. Akesson, M. Bohus, J. Trojnar, M. Abrahamson, I. Olafsson and A. Grubb. 1989. Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. Nature. 337: 385-386.

Brenciani, A., A. Bacciaglia, M. Vecchi, L. A. Vitali, P. E. Varaldo and E. Giovanetti. 2007. Genetic elements carrying erm(B) in *Streptococcus pyogenes* and association with tet(M) tetracycline resistance gene. Antimicrob Agents Chemother. 51: 1209-1216.

Bricker, A. L., V. J. Carey and M. R. Wessels. 2005. Role of NADase in virulence in experimental invasive group A streptococcal infection. Infect Immun. 73: 6562-6566.

Bricker, A. L., C. Cywes, C. D. Ashbaugh and M. R. Wessels. 2002. NAD+-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group A streptococci. Mol Microbiol. 44: 257-269.

Brinkmann, V., U. Reichard, C. Goosmann, B. Fauler, Y. Uhlemann, D. S. Weiss, Y. Weinrauch and A. Zychlinsky. 2004. Neutrophil extracellular traps kill bacteria. Science. 303: 1532-1535.

Broudy, T. B., V. Pancholi and V. A. Fischetti. 2001. Induction of lysogenic bacteriophage and phage-associated toxin from group a streptococci during coculture with human pharyngeal cells. Infect Immun. 69: 1440-1443.

Buchanan, J. T., A. J. Simpson, R. K. Aziz, G. Y. Liu, S. A. Kristian, M. Kotb, J. Feramisco and V. Nizet. 2006. DNase expression allows the pathogen group A Streptococcus to escape killing in neutrophil extracellular traps. Curr Biol. 16: 396-400.

Burns, E. H., Jr., A. M. Marciel and J. M. Musser. 1996. Activation of a 66-kilodalton human endothelial cell matrix metalloprotease by *Streptococcus pyogenes* extracellular cysteine protease. Infect Immun. 64: 4744-4750.

Carapetis, J. R., A. C. Steer, E. K. Mulholland and M. Weber. 2005. The global burden of group A streptococcal diseases. Lancet Infect Dis. 5: 685-694.

Carlsson, F., K. Berggard, M. Stalhammar-Carlemalm and G. Lindahl. 2003. Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. J Exp Med. 198: 1057-1068.

Carlsson, F., C. Sandin and G. Lindahl. 2005. Human fibrinogen bound to *Streptococcus pyogenes* M protein inhibits complement deposition via the classical pathway. Mol Microbiol. 56: 28-39.

Carrico, J. A., C. Silva-Costa, J. Melo-Cristino, F. R. Pinto, H. de Lencastre, J. S. Almeida and M. Ramirez. 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol. 44: 2524-2532.

Carroll, R. K. and J. M. Musser. 2011. From transcription to activation: how group A streptococcus, the flesh-eating pathogen, regulates SpeB cysteine protease production. Mol Microbiol. 81: 588-601.

Carroll, R. K., S. A. Shelburne, 3rd, R. J. Olsen, B. Suber, P. Sahasrabhojane, M. Kumaraswami, S. B. Beres, P. R. Shea, A. R. Flores and J. M. Musser. 2011. Naturally occurring single amino acid replacements in a regulatory protein alter streptococcal gene expression and virulence in mice. J Clin Invest. 121: 1956-1968.

Chartier, C. and E. Grosshans. 1990. Erysipelas. Int J Dermatol. 29: 459-467.

Chatellier, S., N. Ihendyane, R. G. Kansal, F. Khambaty, H. Basma, A. Norrby-Teglund, D. E. Low, A. McGeer and M. Kotb. 2000. Genetic relatedness and superantigen expression in group A streptococcus serotype M1 isolates from patients with severe and nonsevere invasive diseases. Infect Immun. 68: 3523-3534.

Chaussee, M. S., R. L. Cole and J. P. van Putten. 2000. Streptococcal erythrogenic toxin B abrogates fibronectin-dependent internalization of *Streptococcus pyogenes* by cultured mammalian cells. Infect Immun. 68: 3226-3232.

Chaussee, M. S., J. Liu, D. L. Stevens and J. J. Ferretti. 1996. Genetic and phenotypic diversity among isolates of Streptococcus pyogenes from invasive infections. J Infect Dis. 173: 901-908.

Chen, C. Y., S. C. Luo, C. F. Kuo, Y. S. Lin, J. J. Wu, M. T. Lin, C. C. Liu, W. Y. Jeng and W. J. Chuang. 2003. Maturation processing and characterization of streptopain. J Biol Chem. 278: 17336-17343.

Chiang-Ni, C. and J. J. Wu. 2008. Effects of streptococcal pyrogenic exotoxin B on pathogenesis of *Streptococcus pyogenes*. J Formos Med Assoc. 107: 677-685.

Choby, B. A. 2009. Diagnosis and treatment of streptococcal pharyngitis. Am Fam Physician. 79: 383-390.

Chopra, I. and M. Roberts. 2001. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev. 65: 232-260 ; second page, table of contents.

Churchward, G. 2007. The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. Mol Microbiol. 64: 34-41.

Cleary, P. P., E. L. Kaplan, J. P. Handley, A. Wlazlo, M. H. Kim, A. R. Hauser and P. M. Schlievert. 1992. Clonal basis for resurgence of serious *Streptococcus pyogenes* disease in the 1980s. Lancet. 339: 518-521.

Cole, C. and J. Gazewood. 2007. Diagnosis and treatment of impetigo. Am Fam Physician. 75: 859-864.

Cole, J. N., T. C. Barnett, V. Nizet and M. J. Walker. 2011. Molecular insight into invasive group A streptococcal disease. Nat Rev Microbiol. 9: 724-736.

Cole, J. N., J. D. McArthur, F. C. McKay, M. L. Sanderson-Smith, A. J. Cork, M. Ranson, M. Rohde, A. Itzek, H. Sun, D. Ginsburg, M. Kotb, V. Nizet, G. S. Chhatwal and M. J. Walker. 2006. Trigger for group A streptococcal M1T1 invasive disease. FASEB J. 20: 1745-1747.

Cole, J. N., M. A. Pence, M. von Kockritz-Blickwede, A. Hollands, R. L. Gallo, M. J. Walker and V. Nizet. 2010. M protein and hyaluronic acid capsule are essential for in vivo selection of covRS mutations characteristic of invasive serotype M1T1 group A *Streptococcus*. MBio. 1.

Collin, M. and A. Olsen. 2001. Effect of SpeB and EndoS from *Streptococcus pyogenes* on human immunoglobulins. Infect Immun. 69: 7187-7189.

Collin, M., M. D. Svensson, A. G. Sjoholm, J. C. Jensenius, U. Sjobring and A. Olsen. 2002. EndoS and SpeB from *Streptococcus pyogenes* inhibit immunoglobulin-mediated opsonophagocytosis. Infect Immun. 70: 6646-6651.

Commons, R., S. Rogers, T. Gooding, M. Danchin, J. Carapetis, R. Robins-Browne and N. Curtis. 2008. Superantigen genes in group A streptococcal isolates and their relationship with emm types. J Med Microbiol. 57: 1238-1246.

Commons, R. J., P. R. Smeesters, T. Proft, J. D. Fraser, R. Robins-Browne and N. Curtis. 2014. Streptococcal superantigens: categorization and clinical associations. Trends Mol Med. 20: 48-62.

Courtney, H. S., D. L. Hasty and J. B. Dale. 2006. Anti-phagocytic mechanisms of *Streptococcus pyogenes*: binding of fibrinogen to M-related protein. Mol Microbiol. 59: 936-947.

Courtney, H. S., S. Liu, J. B. Dale and D. L. Hasty. 1997. Conversion of M serotype 24 of *Streptococcus pyogenes* to M serotypes 5 and 18: effect on resistance to phagocytosis and adhesion to host cells. Infect Immun. 65: 2472-2474.

Courtney, H. S. and H. J. Pownall. 2010. The structure and function of serum opacity factor: a unique streptococcal virulence determinant that targets high-density lipoproteins. J Biomed Biotechnol. 2010: 956071.

Cue, D., H. Lam and P. P. Cleary. 2001. Genetic dissection of the *Streptococcus pyogenes* M1 protein: regions involved in fibronectin binding and intracellular invasion. Microb Pathog. 31: 231-242.

Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 13: 470-511.

Cunningham, M. W. 2008. Pathogenesis of group A streptococcal infections and their sequelae. Adv Exp Med Biol. 609: 29-42.

Cywes, C., I. Stamenkovic and M. R. Wessels. 2000. CD44 as a receptor for colonization of the pharynx by group A *Streptococcus*. J Clin Invest. 106: 995-1002.

Dale, J. B., M. R. Batzloff, P. P. Cleary, H. S. Courtney, M. F. Good, G. Grandi, S. Halperin, I. Y. Margarit, S. McNeil, M. Pandey, P. R. Smeesters and A. C. Steer. 2016. Current Approaches to Group A Streptococcal Vaccine Development.

Dale, J. B., T. A. Penfound, E. Y. Chiang and W. J. Walton. 2011. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine. 29: 8175-8178.

Dale, J. B., T. A. Penfound, B. Tamboura, S. O. Sow, J. P. Nataro, M. Tapia and K. L. Kotloff. 2013. Potential coverage of a multivalent M protein-based group A streptococcal vaccine. Vaccine. 31: 1576-1581.

Dale, J. B., R. G. Washburn, M. B. Marques and M. R. Wessels. 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. Infect Immun. 64: 1495-1501.

Dalton, T. L. and J. R. Scott. 2004. CovS inactivates CovR and is required for growth under conditions of general stress in *Streptococcus pyogenes*. J Bacteriol. 186: 3928-3937.

de Malmanche, S. A. and D. R. Martin. 1994. Protective immunity to the group A *Streptococcus* may be only strain specific. Med Microbiol Immunol. 183: 299-306.

Descheemaeker, P., F. Van Loock, M. Hauchecorne, P. Vandamme and H. Goossens. 2000. Molecular characterisation of group A streptococci from invasive and non-invasive disease episodes in Belgium during 1993-1994. J Med Microbiol. 49: 467-471.

Dmitriev, A. V., E. J. McDowell and M. S. Chaussee. 2008. Inter- and intraserotypic variation in the *Streptococcus pyogenes* Rgg regulon. FEMS Microbiol Lett. 284: 43-51.

Douthwaite, S. and W. S. Champney. 2001. Structures of ketolides and macrolides determine their mode of interaction with the ribosomal target site. J Antimicrob Chemother. 48 Suppl T1: 1-8.

Efstratiou, A. and T. Lamagni. 2016. Epidemiology of *Streptococcus pyogenes*.

Engleberg, N. C., A. Heath, A. Miller, C. Rivera and V. J. DiRita. 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis. 183: 1043-1054.

Engleberg, N. C., A. Heath, K. Vardaman and V. J. DiRita. 2004. Contribution of CsrR-regulated virulence factors to the progress and outcome of murine skin infections by *Streptococcus pyogenes*. Infect Immun. 72: 623-628.

Enright, M. C., B. G. Spratt, A. Kalia, J. H. Cross and D. E. Bessen. 2001. Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between emm type and clone. Infect Immun. 69: 2416-2427.

Eriksson, A. and M. Norgren. 2003. Cleavage of antigen-bound immunoglobulin G by SpeB contributes to streptococcal persistence in opsonizing blood. Infect Immun. 71: 211-217.

Eriksson, B. K., J. Andersson, S. E. Holm and M. Norgren. 1999. Invasive group A streptococcal infections: T1M1 isolates expressing pyrogenic exotoxins A and B

in combination with selective lack of toxin-neutralizing antibodies are associated with increased risk of streptococcal toxic shock syndrome. J Infect Dis. 180: 410-418.

Facklam, R. 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin Microbiol Rev. 15: 613-630.

Facklam, R. F., D. R. Martin, M. Lovgren, D. R. Johnson, A. Efstratiou, T. A. Thompson, S. Gowan, P. Kriz, G. J. Tyrrell, E. Kaplan and B. Beall. 2002. Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: emm103 to emm124. Clin Infect Dis. 34: 28-38.

Falugi, F., C. Zingaretti, V. Pinto, M. Mariani, L. Amodeo, A. G. Manetti, S. Capo, J. M. Musser, G. Orefici, I. Margarit, J. L. Telford, G. Grandi and M. Mora. 2008. Sequence variation in group A *Streptococcus* pili and association of pilus backbone types with lancefield T serotypes. J Infect Dis. 198: 1834-1841.

Ferretti, J. and W. Kohler. 2016. History of Streptococcal Research.

Fiorentino, T. R., B. Beall, P. Mshar and D. E. Bessen. 1997. A genetic-based evaluation of the principal tissue reservoir for group A streptococci isolated from normally sterile sites. J Infect Dis. 176: 177-182.

Fischetti, V. A. 1989. Streptococcal M protein: molecular design and biological behavior. Clin Microbiol Rev. 2: 285-314.

Fischetti, V. A., V. Pancholi and O. Schneewind. 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. Mol Microbiol. 4: 1603-1605.

Flores, A. R., B. E. Jewell, N. Fittipaldi, S. B. Beres and J. M. Musser. 2012. Human disease isolates of serotype m4 and m22 group a streptococcus lack genes required for hyaluronic Acid capsule biosynthesis. MBio. 3: e00413-00412.

Flores, A. R., B. E. Jewell, R. J. Olsen, S. A. Shelburne, 3rd, N. Fittipaldi, S. B. Beres and J. M. Musser. 2014. Asymptomatic carriage of group A *Streptococcus* is associated with elimination of capsule production. Infect Immun. 82: 3958-3967.

Fox, E. N., R. H. Waldman, M. K. Wittner, A. A. Mauceri and A. Dorfman. 1973. Protective study with a group A streptococcal M protein vaccine. Infectivity challenge of human volunteers. J Clin Invest. 52: 1885-1892.

Friães, A., J. P. Lopes, J. Melo-Cristino and M. Ramirez. 2013a. Changes in *Streptococcus pyogenes* causing invasive disease in Portugal: evidence for superantigen gene loss and acquisition. Int J Med Microbiol. 303: 505-513.

Friães, A., M. P. Machado, C. Pato, J. Carrico, J. Melo-Cristino and M. Ramirez. 2015. Emergence of the Same Successful Clade among Distinct Populations of emm89 *Streptococcus pyogenes* in Multiple Geographic Regions. MBio. 6: e01780-01715.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2012. Group A streptococci clones associated with invasive infections and pharyngitis in Portugal present differences in emm types, superantigen gene content and antimicrobial resistance. BMC Microbiol. 12: 280.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2013b. Superantigen gene complement of *Streptococcus pyogenes*--relationship with other typing methods and short-term stability. Eur J Clin Microbiol Infect Dis. 32: 115-125.

Froehlich, B. J., C. Bates and J. R. Scott. 2009. *Streptococcus pyogenes* CovRS mediates growth in iron starvation and in the presence of the human cationic antimicrobial peptide LL-37. J Bacteriol. 191: 673-677.

Gerlach, D., K. H. Schmidt and B. Fleischer. 2001. Basic streptococcal superantigens (SPEX/SMEZ or SPEC) are responsible for the mitogenic activity of the so-called mitogenic factor (MF). FEMS Immunol Med Microbiol. 30: 209-216.

Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott and J. M. Musser. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. Proc Natl Acad Sci U S A. 99: 13855-13860.

Graham, M. R., K. Virtaneva, S. F. Porcella, W. T. Barry, B. B. Gowen, C. R. Johnson, F. A. Wright and J. M. Musser. 2005. Group A *Streptococcus* transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. Am J Pathol. 166: 455-465.

Gryllos, I., J. C. Levin and M. R. Wessels. 2003. The CsrR/CsrS two-component system of group A Streptococcus responds to environmental Mg2+. Proc Natl Acad Sci U S A. 100: 4227-4232.

Gubbe, K., R. Misselwitz, K. Welfle, W. Reichardt, K. H. Schmidt and H. Welfle. 1997. C repeats of the streptococcal M1 protein achieve the human serum albumin binding ability by flanking regions which stabilize the coiled-coil conformation. Biochemistry. 36: 8107-8113.

Hasegawa, T., A. Okamoto, T. Kamimura, I. Tatsuno, S. N. Hashikawa, M. Yabutani, M. Matsumoto, K. Yamada, M. Isaka, M. Minami and M. Ohta. 2010. Detection of invasive protein profile of *Streptococcus pyogenes* M1 isolates from pharyngitis patients. APMIS. 118: 167-178.

Hasty, D. L., I. Ofek, H. S. Courtney and R. J. Doyle. 1992. Multiple adhesins of streptococci. Infect Immun. 60: 2147-2152.

Hauser, A. R. and P. M. Schlievert. 1990. Nucleotide sequence of the streptococcal pyrogenic exotoxin type B gene and relationship between the toxin and the streptococcal proteinase precursor. J Bacteriol. 172: 4536-4542.

Hauser, A. R., D. L. Stevens, E. L. Kaplan and P. M. Schlievert. 1991. Molecular analysis of pyrogenic exotoxins from *Streptococcus pyogenes* isolates associated with toxic shock-like syndrome. J Clin Microbiol. 29: 1562-1567.

Heath, A., V. J. DiRita, N. L. Barg and N. C. Engleberg. 1999. A two-component regulatory system, CsrR-CsrS, represses expression of three *Streptococcus pyogenes* virulence factors, hyaluronic acid capsule, streptolysin S, and pyrogenic exotoxin B. Infect Immun. 67: 5298-5305.

Herwald, H., M. Collin, W. Muller-Esterl and L. Bjorck. 1996. Streptococcal cysteine proteinase releases kinins: a virulence mechanism. J Exp Med. 184: 665-673.

Herwald, H., H. Cramer, M. Morgelin, W. Russell, U. Sollenberg, A. Norrby-Teglund, H. Flodgaard, L. Lindbom and L. Bjorck. 2004. M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. Cell. 116: 367-379.

Hollands, A., R. K. Aziz, R. Kansal, M. Kotb, V. Nizet and M. J. Walker. 2008. A naturally occurring mutation in ropB suppresses SpeB expression and reduces M1T1 group A streptococcal systemic virulence. PLoS One. 3: e4102.

Hollands, A., D. Gonzalez, E. Leire, C. Donald, R. L. Gallo, M. Sanderson-Smith, P. C. Dorrestein and V. Nizet. 2012. A bacterial pathogen co-opts host plasmin to resist killing by cathelicidin antimicrobial peptides. J Biol Chem. 287: 40891-40897.

Hollands, A., M. A. Pence, A. M. Timmer, S. R. Osvath, L. Turnbull, C. B. Whitchurch, M. J. Walker and V. Nizet. 2010. Genetic switch to hypervirulence reduces colonization phenotypes of the globally disseminated group A *Streptococcus* M1T1 clone. J Infect Dis. 202: 11-19.

Hollingshead, S. K., J. Arnold, T. L. Readdy and D. E. Bessen. 1994. Molecular evolution of a multigene family in group A streptococci. Mol Biol Evol. 11: 208-219.

Hollingshead, S. K., T. L. Readdy, D. L. Yung and D. E. Bessen. 1993. Structural heterogeneity of the emm gene cluster in group A streptococci. Mol Microbiol. 8: 707-717.

Holub, M., M. Helcl, O. Beran, J. Hnykova and J. Prihodova. 2004. Kinetics of immune parameters in a patient with sepsis caused by *Streptococcus pyogenes* treated with activated protein C. Scand J Infect Dis. 36: 485-488.

Hondorp, E. R., S. C. Hou, L. L. Hause, K. Gera, C. E. Lee and K. S. McIver. 2013. PTS phosphorylation of Mga modulates regulon expression and virulence in the group A *Streptococcus*. Mol Microbiol. 88: 1176-1193.

Hooper, D. C. 2002. Fluoroquinolone resistance among Gram-positive cocci. Lancet Infect Dis. 2: 530-538.

Horstmann, N., M. Saldana, P. Sahasrabhojane, H. Yao, X. Su, E. Thompson, A. Koller and S. A. Shelburne, 3rd. 2014. Dual-site phosphorylation of the control of virulence regulator impacts group a streptococcal global gene expression and pathogenesis. PLoS Pathog. 10: e1004088.

Horstmann, R. D., H. J. Sievertsen, M. Leippe and V. A. Fischetti. 1992. Role of fibrinogen in complement inhibition by streptococcal M protein. Infect Immun. 60: 5036-5041.

Hsueh, P. R., J. J. Wu, P. J. Tsai, J. W. Liu, Y. C. Chuang and K. T. Luh. 1998. Invasive group A streptococcal disease in Taiwan is not associated with the presence of streptococcal pyrogenic exotoxin genes. Clin Infect Dis. 26: 584-589.

Iglesias-Gamarra, A., E. A. Mendez, M. L. Cuellar, J. H. Ponce de Leon, C. Jimenez, C. Canas, J. Restrepo, M. Pena, R. Valle and L. R. Espinoza. 2001. Poststreptococcal reactive arthritis in adults: long-term follow-up. Am J Med Sci. 321: 173-177.

Ikebe, T., M. Ato, T. Matsumura, H. Hasegawa, T. Sata, K. Kobayashi and H. Watanabe. 2010. Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. PLoS Pathog. 6: e1000832.

Ikebe, T., A. Wada, Y. Inagaki, K. Sugama, R. Suzuki, D. Tanaka, A. Tamaru, Y. Fujinaga, Y. Abe, Y. Shimizu and H. Watanabe. 2002. Dissemination of the phage-associated novel superantigen gene speL in recent invasive and noninvasive *Streptococcus pyogenes* M3/T3 isolates in Japan. Infect Immun. 70: 3227-3233.

Jansen, T. L., M. Janssen, A. J. de Jong and M. E. Jeurissen. 1999. Post-streptococcal reactive arthritis: a clinical and serological description, revealing its distinction from acute rheumatic fever. J Intern Med. 245: 261-267.

Ji, Y., B. Carlson, A. Kondagunta and P. P. Cleary. 1997. Intranasal immunization with C5a peptidase prevents nasopharyngeal colonization of mice by the group A *Streptococcus*. Infect Immun. 65: 2080-2087.

Ji, Y., L. McLandsborough, A. Kondagunta and P. P. Cleary. 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. Infect Immun. 64: 503-510.

Johansson, L., P. Thulin, D. E. Low and A. Norrby-Teglund. 2010. Getting under the skin: the immunopathogenesis of *Streptococcus pyogenes* deep tissue infections. Clin Infect Dis. 51: 58-65.

Johansson, L., P. Thulin, P. Sendi, E. Hertzen, A. Linder, P. Akesson, D. E. Low, B. Agerberth and A. Norrby-Teglund. 2008. Cathelicidin LL-37 in severe *Streptococcus pyogenes* soft tissue infections in humans. Infect Immun. 76: 3399-3404.

Johnsson, E., K. Berggard, H. Kotarsky, J. Hellwage, P. F. Zipfel, U. Sjobring and G. Lindahl. 1998. Role of the hypervariable region in streptococcal M proteins: binding of a human complement inhibitor. J Immunol. 161: 4894-4901.

Johnsson, E., A. Thern, B. Dahlback, L. O. Heden, M. Wikstrom and G. Lindahl. 1996. A highly variable region in members of the streptococcal M protein family binds the human complement regulator C4BP. J Immunol. 157: 3021-3029.

Johnston, K. H. and J. B. Zabriskie. 1986. Purification and partial characterization of the nephritis strain-associated protein from *Streptococcus pyogenes*, group A. J Exp Med. 163: 697-712.

Kansal, R. G., A. McGeer, D. E. Low, A. Norrby-Teglund and M. Kotb. 2000. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. Infect Immun. 68: 6362-6369.

Kansal, R. G., V. Nizet, A. Jeng, W. J. Chuang and M. Kotb. 2003. Selective modulation of superantigen-induced responses by streptococcal cysteine protease. J Infect Dis. 187: 398-407.

Kapur, V., J. T. Maffei, R. S. Greer, L. L. Li, G. J. Adams and J. M. Musser. 1994. Vaccination with streptococcal extracellular cysteine protease (interleukin-1 beta convertase) protects mice against challenge with heterologous group A streptococci. Microb Pathog. 16: 443-450.

Kapur, V., S. Topouzis, M. W. Majesky, L. L. Li, M. R. Hamrick, R. J. Hamill, J. M. Patti and J. M. Musser. 1993. A conserved *Streptococcus pyogenes* extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. Microb Pathog. 15: 327-346.

Kazmi, S. U., R. Kansal, R. K. Aziz, M. Hooshdaran, A. Norrby-Teglund, D. E. Low, A. B. Halim and M. Kotb. 2001. Reciprocal, temporal expression of SpeA and SpeB by invasive M1T1 group a streptococcal isolates in vivo. Infect Immun. 69: 4988-4995.

Kirvan, C. A., S. E. Swedo, D. Kurahara and M. W. Cunningham. 2006. Streptococcal mimicry and antibody-mediated cell signaling in the pathogenesis of Sydenham's chorea. Autoimmunity. 39: 21-29.

Kittang, B. R., S. Skrede, N. Langeland, C. G. Haanshuus and H. Mylvaganam. 2011. 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. 30: 423-433.

Koller, T., A. G. Manetti, B. Kreikemeyer, C. Lembke, I. Margarit, G. Grandi and A. Podbielski. 2010. Typing of the pilus-protein-encoding FCT region and biofilm formation as novel parameters in epidemiological investigations of *Streptococcus pyogenes* isolates from various infection sites. J Med Microbiol. 59: 442-452.

Kotarsky, H., A. Thern, G. Lindahl and U. Sjobring. 2000. Strain-specific restriction of the antiphagocytic property of group A streptococcal M proteins. Infect Immun. 68: 107-112.

Kotb, M., A. Norrby-Teglund, A. McGeer, H. El-Sherbini, M. T. Dorak, A. Khurshid, K. Green, J. Peeples, J. Wade, G. Thomson, B. Schwartz and D. E. Low. 2002. An immunogenetic and molecular basis for differences in outcomes of invasive group A streptococcal infections. Nat Med. 8: 1398-1404.

Kratovac, Z., A. Manoharan, F. Luo, S. Lizano and D. E. Bessen. 2007. Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. J Bacteriol. 189: 1299-1310.

Kreikemeyer, B., M. Nakata, T. Koller, H. Hildisch, V. Kourakos, K. Standar, S. Kawabata, M. O. Glocker and A. Podbielski. 2007. The *Streptococcus pyogenes* serotype M49 Nra-Ralp3 transcriptional regulatory network and its control of virulence factor expression from the novel eno ralp3 epf sagA pathogenicity region. Infect Immun. 75: 5698-5710.

Lamagni, T. L., J. Darenberg, B. Luca-Harari, T. Siljander, A. Efstratiou, B. Henriques-Normark, J. Vuopio-Varkila, A. Bouvet, R. Creti, K. Ekelund, M. Koliou, R. R. Reinert, A. Stathi, L. Strakova, V. Ungureanu, C. Schalen and A. Jasir. 2008. Epidemiology of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 46: 2359-2367.

Lancefield, R. C. 1928. The Antigenic Complex of Streptococcus Haemolyticus : I. Demonstration of a Type-Specific Substance in Extracts of *Streptococcus Haemolyticus*. J Exp Med. 47: 91-103.

Lancefield, R. C. 1933. A Serological Differentiation of Human and Other Groups of Hemolytic Streptococci. J Exp Med. 57: 571-595.

Lancefield, R. C. 1959. Persistence of type-specific antibodies in man following infection with group A streptococci. J Exp Med. 110: 271-292.

Lancefield, R. C. 1962. Current knowledge of type-specific M antigens of group A streptococci. J Immunol. 89: 307-313.

Lancefield, R. C. and V. P. Dole. 1946. The Properties of T Antigens Extracted from Group a Hemolytic Streptococci. J Exp Med. 84: 449-471.

Lauth, X., M. von Kockritz-Blickwede, C. W. McNamara, S. Myskowski, A. S. Zinkernagel, B. Beall, P. Ghosh, R. L. Gallo and V. Nizet. 2009. M1 protein allows Group A streptococcal survival in phagocyte extracellular traps through cathelicidin inhibition. J Innate Immun. 1: 202-214.

Le Hello, S., A. Doloy, F. Baumann, N. Roques, P. Coudene, B. Rouchon, F. Lacassin and A. Bouvet. 2010. Clinical and microbial characteristics of invasive *Streptococcus pyogenes* disease in New Caledonia, a region in Oceania with a high incidence of acute rheumatic fever. J Clin Microbiol. 48: 526-530.

Lee, J. L., S. M. Naguwa, G. S. Cheema and M. E. Gershwin. 2009. Acute rheumatic fever and its consequences: a persistent threat to developing nations in the 21st century. Autoimmun Rev. 9: 117-123.

Levin, J. C. and M. R. Wessels. 1998. Identification of csrR/csrS, a genetic locus that regulates hyaluronic acid capsule synthesis in group A *Streptococcus*. Mol Microbiol. 30: 209-219.

Lin, J. N., L. L. Chang, C. H. Lai, H. H. Lin and Y. H. Chen. 2013. Group A streptococcal necrotizing fasciitis in the emergency department. J Emerg Med. 45: 781-788.

Lin, J. N., L. L. Chang, C. H. Lai, H. H. Lin and Y. H. Chen. 2014. Association between polymorphisms in the csrRS two-component regulatory system and invasive group A streptococcal infection. Eur J Clin Microbiol Infect Dis. 33: 735-743.

Lintges, M., M. van der Linden, R. D. Hilgers, S. Arlt, A. Al-Lahham, R. R. Reinert, S. Plucken and L. Rink. 2010. Superantigen genes are more important than the emm type for the invasiveness of group A *Streptococcus* infection. J Infect Dis. 202: 20-28.

Liu, T. Y. and S. D. Elliott. 1965. Streptococcal Proteinase: The Zymogen to Enzyme Transfromation. J Biol Chem. 240: 1138-1142.

Loughman, J. A. and M. Caparon. 2006. Regulation of SpeB in *Streptococcus pyogenes* by pH and NaCl: a model for in vivo gene expression. J Bacteriol. 188: 399-408.

Luca-Harari, B., J. Darenberg, S. Neal, T. Siljander, L. Strakova, A. Tanna, R. Creti, K. Ekelund, M. Koliou, P. T. Tassios, M. van der Linden, M. Straut, J. Vuopio-Varkila, A. Bouvet, A. Efstratiou, C. Schalen, B. Henriques-Normark and A. Jasir. 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 47: 1155-1165.

Ly, D., J. M. Taylor, J. A. Tsatsaronis, M. M. Monteleone, A. S. Skora, C. A. Donald, T. Maddocks, V. Nizet, N. P. West, M. Ranson, M. J. Walker, J. D. McArthur and M. L. Sanderson-Smith. 2014. Plasmin(ogen) acquisition by group A *Streptococcus* protects against C3b-mediated neutrophil killing. J Innate Immun. 6: 240-250.

Lynskey, N. N., D. Goulding, M. Gierula, C. E. Turner, G. Dougan, R. J. Edwards and S. Sriskandan. 2013. RocA truncation underpins hyper-encapsulation, carriage longevity and transmissibility of serotype M18 group A streptococci. PLoS Pathog. 9: e1003842.

Maamary, P. G., N. L. Ben Zakour, J. N. Cole, A. Hollands, R. K. Aziz, T. C. Barnett, A. J. Cork, A. Henningham, M. Sanderson-Smith, J. D. McArthur, C. Venturini, C. M. Gillen, J. K. Kirk, D. R. Johnson, W. L. Taylor, E. L. Kaplan, M. Kotb, V. Nizet, S. A. Beatson and M. J. Walker. 2012. Tracing the evolutionary history of the pandemic group A streptococcal M1T1 clone. FASEB J. 26: 4675-4684.

Maamary, P. G., M. L. Sanderson-Smith, R. K. Aziz, A. Hollands, J. N. Cole, F. C. McKay, J. D. McArthur, J. K. Kirk, A. J. Cork, R. J. Keefe, R. G. Kansal, H. Sun, W. L. Taylor, G. S. Chhatwal, D. Ginsburg, V. Nizet, M. Kotb and M. J. Walker. 2010. Parameters governing invasive disease propensity of non-M1 serotype group A streptococci. J Innate Immun. 2: 596-606.

Madden, J. C., N. Ruiz and M. Caparon. 2001. Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria. Cell. 104: 143-152.

Malhotra-Kumar, S., S. Wang, C. Lammens, S. Chapelle and H. Goossens. 2003. Bacitracin-resistant clone of *Streptococcus pyogenes* isolated from pharyngitis patients in Belgium. J Clin Microbiol. 41: 5282-5284.

Manetti, A. G., C. Zingaretti, F. Falugi, S. Capo, M. Bombaci, F. Bagnoli, G. Gambellini, G. Bensi, M. Mora, A. M. Edwards, J. M. Musser, E. A. Graviss, J. L. Telford, G. Grandi and I. Margarit. 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. Mol Microbiol. 64: 968-983.

Maripuu, L., A. Eriksson and M. Norgren. 2008. Superantigen gene profile diversity among clinical group A streptococcal isolates. FEMS Immunol Med Microbiol. 54: 236-244.

Martin, J. M., M. Green, K. A. Barbadora and E. R. Wald. 2004. Group A streptococci among school-aged children: clinical characteristics and the carrier state. Pediatrics. 114: 1212-1219.

Massell, B. F., L. H. Honikman and J. Amezcua. 1969. Rheumatic fever following streptococcal vaccination. Report of three cases. JAMA. 207: 1115-1119.

Massell, B. F., J. G. Michael, J. Amezcua and M. Siner. 1968. Secondary and apparent primary antibody responses after group A streptococcal vaccination of 21 children. Appl Microbiol. 16: 509-518.

Maxted, W. R., J. P. Widdowson and C. A. Fraser. 1973. Antibody to streptococcal opacity factor in human sera. J Hyg (Lond). 71: 35-42.

Mayfield, J. A., Z. Liang, G. Agrahari, S. W. Lee, D. L. Donahue, V. A. Ploplis and F. J. Castellino. 2014. Mutations in the Control of Virulence Sensor Gene from *Streptococcus pyogenes* after Infection in Mice Lead to Clonal Bacterial Variants with Altered Gene Regulatory Activity and Virulence. PLoS One. 9: e100698.

McArthur, J. D., F. C. McKay, V. Ramachandran, P. Shyam, A. J. Cork, M. L. Sanderson-Smith, J. N. Cole, U. Ringdahl, U. Sjobring, M. Ranson and M. J. Walker. 2008. Allelic variants of streptokinase from *Streptococcus pyogenes* display functional differences in plasminogen activation. FASEB J. 22: 3146-3153.

McCormick, J. K., J. M. Yarwood and P. M. Schlievert. 2001. Toxic shock syndrome and bacterial superantigens: an update. Annu Rev Microbiol. 55: 77-104.

McGregor, K. F., B. G. Spratt, A. Kalia, A. Bennett, N. Bilek, B. Beall and D. E. Bessen. 2004. Multilocus sequence typing of *Streptococcus pyogenes* representing most known emm types and distinctions among subpopulation genetic structures. J Bacteriol. 186: 4285-4294.

McKay, F. C., J. D. McArthur, M. L. Sanderson-Smith, S. Gardam, B. J. Currie, K. S. Sriprakash, P. K. Fagan, R. J. Towers, M. R. Batzloff, G. S. Chhatwal, M. Ranson and M. J. Walker. 2004. Plasminogen binding by group A streptococcal isolates from a region of hyperendemicity for streptococcal skin infection and a high incidence of invasive infection. Infect Immun. 72: 364-370.

McMillan, D. J., P. A. Dreze, T. Vu, D. E. Bessen, J. Guglielmini, A. C. Steer, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2013a. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. Clin Microbiol Infect. 19: E222-229.

McMillan, D. J., M. L. Sanderson-Smith, P. R. Smeesters and K. S. Sriprakash. 2013b. Molecular markers for the study of streptococcal epidemiology. Curr Top Microbiol Immunol. 368: 29-48.

Miyoshi-Akiyama, T., D. Takamatsu, M. Koyanagi, J. Zhao, K. Imanishi and T. Uchiyama. 2005. Cytocidal effect of *Streptococcus pyogenes* on mouse neutrophils in vivo and the critical role of streptolysin S. J Infect Dis. 192: 107-116.

Molloy, E. M., P. D. Cotter, C. Hill, D. A. Mitchell and R. P. Ross. 2011. Streptolysin S-like virulence factors: the continuing sagA. Nat Rev Microbiol. 9: 670-681.

Montes, M., E. Tamayo, B. Orden, J. Larruskain and E. Perez-Trallero. 2010. Prevalence and clonal characterization of *Streptococcus pyogenes* clinical isolates with reduced fluoroquinolone susceptibility in Spain. Antimicrob Agents Chemother. 54: 93-97.

Mora, M., G. Bensi, S. Capo, F. Falugi, C. Zingaretti, A. G. Manetti, T. Maggi, A. R. Taddei, G. Grandi and J. L. Telford. 2005. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc Natl Acad Sci U S A. 102: 15641-15646.

Moses, A. E., M. R. Wessels, K. Zalcman, S. Alberti, S. Natanson-Yaron, T. Menes and E. Hanski. 1997. Relative contributions of hyaluronic acid capsule and M protein to virulence in a mucoid strain of the group A *Streptococcus*. Infect Immun. 65: 64-71.

Mulla, Z. D., P. E. Leaverton and S. T. Wiersma. 2003. Invasive group A streptococcal infections in Florida. South Med J. 96: 968-973.

Murakami, J., S. Kawabata, Y. Terao, K. Kikuchi, K. Totsuka, A. Tamaru, C. Katsukawa, K. Moriya, I. Nakagawa, I. Morisaki and S. Hamada. 2002. Distribution of emm genotypes and superantigen genes of *Streptococcus pyogenes* isolated in Japan, 1994-9. Epidemiol Infect. 128: 397-404.

Murphy, T. K., R. Kurlan and J. Leckman. 2010. The immunobiology of Tourette's disorder, pediatric autoimmune neuropsychiatric disorders associated with *Streptococcus*, and related disorders: a way forward. J Child Adolesc Psychopharmacol. 20: 317-331.

Murray, P. R., K. S. Rosenthal and M. A. Pfaller. 2013. Chapter 19. *Streptococcus*, p.188-196. In Medical Microbiology, 7th ed. Elsevier Saunders Inc., Philadelphia

Mylvaganam, H., B. Bjorvatn and A. Osland. 2000. Distribution and sequence variations of selected virulence genes among group A streptococcal isolates from western Norway. APMIS. 108: 771-778.

Nakata, M., T. Koller, K. Moritz, D. Ribardo, L. Jonas, K. S. McIver, T. Sumitomo, Y. Terao, S. Kawabata, A. Podbielski and B. Kreikemeyer. 2009. Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. Infect Immun. 77: 32-44.

Nakata, M., A. Podbielski and B. Kreikemeyer. 2005. MsmR, a specific positive regulator of the *Streptococcus pyogenes* FCT pathogenicity region and cytolysin-mediated translocation system genes. Mol Microbiol. 57: 786-803.

Neely, M. N., W. R. Lyon, D. L. Runft and M. Caparon. 2003. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. J Bacteriol. 185: 5166-5174.

Nilson, B. H., I. M. Frick, P. Akesson, S. Forsen, L. Bjorck, B. Akerstrom and M. Wikstrom. 1995. Structure and stability of protein H and the M1 protein from *Streptococcus pyogenes*. Implications for other surface proteins of gram-positive bacteria. Biochemistry. 34: 13688-13698.

Nomizu, M., G. Pietrzynski, T. Kato, P. Lachance, R. Menard and E. Ziomek. 2001. Substrate specificity of the streptococcal cysteine protease. J Biol Chem. 276: 44551-44556.

Nooh, M. M., R. K. Aziz, M. Kotb, A. Eroshkin, W. J. Chuang, T. Proft and R. Kansal. 2006. Streptococcal mitogenic exotoxin, SmeZ, is the most susceptible M1T1 streptococcal superantigen to degradation by the streptococcal cysteine protease, SpeB. J Biol Chem. 281: 35281-35288.

Nooh, M. M., S. Nookala, R. Kansal and M. Kotb. 2011. Individual genetic variations directly effect polarization of cytokine responses to superantigens associated with streptococcal sepsis: implications for customized patient care. J Immunol. 186: 3156-3163.

Norrby-Teglund, A., K. Pauksens, S. E. Holm and M. Norgren. 1994. Relation between low capacity of human sera to inhibit streptococcal mitogens and serious manifestation of disease. J Infect Dis. 170: 585-591.

Nyberg, P., M. Rasmussen and L. Bjorck. 2004a. alpha2-Macroglobulin-proteinase complexes protect *Streptococcus pyogenes* from killing by the antimicrobial peptide LL-37. J Biol Chem. 279: 52820-52823.

Nyberg, P., M. Rasmussen, U. Von Pawel-Rammingen and L. Bjorck. 2004b. SpeB modulates fibronectin-dependent internalization of *Streptococcus pyogenes* by efficient proteolysis of cell-wall-anchored protein F1. Microbiology. 150: 1559-1569.

O'Loughlin, R. E., A. Roberson, P. R. Cieslak, R. Lynfield, K. Gershman, A. Craig, B. A. Albanese, M. M. Farley, N. L. Barrett, N. L. Spina, B. Beall, L. H. Harrison, A. Reingold and C. Van Beneden. 2007. The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. Clin Infect Dis. 45: 853-862.

Oehmcke, S., O. Shannon, M. Morgelin and H. Herwald. 2010. Streptococcal M proteins and their role as virulence determinants. Clin Chim Acta. 411: 1172-1180.

Okada, N., M. K. Liszewski, J. P. Atkinson and M. Caparon. 1995. Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A *Streptococcus*. Proc Natl Acad Sci U S A. 92: 2489-2493.

Olsen, R. J., D. R. Laucirica, M. E. Watkins, M. L. Feske, J. R. Garcia-Bustillos, C. Vu, C. Cantu, S. A. Shelburne, 3rd, N. Fittipaldi, M. Kumaraswami, P. R. Shea, A. R. Flores, S. B. Beres, M. Lovgren, G. J. Tyrrell, A. Efstratiou, D. E. Low, C. A.

Van Beneden and J. M. Musser. 2012. Polymorphisms in regulator of protease B (RopB) alter disease phenotype and strain virulence of serotype M3 group A *Streptococcus*. J Infect Dis. 205: 1719-1729.

Olsen, R. J. and J. M. Musser. 2010. Molecular pathogenesis of necrotizing fasciitis. Annu Rev Pathol. 5: 1-31.

Olsen, R. J., A. Raghuram, C. Cantu, M. H. Hartman, F. E. Jimenez, S. Lee, A. Ngo, K. A. Rice, D. Saddington, H. Spillman, C. Valson, A. R. Flores, S. B. Beres, S. W. Long, W. Nasser and J. M. Musser. 2015. The Majority of 9,729 Group A *Streptococcus* Strains Causing Disease Secrete SpeB Cysteine Protease: Pathogenesis Implications. Infect Immun. 83: 4750-4758.

Pahlman, L. I., M. Morgelin, J. Eckert, L. Johansson, W. Russell, K. Riesbeck, O. Soehnlein, L. Lindbom, A. Norrby-Teglund, R. R. Schumann, L. Bjorck and H. Herwald. 2006. Streptococcal M protein: a multipotent and powerful inducer of inflammation. J Immunol. 177: 1221-1228.

Pahlman, L. I., A. I. Olin, J. Darenberg, M. Morgelin, M. Kotb, H. Herwald and A. Norrby-Teglund. 2008. Soluble M1 protein of *Streptococcus pyogenes* triggers potent T cell activation. Cell Microbiol. 10: 404-414.

Perez-Caballero, D., S. Alberti, F. Vivanco, P. Sanchez-Corral and S. Rodriguez de Cordoba. 2000. Assessment of the interaction of human complement regulatory proteins with group A *Streptococcus*. Identification of a high-affinity group A *Streptococcus* binding site in FHL-1. Eur J Immunol. 30: 1243-1253.

Pires, R., C. Ardanuy, D. Rolo, A. Morais, A. Brito-Avo, J. Goncalo-Marques, J. Linares and I. Santos-Sanches. 2010. Emergence of ciprofloxacin-nonsusceptible *Streptococcus pyogenes* isolates from healthy children and pediatric patients in Portugal. Antimicrob Agents Chemother. 54: 2677-2680.

Pires, R., D. Rolo, R. Mato, J. Feio de Almeida, C. Johansson, B. Henriques-Normark, A. Morais, A. Brito-Avo, J. Goncalo-Marques and I. Santos-Sanches. 2009. Resistance to bacitracin in *Streptococcus pyogenes* from oropharyngeal colonization and noninvasive infections in Portugal was caused by two clones of distinct virulence genotypes. FEMS Microbiol Lett. 296: 235-240.

Pires, R., D. Rolo, A. Morais, A. Brito-Avo, C. Johansson, B. Henriques-Normark, J. Goncalo-Marques and I. Santos-Sanches. 2011. Description of macrolide-resistant and potential virulent clones of *Streptococcus pyogenes* causing asymptomatic colonization during 2000-2006 in the Lisbon area. Eur J Clin Microbiol Infect Dis.

Pletz, M. W., L. McGee, C. A. Van Beneden, S. Petit, M. Bardsley, M. Barlow and K. P. Klugman. 2006. Fluoroquinolone resistance in invasive Streptococcus *pyogenes* isolates due to spontaneous mutation and horizontal gene transfer. Antimicrob Agents Chemother. 50: 943-948.

Polly, S. M., R. H. Waldman, P. High, M. K. Wittner and A. Dorfman. 1975. Protective studies with a group A streptococcal M protein vaccine. II. Challange of volenteers after local immunization in the upper respiratory tract. J Infect Dis. 131: 217-224.

Proft, T. and J. D. Fraser. 2016. Streptococcal Superantigens: Biological properties and potential role in disease.

Proft, T., S. Sriskandan, L. Yang and J. D. Fraser. 2003. Superantigens and streptococcal toxic shock syndrome. Emerg Infect Dis. 9: 1211-1218.

Raeder, R., M. Woischnik, A. Podbielski and M. D. Boyle. 1998. A secreted streptococcal cysteine protease can cleave a surface-expressed M1 protein and alter the immunoglobulin binding properties. Res Microbiol. 149: 539-548.

Ranjbar, R., A. Karami, S. Farshad, G. M. Giammanco and C. Mammina. 2014. Typing methods used in the molecular epidemiology of microbial pathogens: a how-to guide. New Microbiol. 37: 1-15.

Reglinski, M. and S. Sriskandan. 2014. The contribution of group A streptococcal virulence determinants to the pathogenesis of sepsis. Virulence. 5: 127-136.

Ringdahl, U., H. G. Svensson, H. Kotarsky, M. Gustafsson, M. Weineisen and U. Sjobring. 2000. A role for the fibrinogen-binding regions of streptococcal M proteins in phagocytosis resistance. Mol Microbiol. 37: 1318-1326.

Roberts, M. C. 2005. Update on acquired tetracycline resistance genes. FEMS Microbiol Lett. 245: 195-203.

Rouadi, P., F. M. Baroody, D. Abbott, E. Naureckas, J. Solway and R. M. Naclerio. 1999. A technique to measure the ability of the human nose to warm and humidify air. J Appl Physiol (1985). 87: 400-406.

Sanderson-Smith, M., D. M. De Oliveira, J. Guglielmini, D. J. McMillan, T. Vu, J. K. Holien, A. Henningham, A. C. Steer, D. E. Bessen, J. B. Dale, N. Curtis, B. W. Beall, M. J. Walker, M. W. Parker, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2014. A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. J Infect Dis. 210: 1325-1338.

Sandin, C., F. Carlsson and G. Lindahl. 2006. Binding of human plasma proteins to *Streptococcus pyogenes* M protein determines the location of opsonic and non-opsonic epitopes. Mol Microbiol. 59: 20-30.

Sawai, J., T. Hasegawa, T. Kamimura, A. Okamoto, D. Ohmori, N. Nosaka, K. Yamada, K. Torii and M. Ohta. 2007. Growth phase-dependent effect of clindamycin on production of exoproteins by *Streptococcus pyogenes*. Antimicrob Agents Chemother. 51: 461-467.

Schmitz, F. J., A. Beyer, E. Charpentier, B. H. Normark, M. Schade, A. C. Fluit, D. Hafner and R. Novak. 2003. Toxin-gene profile heterogeneity among endemic invasive European group A streptococcal isolates. J Infect Dis. 188: 1578-1586.

Shaikh, N., E. Leonard and J. M. Martin. 2010. Prevalence of streptococcal pharyngitis and streptococcal carriage in children: a meta-analysis. Pediatrics. 126: e557-564.

Shea, P. R., S. B. Beres, A. R. Flores, A. L. Ewbank, J. H. Gonzalez-Lugo, A. J. Martagon-Rosado, J. C. Martinez-Gutierrez, H. A. Rehman, M. Serrano-

Gonzalez, N. Fittipaldi, S. D. Ayers, P. Webb, B. M. Willey, D. E. Low and J. M. Musser. 2011. Distinct signatures of diversifying selection revealed by genome analysis of respiratory tract and invasive bacterial populations. Proc Natl Acad Sci U S A. 108: 5039-5044.

Shelburne, S. A., 3rd, C. Granville, M. Tokuyama, I. Sitkiewicz, P. Patel and J. M. Musser. 2005. Growth characteristics of and virulence factor production by group A *Streptococcus* during cultivation in human saliva. Infect Immun. 73: 4723-4731.

Shelburne, S. A., 3rd, R. J. Olsen, N. Makthal, N. G. Brown, P. Sahasrabhojane, E. M. Watkins, T. Palzkill, J. M. Musser and M. Kumaraswami. 2011. An amino-terminal signal peptide of Vfr protein negatively influences RopB-dependent SpeB expression and attenuates virulence in *Streptococcus pyogenes*. Mol Microbiol. 82: 1481-1495.

Shulman, S. T., A. L. Bisno, H. W. Clegg, M. A. Gerber, E. L. Kaplan, G. Lee, J. M. Martin and C. Van Beneden. 2012. Clinical practice guideline for the diagnosis and management of group A streptococcal pharyngitis: 2012 update by the Infectious Diseases Society of America. Clin Infect Dis. 55: 1279-1282.

Shulman, S. T., R. R. Tanz, J. B. Dale, B. Beall, W. Kabat, K. Kabat, E. Cederlund, D. Patel, J. Rippe, Z. Li and V. Sakota. 2009. Seven-year surveillance of north american pediatric group a streptococcal pharyngitis isolates. Clin Infect Dis. 49: 78-84.

Shulman, S. T., R. R. Tanz, W. Kabat, K. Kabat, E. Cederlund, D. Patel, Z. Li, V. Sakota, J. B. Dale and B. Beall. 2004. Group A streptococcal pharyngitis serotype surveillance in North America, 2000-2002. Clin Infect Dis. 39: 325-332.

Siemens, N., N. Patenge, J. Otto, T. Fiedler and B. Kreikemeyer. 2011. *Streptococcus pyogenes* M49 plasminogen/plasmin binding facilitates keratinocyte invasion via integrin-integrin-linked kinase (ILK) pathways and protects from macrophage killing. J Biol Chem. 286: 21612-21622.

Silva-Costa, C., J. A. Carrico, M. Ramirez and J. Melo-Cristino. 2014. Scarlet fever is caused by a limited number of *Streptococcus pyogenes* lineages and is associated with the exotoxin genes ssa, speA and speC. Pediatr Infect Dis J. 33: 306-310.

Silva-Costa, C., A. Friães, M. Ramirez and J. Melo-Cristino. 2015a. Macrolide-resistant *Streptococcus pyogenes*: prevalence and treatment strategies. Expert Rev Anti Infect Ther. 13: 615-628.

Silva-Costa, C., F. R. Pinto, M. Ramirez and J. Melo-Cristino. 2008. Decrease in macrolide resistance and clonal instability among *Streptococcus pyogenes* in Portugal. Clin Microbiol Infect. 14: 1152-1159.

Silva-Costa, C., M. Ramirez and J. Melo-Cristino. 2015b. Declining macrolide resistance in *Streptococcus pyogenes* in Portugal (2007-13) was accompanied by continuous clonal changes. J Antimicrob Chemother. 70: 2729-2733.

Sladden, M. J. and G. A. Johnston. 2004. Common skin infections in children. BMJ. 329: 95-99.

Smeesters, P. R., M. Dramaix and L. Van Melderen. 2010a. The *emm*-type diversity does not always reflect the M protein genetic diversity--is there a case for designer vaccine against GAS. Vaccine. 28: 883-885.

Smeesters, P. R., D. J. McMillan and K. S. Sriprakash. 2010b. The streptococcal M protein: a highly versatile molecule. Trends Microbiol. 18: 275-282.

Sriskandan, S., L. Faulkner and P. Hopkins. 2007. Streptococcus pyogenes: Insight into the function of the streptococcal superantigens. Int J Biochem Cell Biol. 39: 12-19.

Staali, L., S. Bauer, M. Morgelin, L. Bjorck and H. Tapper. 2006. *Streptococcus pyogenes* bacteria modulate membrane traffic in human neutrophils and selectively inhibit azurophilic granule fusion with phagosomes. Cell Microbiol. 8: 690-703.

Stevens, D. L. 2000. Streptococcal toxic shock syndrome associated with necrotizing fasciitis. Annu Rev Med. 51: 271-288.

Stevens, D. L., A. L. Bisno, H. F. Chambers, E. D. Everett, P. Dellinger, E. J. Goldstein, S. L. Gorbach, J. V. Hirschmann, E. L. Kaplan, J. G. Montoya and J. C. Wade. 2005. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. Clin Infect Dis. 41: 1373-1406.

Stevens, D. L., A. E. Gibbons, R. Bergstrom and V. Winn. 1988. The Eagle effect revisited: efficacy of clindamycin, erythromycin, and penicillin in the treatment of streptococcal myositis. J Infect Dis. 158: 23-28.

Sugareva, V., R. Arlt, T. Fiedler, C. Riani, A. Podbielski and B. Kreikemeyer. 2010. Serotype- and strain- dependent contribution of the sensor kinase CovS of the CovRS two-component system to *Streptococcus pyogenes* pathogenesis. BMC Microbiol. 10: 34.

Sumby, P., K. D. Barbian, D. J. Gardner, A. R. Whitney, D. M. Welty, R. D. Long, J. R. Bailey, M. J. Parnell, N. P. Hoe, G. G. Adams, F. R. Deleo and J. M. Musser. 2005a. Extracellular deoxyribonuclease made by group A *Streptococcus* assists pathogenesis by enhancing evasion of the innate immune response. Proc Natl Acad Sci U S A. 102: 1679-1684.

Sumby, P., S. F. Porcella, A. G. Madrigal, K. D. Barbian, K. Virtaneva, S. M. Ricklefs, D. E. Sturdevant, M. R. Graham, J. Vuopio-Varkila, N. P. Hoe and J. M. Musser. 2005b. Evolutionary origin and emergence of a highly successful clone of serotype M1 group a *Streptococcus* involved multiple horizontal gene transfer events. J Infect Dis. 192: 771-782.

Sumby, P., A. R. Whitney, E. A. Graviss, F. R. DeLeo and J. M. Musser. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog. 2: e5.

Sun, H., U. Ringdahl, J. W. Homeister, W. P. Fay, N. C. Engleberg, A. Y. Yang, L. S. Rozek, X. Wang, U. Sjobring and D. Ginsburg. 2004. Plasminogen is a critical

host pathogenicity factor for group A streptococcal infection. Science. 305: 1283-1286.

Svensson, M. D., D. A. Scaramuzzino, U. Sjobring, A. Olsen, C. Frank and D. E. Bessen. 2000. Role for a secreted cysteine proteinase in the establishment of host tissue tropism by group A streptococci. Mol Microbiol. 38: 242-253.

Svensson, M. D., U. Sjobring, F. Luo and D. E. Bessen. 2002. Roles of the plasminogen activator streptokinase and the plasminogen-associated M protein in an experimental model for streptococcal impetigo. Microbiology. 148: 3933-3945.

Talkington, D. F., B. Schwartz, C. M. Black, J. K. Todd, J. Elliott, R. F. Breiman and R. R. Facklam. 1993. Association of phenotypic and genotypic characteristics of invasive *Streptococcus pyogenes* isolates with clinical components of streptococcal toxic shock syndrome. Infect Immun. 61: 3369-3374.

Tamura, F., R. Nakagawa, T. Akuta, S. Okamoto, S. Hamada, H. Maeda, S. Kawabata and T. Akaike. 2004. Proapoptotic effect of proteolytic activation of matrix metalloproteinases by *Streptococcus pyogenes* thiol proteinase (Streptococcus pyrogenic exotoxin B). Infect Immun. 72: 4836-4847.

Terao, Y., Y. Mori, M. Yamaguchi, Y. Shimizu, K. Ooe, S. Hamada and S. Kawabata. 2008. Group A streptococcal cysteine protease degrades C3 (C3b) and contributes to evasion of innate immunity. J Biol Chem. 283: 6253-6260.

Trevino, J., N. Perez, E. Ramirez-Pena, Z. Liu, S. A. Shelburne, 3rd, J. M. Musser and P. Sumby. 2009. CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A *Streptococcus* virulence factor-encoding genes. Infect Immun. 77: 3141-3149.

Turner, C. E., J. Abbott, T. Lamagni, M. T. Holden, S. David, M. D. Jones, L. Game, A. Efstratiou and S. Sriskandan. 2015. Emergence of a New Highly Successful Acapsular Group A *Streptococcus* Clade of Genotype *emm89* in the United Kingdom. MBio. 6: e00622.

Turner, C. E., P. Kurupati, M. D. Jones, R. J. Edwards and S. Sriskandan. 2009. Emerging role of the interleukin-8 cleaving enzyme SpyCEP in clinical *Streptococcus pyogenes* infection. J Infect Dis. 200: 555-563.

Turner, C. E., M. Sommerlad, K. McGregor, F. J. Davies, B. Pichon, D. L. Chong, L. Farzaneh, M. T. Holden, B. G. Spratt, A. Efstratiou and S. Sriskandan. 2012. Superantigenic activity of *emm3 Streptococcus pyogenes* is abrogated by a conserved, naturally occurring *smeZ* mutation. PLoS One. 7: e46376.

Uchiyama, S., F. Andreoni, R. A. Schuepbach, V. Nizet and A. S. Zinkernagel. 2012. DNase Sda1 Allows Invasive M1T1 Group A *Streptococcus* to Prevent TLR9-Dependent Recognition. PLoS Pathog. 8: e1002736.

Ulrich, R. G. 2008. Vaccine based on a ubiquitous cysteinyl protease and streptococcal pyrogenic exotoxin A protects against *Streptococcus pyogenes* sepsis and toxic shock. J Immune Based Ther Vaccines. 6: 8.

van Belkum, A., P. T. Tassios, L. Dijkshoorn, S. Haeggman, B. Cookson, N. K. Fry, V. Fussing, J. Green, E. Feil, P. Gerner-Smidt, S. Brisse and M. Struelens. 2007.

Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 13 Suppl 3: 1-46.

Varaldo, P. E., M. P. Montanari and E. Giovanetti. 2009. Genetic elements responsible for erythromycin resistance in streptococci. Antimicrob Agents Chemother. 53: 343-353.

Vlaminckx, B. J., E. M. Mascini, J. Schellekens, L. M. Schouls, A. Paauw, A. C. Fluit, R. Novak, J. Verhoef and F. J. Schmitz. 2003. Site-specific manifestations of invasive group a streptococcal disease: type distribution and corresponding patterns of virulence determinants. J Clin Microbiol. 41: 4941-4949.

Voyich, J. M., K. R. Braughton, D. E. Sturdevant, C. Vuong, S. D. Kobayashi, S. F. Porcella, M. Otto, J. M. Musser and F. R. DeLeo. 2004. Engagement of the pathogen survival response used by group A *Streptococcus* to avert destruction by innate host defense. J Immunol. 173: 1194-1201.

Walker, M. J., T. C. Barnett, J. D. McArthur, J. N. Cole, C. M. Gillen, A. Henningham, K. S. Sriprakash, M. L. Sanderson-Smith and V. Nizet. 2014. Disease manifestations and pathogenic mechanisms of group a *Streptococcus*. Clin Microbiol Rev. 27: 264-301.

Walker, M. J., A. Hollands, M. L. Sanderson-Smith, J. N. Cole, J. K. Kirk, A. Henningham, J. D. McArthur, K. Dinkla, R. K. Aziz, R. G. Kansal, A. J. Simpson, J. T. Buchanan, G. S. Chhatwal, M. Kotb and V. Nizet. 2007. DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat Med. 13: 981-985.

Wang, H., R. Lottenberg and M. D. Boyle. 1995a. Analysis of the interaction of group A streptococci with fibrinogen, streptokinase and plasminogen. Microb Pathog. 18: 153-166.

Wang, H., R. Lottenberg and M. D. Boyle. 1995b. A role for fibrinogen in the streptokinase-dependent acquisition of plasmin(ogen) by group A streptococci. J Infect Dis. 171: 85-92.

Wei, L., V. Pandiripally, E. Gregory, M. Clymer and D. Cue. 2005. Impact of the SpeB protease on binding of the complement regulatory proteins factor H and factor H-like protein 1 by *Streptococcus pyogenes*. Infect Immun. 73: 2040-2050.

Wessels, M. R. 2011. Clinical practice. Streptococcal pharyngitis. N Engl J Med. 364: 648-655.

Wessels, M. R., A. E. Moses, J. B. Goldberg and T. J. DiCesare. 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. Proc Natl Acad Sci U S A. 88: 8317-8321.

Yamaguchi, M., Y. Terao and S. Kawabata. 2013. Pleiotropic virulence factor - *Streptococcus pyogenes* fibronectin-binding proteins. Cell Microbiol. 15: 503-511.

Yoshino, M., S. Y. Murayama, K. Sunaoshi, T. Wajima, M. Takahashi, J. Masaki, I. Kurokawa and K. Ubukata. 2010. Nonhemolytic *Streptococcus pyogenes*

isolates that lack large regions of the sag operon mediating streptolysin S production. J Clin Microbiol. 48: 635-638.

Yu, C. E. and J. J. Ferretti. 1989. Molecular epidemiologic analysis of the type A streptococcal exotoxin (erythrogenic toxin) gene (*speA*) in clinical *Streptococcus pyogenes* strains. Infect Immun. 57: 3715-3719.

Zhu, L., R. J. Olsen, W. Nasser, I. de la Riva Morales and J. M. Musser. 2015. Trading Capsule for Increased Cytotoxin Production: Contribution to Virulence of a Newly Emerged Clade of *emm89 Streptococcus pyogenes*. MBio. 6: e01378-01315.

Zimbelman, J., A. Palmer and J. Todd. 1999. Improved outcome of clindamycin compared with beta-lactam antibiotic treatment for invasive *Streptococcus pyogenes* infection. Pediatr Infect Dis J. 18: 1096-1100.

AIMS OF THE THESIS

According to WHO, GAS is considered one of the pathogens responsible for the highest rates of mortality with an estimated 500 000 deaths per year. Its ability to cause a wide variety of infections, differing according to tissue type and severity, associated with its autoimmune complications, is one of the most particular features of GAS.

Surveillance studies evaluating the phenotypic and molecular characteristics of the isolates are the basis of all GAS research. Additionally to the overall prevalence and evaluation of antimicrobial resistance rates, these studies, through the use of typing methods, allow the identification of lineages, as well as their characteristics and their distribution among the different infections. Globally, this information highlighted that some lineages are more associated with certain types of infection. However, despite several theories, the specific characteristics responsible for a tissue preference or for an enhanced invasive potential of an isolate remain unknown. Among these theories, is the acquisition of spontaneous mutations in the CovRS regulator as a potential mechanism to increase the invasive potential. However, contrasting results have been reported regarding the prevalence of these mutations among invasive and non-invasive isolates. Therefore, we aimed to evaluate the prevalence of these mutations in isolates recovered from different infections and their distribution among different lineages. In order to address these questions, we sequenced the *covRS* operon in a collection of 191 isolates comprising isolates from distinct lineages recovered from invasive infections and pharyngitis. Since these mutations are associated with different expression patterns of several virulence factors, we also aimed to evaluate if this occurs in our isolates. Hence, we screened all isolates for the activity of SpeB, NADase, and SLS. The downregulation of *speB*, as a consequence of *covRS* mutations, has been considered the determinant step in the transition to an invasive phenotype. Additionally to CovRS, *speB* is also directly controlled by RopB. Similarity to CovRS, *ropB* mutations have been also recovered from humans infections, and therefore we also wanted to know their importance. For that, we sequenced the *ropB* gene of all isolates.

GAS invasive diseases are frequently considered as a complication of skin and soft tissue infections. However, in comparison with the amount of knowledge referring to invasive and pharyngitis isolates there is little information regarding isolates recovered from SSTI. Therefore we characterized, by multiple typing methods, 320 isolates from purulent exudates recovered in Portugal and performed a comparison to

313 contemporary invasive isolates, which were previously characterized. Additionally, we aimed to know if there were differences in comparison with invasive and also pharyngitis isolates regarding the prevalence of mutations predicted to eliminate protein function in *covRS* and *ropB* genes and also regarding SpeB activity. With this purpose, we tested all 320 SSTI isolates for SpeB activity and in those without detectable activity we sequenced the *covRS* and *ropB* genes.

CHAPTER II

CONSEQUENCES OF THE VARIABILITY OF THE OF COVRS AND ROPB REGULATORS AMONG *STREPTOCOCCUS PYOGENES* CAUSING HUMAN INFECTIONS

This chapter is a reproduction of the following publication:

Friães, A., C. Pato, J. Melo-Cristino and M. Ramirez. 2015. Consequences of the variability of the CovRS and RopB regulators among *Streptococcus pyogenes* causing human infections. Sci Rep. 5: 12057.

The supplementary information referred throughout the text can be consulted in this chapter before the section of references.

My contribution to this publication included the decision of choosing SpeB, NADase and SLS as the virulence factors under the control of CovRS, in order to evaluate the impact of mutations in this regulator. I optimized and applied the two methods used for screen the SpeB activity and the western blot for its detection. I started the optimization of the methods for measure the activity of NADase and SLS. The majority of statistical analysis was also my responsibility. Regarding the draft of manuscript, I wrote the parts related with SpeB, and was involved during all the revision process.

Consequences of the variability of the of CovRS and RopB regulators among *Streptococcus pyogenes* causing human infections[1]

Ana Friães*, Catarina Pato*, José Melo-Cristino and Mário Ramirez

Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

*These authors contributed equally to this work

SUMMARY

To evaluate the importance of *covRS* and *ropB* mutations in invasive disease caused by Group A Streptococci (GAS), we determined the sequence of the *covRS* and *ropB* genes of 191 isolates from invasive infections and pharyngitis, comprising a diverse set of *emm* types and multilocus sequence types. The production of SpeB and the activity of NAD glycohydrolase (NADase) and streptolysin S (SLS) were evaluated. The results support the acquisition of null *covS* alleles (predicted to eliminate protein function), resulting in downregulation of SpeB and upregulation of NADase and SLS, as a mechanism possibly contributing to higher invasiveness. Among the isolates tested, this mechanism was found to be uncommon (10% of invasive isolates) and was not more prevalent among clones with enhanced invasiveness (including M1T1) but occurred in diverse genetic backgrounds. In lineages such as *emm*64, these changes did not result in upregulation of NADase and SLS, highlighting the diversity of regulatory pathways in GAS. Despite abrogating SpeB production, null alleles in *ropB* were not associated with invasive infection. The *covRS* and *ropB* genes are under stabilising selection and no expansion of isolates carrying null alleles has been observed, suggesting that the presence of these regulators is important for overall fitness.

[1] A facsimile of this publication is found at "APPENDIX I".

INTRODUCTION

Streptococcus pyogenes (Group A Streptococci, GAS) is a human pathogen that can asymptomatically colonize the oropharynx, but is also responsible for a variety of human diseases, ranging from uncomplicated superficial infections of the respiratory tract and skin, such as pharyngitis and impetigo, to severe invasive infections associated with high morbidity and mortality, like necrotizing fasciitis and streptococcal toxic shock syndrome (STSS) (Cunningham, 2000).

The recognition of the ability of strains belonging to the highly invasive M1T1 clone to acquire mutations in the *covRS* genes during skin and soft tissue infection in mice suggested that GAS could increase its capacity to invade deeper tissues by altering its regulatory networks in order to produce a switch to an invasive transcriptome profile (Engleberg *et al.*, 2001; Sumby *et al.*, 2006). The adjacent and cotranscribed genes *covR* and *covS* encode the two-component regulatory system CovRS (also known as CsrRS). The DNA-binding response regulator CovR acts mostly as a repressor of transcription upon phosphorylation by the sensor kinase/phosphatase CovS, which responds to stress factors, such as elevated temperatures, high saline concentrations, and decreased pH (Churchward, 2007). CovRS is estimated to directly or indirectly influence the expression of 10–15% of the GAS genome. Although it is also involved in the regulation of metabolic genes, it is mostly known for controlling the expression of a large number of genes encoding factors that promote GAS virulence and evasion of the host immune system (Graham *et al.*, 2002; Horstmann *et al.*, 2014). Most studies indicate that mutations impairing CovRS function result in the upregulation of genes encoding the hyaluronic acid capsule, streptolysins S and O, streptokinase, DNases, the interleukin-protease SpyCEP, and NAD-glycohydrolase (NADase), among others, and in the downregulation of proteins like the extracellular cysteine protease SpeB and the protein-G-related α 2-macroglobulin-binding protein GRAB (Hasegawa *et al.*, 2010; Aziz *et al.*, 2010). However, contrasting results have been reported regarding the effect of *covRS* mutations in some of those virulence factors, which may be partly explained by the interaction of CovRS with other transcriptional regulators, resulting in complex regulation patterns that can vary between different strains (Aziz *et al.*, 2010; Sugareva *et al.*, 2010).

Mutations in CovRS and in the stand-alone transcriptional regulator RopB have been identified in isolates recovered from human infections (Hasegawa *et al.*, 2010; Carroll *et al.*, 2011; Ikebe *et al.*, 2010; Lin *et al.*, 2014). RopB (also known as Rgg) is encoded by the *ropB* (*rgg*) gene, which is located 940 bp away from *speB*, in the opposite DNA strand, and directly binds the promoter of the latter gene to activate its transcription (Neely *et al.*, 2003). In some strains, this regulator has been reported to affect the transcription of other virulence factors, including DNases, the hyaluronic acid capsule, NADase, streptokinase, streptolysins, and phage-encoded superantigens, among others (Carroll *et al.*, 2011; Ikebe *et al.*, 2010).

The downregulation of SpeB, which is usually observed as a consequence of CovRS and RopB mutations, is considered to be a determining step in the transition to an invasive phenotype, since this potent protease degrades several extracellular GAS virulence factors that play an important role in the invasive process, including the M protein, the F1 protein, C5a peptidase, streptokinase, and SmeZ (Carroll and Musser, 2011). In agreement, SpeB production has been inversely correlated with disease severity, in both human infections and murine models (Kansal *et al.*, 2000; Ashbaugh *et al.*, 1998). SpeB is initially synthesised and secreted as an inactive 42-kDa zymogen, which is then autocatalytically processed through a series of intermediates to form the 28-kDa mature active protease. Each stage from transcription to mature SpeB involves tight regulatory controls, including post-transcriptional and post-translational mechanisms (Carroll and Musser, 2011).

Some studies suggest that *covRS* and *ropB* mutations occur more frequently among isolates from invasive infections, while others found them in a similar proportion among non-invasive isolates (Lin *et al.*, 2014; Hasegawa *et al.*, 2010; Ikebe *et al.*, 2010). Contrasting observations have also been reported regarding the impact of alterations in these regulators on virulence using animal models, as well as on the expression of virulence factors like SpeB and the streptolysins (Engleberg *et al.*, 2001; Sumby *et al.*, 2006; Graham *et al.*, 2002; Hasegawa *et al.*, 2010; Aziz *et al.*, 2010; Ikebe *et al.*, 2010; Mayfield *et al.*, 2014; Miller *et al.*, 2001; Trevino *et al.*, 2009; Ravins *et al.*, 2000). Given these discrepancies and the fact that most of the studies have focused on particular *emm* types, especially *emm*1 and *emm*3, the importance of the acquisition of these mutations for the overall invasiveness and virulence of GAS strains is still not completely understood.

In order to address these questions, we sequenced the *covRS* and *ropB* genes in a collection of 191 GAS isolates presenting a high diversity of *emm* types and multilocus sequence types (STs), including the most prevalent clones causing pharyngitis and invasive infections in Portugal and in a majority of other countries from temperate climate regions (Friães *et al.*, 2012; Shulman *et al.*, 2004; Luca-Harari *et al.*, 2009). The genetic diversity of the genes was evaluated, and the respective alleles were correlated with the activity of SpeB and of two other virulence factors whose expression has been suggested to be under the influence of CovRS and RopB, at least in some strains, namely the NADase and the streptolysin S (SLS) (Hasegawa *et al.*, 2010; Ikebe *et al.*, 2010; Miller *et al.*, 2001; Hollands *et al.*, 2008).

RESULTS

Genetic variation of the *covR*, *covS* and *ropB* genes.

In the present work, the sequence of the *covR*, *covS*, and *ropB* genes was determined for a collection of 191 isolates (Supplementary Table S1) comprising one third of the isolates of each *emm* type present in a larger collection of strains recovered from pharyngitis and invasive infections in Portugal, which has been characterized elsewhere (Friães *et al.*, 2012). In addition, all isolates of *emm* types 1 and 64 were included in this study, since these two *emm* types were shown to be significantly associated with invasive disease.

The three genes presented a high allelic diversity in the studied GAS collection, with Simpson's index of diversity (SID) values close to the ones obtained for *emm* type and ST (Table 2). The allelic diversity of the *covR* genes was lower than that of *covS* ($P = 0.0140$) and of the ST ($P = 0.0137$). Overall, isolates sharing the same *emm* type or ST frequently shared the same *covR*, *covS*, and *ropB* alleles (Supplementary Table S2). This association was stronger for *covR*, in line with the lower allelic diversity of this gene. These results indicate that the *covR*, *covS*, and *ropB* alleles present in a given isolate were essentially clonal properties, although some level of intra-clonal diversity was observed, particularly for *covS* and *ropB*.

Table 2. Simpson's index of diversity (SID) and corresponding 95% confidence intervals ($CI_{95\%}$) of the alleles of *covR*, *covS*, and *ropB*, the *emm* types, and the STs identified among the 191 GAS isolates analyzed in this study.

Gene	All isolates (n = 191)		Invasive isolates (n = 87)		Pharyngitis isolates (n = 104)	
	No. partitions	SID [$CI_{95\%}$]	No. partitions	SID [$CI_{95\%}$]	No. partitions	SID [$CI_{95\%}$]
covR	19	0.831 [0.793-0.870]	17	0.816 [0.752-0.880]	13	0.839 [0.794-0.885]
covS	43	0.893 [0.861-0.925]	31	0.870 [0.811-0.930]	25	0.905 [0.873-0.938]
ropB	38	0.850 [0.805-0.896]	23	0.806 [0.729-0.883]	29	0.879 [0.828-0.929]
emm type	26	0.877 [0.844-0.910]	20	0.833 [0.769-0.898]	20	0.900 [0.868-0.932]
ST	41	0.894 [0.861-0.927]	27	0.860 [0.798-0.922]	32	0.910 [0.876-0.945]

Together with the alleles previously deposited in GenBank, we identified a total of 54 alleles for *covR*, 121 for *covS*, and 78 for *ropB* (Table 1). One of the isolates in our study is devoid of the *ropB* gene due to a previously characterized deletion that also involves the *speB* gene (Friães *et al.*, 2013). Alterations relative to the nucleotide sequence of the three genes in strain SF370 are depicted in Fig. 1. For each of the genes, more than half of the alleles result in alterations in the amino acid sequence, with *ropB* presenting the highest proportion of distinct amino acid sequences relative to the total number of alleles. However, all the *K*a/*K*s ratios were significantly lower than 1 (Table 1), indicating that all three genes are under stabilising selection. *K*a/*K*s is the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site. This measure is often used to evaluate deviations from neutral evolution. A value of *K*a/*K*s $< < 1$ indicates that the protein is under stabilising selection, i.e. selection favours alleles that do not change the amino acid sequence. In contrast, a value of *K*a/*K*s $> > 1$ is strong evidence for positive selection, indicating that selection drove changes to the protein (Hurst, 2002).

Table 1. Type of amino acid alteration, nucleotide diversity, and *Ka*/*Ks* values associated with the *covR*, *covS*, and *ropB* alleles described in this study and those previously reported in GenBank.

Gene	No. alleles (this study)	No. amino acid sequences (this study)	No. alleles with indels	No. alleles with nonsense mutations	Nucleotide diversity	*K*a/*K*s	*P* value
covR	54 (19)	32 (5)	1	2	0.0058	0.19	0.0032
covS	121 (43)	83 (26)	28[a]	11	0.0043	0.14	0.0029
ropB	78 (37[c])	56 (21[c])	7[b]	2	0.0055	0.23	0.0015

[a] Three of the *covS* indels are in frame.
[b] Two of the *ropB* indels are in frame, but one of these deletes a considerable portion of the protein (119 residues).
[c] Excluding the isolate with a complete *ropB* deletion.

Nucleotide diversity was also similar for the three genes, with differences in the total number of alleles basically reflecting the different sizes of the genes (687 bp for *covR*, 1503 bp for *covS*, and 843 bp for *rop*B). Null alleles were defined as those presenting changes predicted to result in a complete absence of protein function - including nonsense mutations, indels that result in frameshifts, large in-frame indels ($\geq$ 357 bp), and complete gene deletions. In spite of similar diversity, the *covS* gene presented more null alleles than *covR* ($P = 0.0004$).

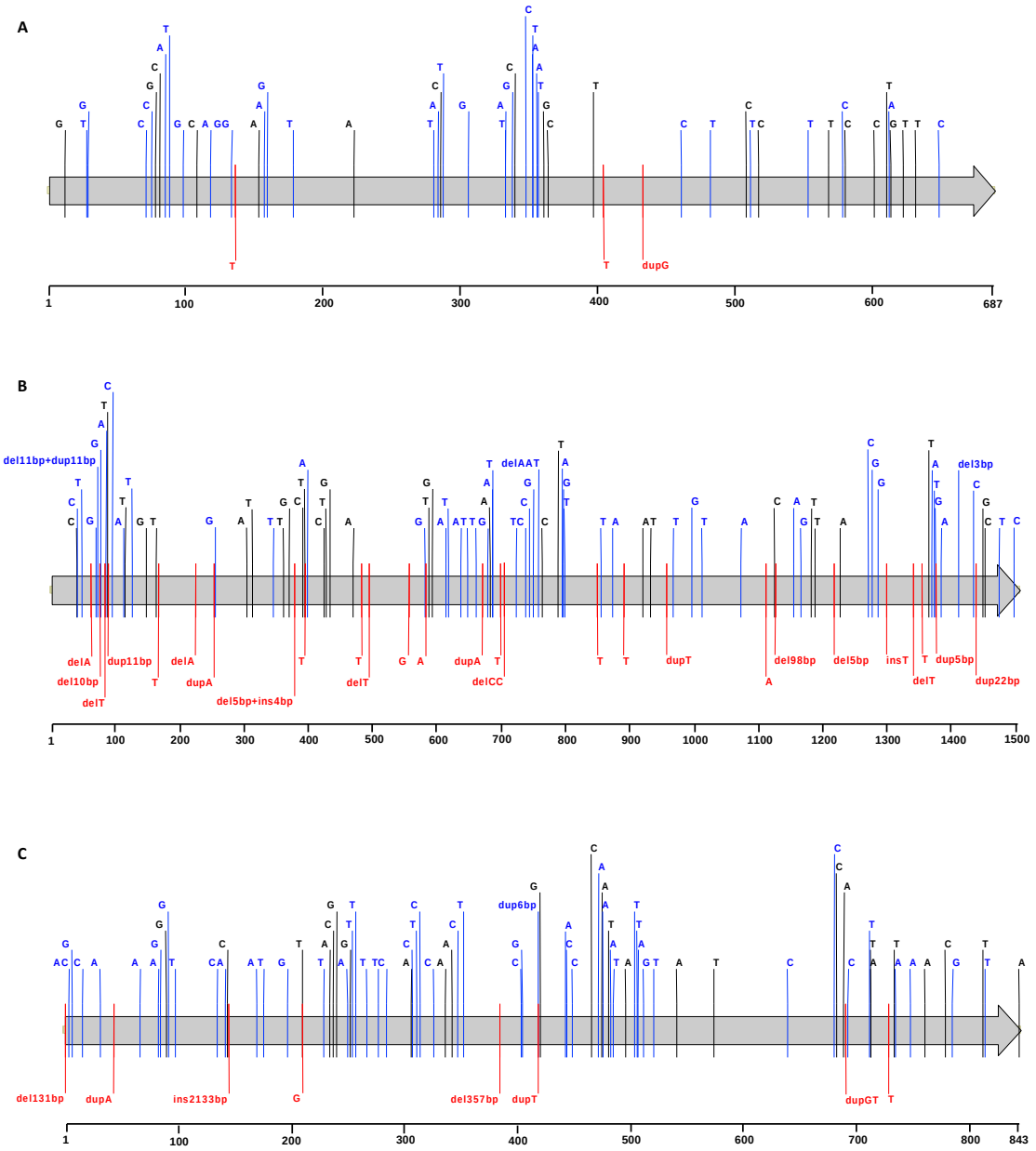


Figure 1. Nucleotide alterations identified in the covR (A), covS (B), and ropB (C) alleles reported in this study and previously deposited in GenBank, relative to the alleles present in strain SF370 (AE004092). Each gene is indicated by a grey arrow and a nucleotide numbering scale is represented below each one. Each sequence variation is indicated in the respective nucleotide position by a letter corresponding to the variant nucleotide, or the sequence or number of base-pairs inserted (ins), deleted (del), or duplicated (dup). Synonymous nucleotide changes are represented in black; missense mutations and short in-frame indels (≤ 6 bp) are represented in blue; changes predicted to result in null alleles, including nonsense mutations, indels that generate frameshifts, and long in-frame indels (≥ 357 bp) are represented in red.

Mutations in *covR*, *covS*, and *ropB*, and association with *emm* type and invasiveness.

For *covR* and *ropB*, the reference alleles (those present in SF370) were the most common among the 191 GAS isolates analysed in this study [*n* = 67 (35%), and *n* = 69 (36%), respectively]. The SF370 *covS* allele was present in only five isolates, but 68 isolates (36%) presented only synonymous nucleotide changes (15 alleles). Since our collection was enriched in *emm*1 isolates, the most common *covS* allele (*covS*-02, 28%) included a missense mutation characteristic of the M1T1 clone, namely I332V. Other missense mutations were also commonly identified in isolates of specific *emm* types, in both *covS* and *ropB* genes.

In addition to the coding sequence of the genes, the region upstream of *covR*, which includes the *covRS* promoter, was also analysed for all isolates, but no nucleotide changes were found in the identified -35 and -10 regions, in the transcription start site, nor in the consensus sequences that have been identified as binding targets for CovR auto-repression (Gusa and Scott, 2005).

The overall distribution of *emm* types differed significantly between the isolates carrying missense or null alleles of *covR*, *covS*, or *ropB* and those with no amino acid alterations ($P < 0.0001$). Null alleles were not associated with any particular *emm* type, occurring in isolates of seven different *emm* types for *covS* (SID = 0.909 [CI95% 0.822 -0.996]) and of four different *emm* types in the case of *ropB* (SID = 0.867 [CI95% 0.738–0.995]). Missense mutations in *covS* and *ropB* were significantly more prevalent in isolates of certain *emm* types due to the presence of clonal alleles in specific lineages, i.e. alleles carrying amino acid changes that occur in several isolates of that *emm* type (Supplementary Table S1).

According to the SID values, the allele diversity of the three genes was similar for the subsets of isolates associated with pharyngitis and with invasive infections (Table 2). The overall distribution of *covS* and *ropB* alleles was significantly different between the two subsets ($P < 0.0001$), but no specific *covR*, *covS*, or *ropB* allele presented a significant association with either infection type after correcting for the false discovery rate (FDR), except for *ropB*-23, which was significantly associated with invasive infections ($P = 0.0109$). This association is explained by the fact that *ropB*-23 was present in all *emm*64 isolates and not found in any other *emm* type (Supplementary Table S1), and *emm*64 was significantly associated with invasive infections in Portugal

(Friães *et al.*, 2012). This allele presents only two synonymous nucleotide changes, which are not likely to be the cause of the high invasiveness of this lineage.

The presence of an altered amino acid sequence in general was not associated with infection type for any of the genes. However, null *covS* alleles were significantly overrepresented among the invasive isolates (P = 0.0247) (Table 3), even though they were only present in nine of these isolates (10%).

Table 3. Number of isolates from invasive infections and pharyngitis presenting null and missense alleles in the *covR*, *covS*, and *ropB* genes.

gene	Allele type	Invasive (n = 87)	Pharyngitis (n = 104)
covR	Null	0	0
	Missense	2	2
covS	Null	9	2
	Missense	51	61
ropB	Null	3	3
	Missense	13	26
covRS/ropB	Null	12	5
	Missense	56	72

SpeB production.

In this study, the presence or absence of extracellular proteolytic activity was tested for all isolates using a spectrophotometric assay based on the degradation of azocasein by culture supernatants, as well as by a plate assay in which the strains were cultured in solid medium containing casein. To confirm that the proteolytic activity detected by these two methods was essentially due to SpeB, which has been described as the major extracellular protease of GAS (Chaussee *et al.*, 1993), Western blot analysis using monoclonal anti-SpeB antibodies was performed in a subset of six isolates of *emm* types 1, 64 and 89. For each *emm* type, one protease-positive and one protease-negative isolate were randomly chosen. The presence of a 28-kDa band similar to the one detected for SF370 was considered as a positive result (Fig. 2). The Western blot results were concordant with those of the two proteolytic activity assays for this subset of isolates. The *emm*1 strains SF370 and MGAS5005 were used as controls in all assays. As expected, strain SF370, which encodes a functional CovRS, presented proteolytic activity in both assays and was positive for mature SpeB production by

Western blotting. In contrast, almost undetectable amounts of mature SpeB and no proteolytic activity were observed for MGAS5005, which harbours a frameshift mutation in *covS* (Fig. 2 and Fig. 3). Taken together, these results confirm that SpeB is the major protease of GAS, allowing us to use the total proteolytic activity as a proxy for SpeB activity. Western blot was also performed for all isolates in which the azocasein and casein-plate results were discordant (n = 31). In these cases, the Western blot result was considered as final.

Of a total of 191 isolates, 153 (80%) presented SpeB activity, while for the remaining 38 (20%) no SpeB activity could be detected (Fig. 3 and Supplementary Table S1). All isolates with null alleles, either in *covS* or in *ropB*, were associated with an absence of SpeB activity (P < 0.0001 for both). In contrast, the prevalence of missense alleles among the SpeB-negative and SpeB-positive isolates was similar, except for non-clonal missense alleles in *ropB* (i.e. missense alleles that were not identified in other isolates of the same *emm* type in this study or in the GenBank database), which were associated with the absence of SpeB production (P = 0.0008). Only two SpeB-negative isolates presented alleles encoding the reference amino acid sequence in all three genes. The absence of SpeB activity in isolates carrying non-null alleles in the three genes could be due to alterations in the *speB* gene, its promoter, or in other mechanisms involved in the post-transcriptional or post-translational regulation of SpeB (Carroll and Musser, 2011).

There was no significant association between the SpeB activity of the isolates and infection type. The overall distribution of *emm* types differed significantly between SpeB-producing isolates and those without protease activity (P = 0.0030), but after FDR correction, only *emm* types 89 and 6 were significantly associated with the absence of SpeB (P = 0.0354 for both).

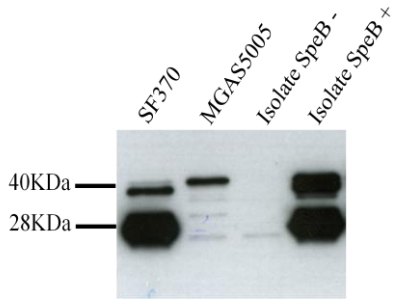


Figure 2. Representative Western blot result for detection of mature SpeB (28 kDa) expression by GAS isolates. The blot shown was cropped to the region of interest.

NADase activity.

In this work, the NADase activity of the isolates was measured by an endpoint titre method based on the degradation of β -NAD by culture supernatants, which results in a reduction of the fluorescence emitted by the reduced form of β -NAD. GAS *emm*1 strains SF370 and MGAS5005 were used as controls of the assay. As expected, SF370 presented a low NADase activity (NADase ≤ 3), while MGAS5005 was found to have the highest NADase activity level measured (NADase = 192) (Fig. 3). The high NADase activity of the latter strain can be attributed to the null *covS* allele, as well as to a different NADase locus from the one carried by old *emm*1 strains such as SF370 (Sumby *et al.*, 2005).

The gene encoding the GAS NADase, known as *spn* or *nga*, presents multiple variants and has been shown to be diverging into NADase-active and -inactive subtypes, which are correlated with *emm* patterns and tissue tropism (Riddle *et al.*, 2010). Therefore, variations in the levels of NADase activity exhibited by distinct GAS lineages, which may encode different *nga* alleles, are expected and were observed among the isolates analysed in this study (Fig. 3 and Supplementary Table S1). Despite this lineage-specific variation, it was possible to identify a significant association between the highest activity values, namely NADase = 96 and NADase = 192, and the presence of null alleles in *covS* (P = 0.0005), while *covR* and *ropB* changes did not significantly contribute to an increased NADase activity. In agreement, all the strains carrying *covS* null alleles expressed higher levels of NADase activity than the majority of the isolates of the same *emm* type, except for two *emm*64 isolates (Fig. 3). The absence of NADase activity could be an intrinsic characteristic of this lineage, regardless of the CovRS alleles, due to alterations in the *nga* gene, its promoter, or other

regulatory pathways. It is also possible that *nga* is not part of the CovRS regulon in the *emm*64 clone. Both inter- and intra-serotype differences in the regulatory activity of CovRS have been proposed (Sugareva *et al*., 2010).

Despite the association between high levels of NADase activity and *covS* null mutations, it was not possible to detect any significant association between the level of NADase activity and invasive disease.

Streptolysin S activity.

The SLS activity of the isolates was evaluated by an endpoint titre method based on the amount of released haemoglobin from sheep erythrocytes during incubation with culture supernatants. Although the majority of the isolates (*n* = 179, 94%) presented an SLS activity ≤ 3, it was possible to identify multiple isolates with increased SLS activity and that also carried alterations in CovRS (Fig. 3). Accordingly, the SLS activity determined for the reference strains SF370 (SLS ≤ 3) and MGAS5005 (SLS = 12) were consistent with a de-repression of the *sag* operon in the latter due to the null allele in *covS*, and a significant association between activity values of SLS = 12 and SLS = 48 and *covS* null alleles was observed (*P* = 0.0003 and 0.0040, respectively). In fact, all isolates with an SLS ≥ 12 presented null alleles in *covS* (Fig. 3 and Supplementary Table S1). Four isolates carrying null alleles and belonging to *emm* types 6, 44, and 64 did not exhibit an increased SLS activity. The coherent low NADase and SLS activity among *emm*64 isolates independent of the CovRS alleles is consistent with a potentially different CovRS regulon in these isolates. Among *emm*6 and *emm*44 isolates, it is possible that the *sag* operon, which encodes SLS activity, is not regulated by CovRS. Alternatively, SLS activity may not be significantly increased in these isolates due to mutations in the *sag* operon, which have been infrequently reported among *S. pyogenes* isolates (Yoshino *et al.*, 2010).

It was not possible to detect any significant associations of SLS activity levels with the presence or absence of changes in *covR* and *ropB*, nor with infection type.

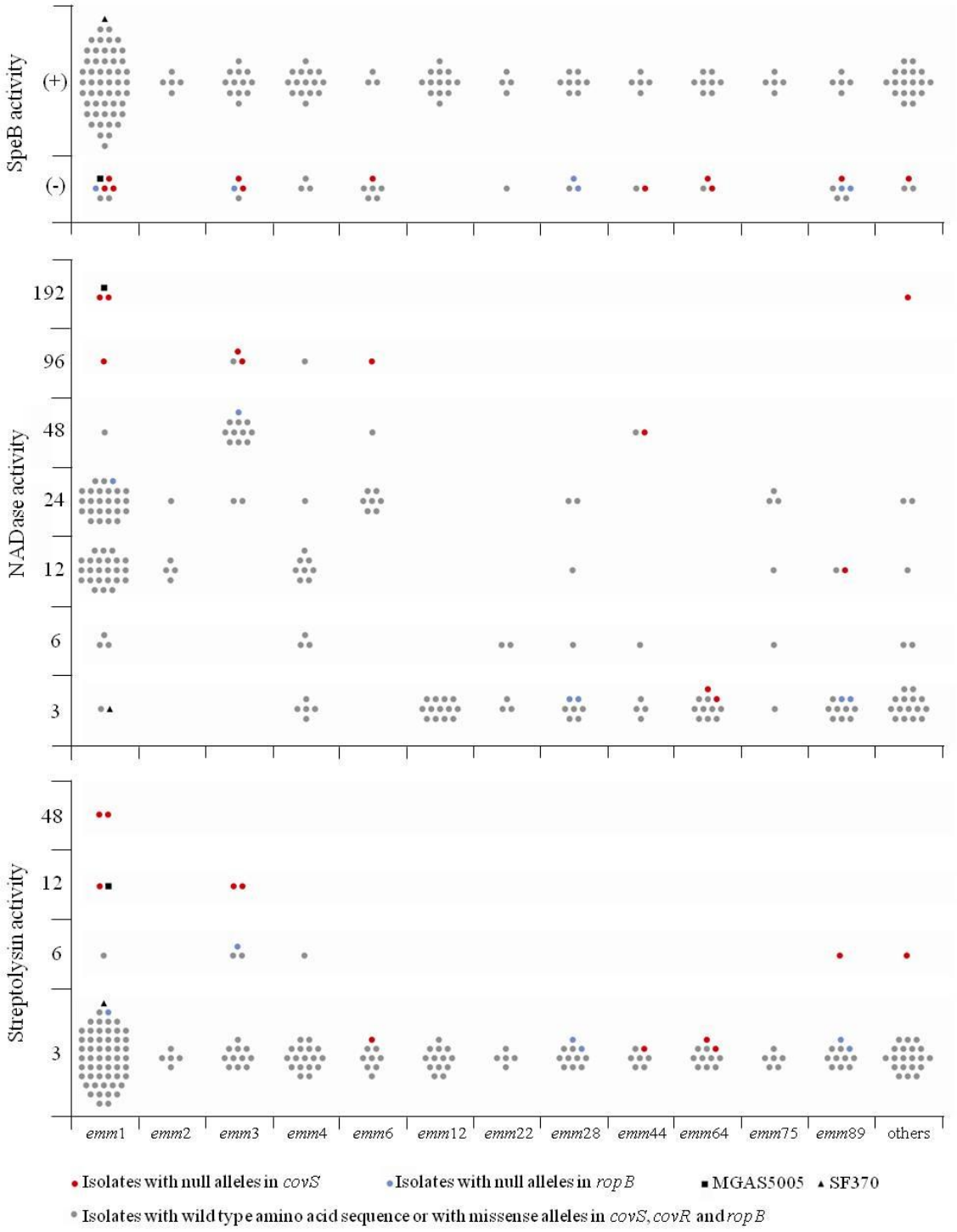


Figure 3. SpeB production and NADase and SLS activity determined for the 191 isolates analyzed in this study, according to *emm* type. The *emm* types represented by < 5 isolates are grouped into "Others" and include *emm* 9, 11, 18, 29, 43, 53, 58, 74, 77, 78, 87, 94, 102 and 113. The results of reference strains SF370 and MGAS5005 are also presented.

DISCUSSION

Variations at the level of the regulatory networks governing GAS gene expression may constitute key elements differentiating strains with a high invasive ability from those that carry the same virulence genes, but cause mild infections or asymptomatically colonise the host. Mutations impairing the function of the two-component system CovRS and the stand-alone regulator RopB could play this role and have been shown to be important in animal models of infection, mostly using representatives of the M1T1 clone (Engleberg *et al.*, 2001; Sumby *et al.*, 2006; Ikebe *et al.*, 2010; Mayfield *et al.*, 2014). However, the actual impact of these mechanisms in invasive human GAS infections remains unclear.

In the characterised collection, all three genes (*covR*, *covS*, and *ropB*) presented a high genetic diversity and the allele distribution was closely associated with *emm* type and ST. Some studies, mostly based on murine infection models, suggest that isolates of the highly invasive M1T1 clone have a higher ability to acquire *covRS* mutations than strains of other *emm* types (Mayfield *et al.*, 2014; Maamary *et al.*, 2010). However, in our study mutations in the *covR*, *covS*, and *ropB* genes were not more prevalent in *emm*1 or any other *emm* type. We conclude that among the GAS population causing human infections, the acquisition of mutations in any of the three genes can occur in isolates of diverse lineages, in agreement with data from Japan and Taiwan (Ikebe *et al.*, 2010; Lin *et al.*, 2014). We also did not find a higher prevalence of *covRS* mutations in isolates carrying the DNase gene *sda1* (data not shown), although it has been suggested that the presence of the phage encoding this gene exerts a selective pressure that favours the acquisition of *covRS* mutations, at least in the M1T1 genetic background (Walker *et al.*, 2007).

Null *covS* alleles were found to have a significant association with invasive infections, while the missense alleles in *covS* and all null and missense alleles in *covR* and *ropB* were not differently distributed among isolates from pharyngitis and from invasive disease. According to these results, only the null alleles in *covS* would contribute to the transition to invasive infection. In addition, the fact that null alleles are significantly more common in *covS* than in *covR* supports the hypothesis that during the infection process, the acquisition of mutations that impair signalling through CovS, while keeping CovR functional and possibly responsive to phosphorylation by other

kinases, such as SP-STK, may favour the progression to invasive infection (Trevino *et al.*, 2009; Agarwal *et al.*, 2011). However, this mechanism of transition to an invasive phenotype seems to be uncommon, since it was found in a minority of invasive isolates (10%).

The absence of expansion of isolates carrying null alleles in any of the three analysed genes and the fact that they are all under stabilising selection suggests that although mutations compromising the activity of these regulators may favour the progression to invasive disease, they are not beneficial to the overall fitness of *S. pyogenes*. Strains carrying *covRS* null alleles presented decreased ability to survive in human saliva and to persist in the murine nasopharynx, leading to the proposal that these strains would have an impaired colonisation and transmission capacity (Trevino *et al.*, 2009; Alam *et al.*, 2013). In a murine skin and soft tissue infection model, infection with a mixed population containing both the wild-type strain and a derivative *covRS* mutant resulted in higher virulence than infection with either strain alone, highlighting the importance of the presence of isolates carrying a functional CovRS (Engleberg *et al.*, 2001). Therefore, in spite of promoting a switch to a phenotype that favours the survival of GAS in deeper tissues, the loss of CovS may affect the success of the initial stages of infection (Aziz *et al.*, 2010).

In order to evaluate the phenotypic impact of mutations in the *covRS* and *ropB* genes, we determined the activity of three extracellular proteins that are known to be involved in GAS pathogenesis and to be under the direct or indirect influence of these regulators, namely SpeB, NADase, and SLS. A functional RopB is considered an absolute requirement for *speB* transcription (Neely *et al.*, 2003), while the influence of *covS* mutations in *speB* expression is more controversial (Engleberg *et al.*, 2001; Graham *et al.*, 2002; Hasegawa *et al.*, 2010; Aziz *et al.*, 2010; Mayfield *et al.*, 2014; Miller *et al.*, 2001; Trevino *et al.*, 2009; Maamary *et al.*, 2010). Our results support a role of both CovRS and RopB in the expression of SpeB, since null alleles in both *covS* and *ropB* were associated with the absence of SpeB production. Overall, the absence of SpeB activity was significantly more common among *emm* types 6 and 89, both frequently reported among invasive infections, especially *emm*89, which is one of the leading *emm* types among invasive isolates in several European countries (Luca-Harari *et al.*, 2009). However, the *emm* types that have been found to be significantly overrepresented among invasive infections in Portugal, namely *emm*1 and *emm*64 (Friães *et al.*, 2012), were not associated with an increased proportion of SpeB-negative

isolates. In agreement, the absence of SpeB activity was not associated with invasive infection, indicating that, in spite of preserving several virulence factors that are regarded as important for the invasive process, the abrogation of SpeB activity cannot explain by itself the higher ability of certain clones to cause invasive disease.

Several studies report a significant influence of RopB on the expression of virulence factors other than SpeB, including NADase and SLS, either directly or due to the regulation of SpeB or the interaction with other transcriptional regulators (Carroll *et al.*, 2011; Ikebe *et al.*, 2010; Hollands *et al.*, 2008). However, it has been demonstrated that the RopB regulon is highly variable, even among strains of the same *emm* type and that the core regulon is limited to the *speB* and *spi* genes (Dmitriev *et al.*, 2008). Our results are in agreement with this observation, since changes in *ropB* were not significantly associated with altered NADase or SLS activities. This suggests that, despite the influence that the regulator may have in the expression of these proteins in particular lineages, in the overall GAS population and under the growth conditions used in this study, RopB does not contribute significantly to the expression of the *nga* and *sag* operons. In contrast, *covS* null alleles were significantly associated with an increased activity of both NADase and SLS, which is consistent with a repression of the corresponding operons by phosphorylated CovR, as demonstrated for the *sag* promoter (Miller *et al*., 2001). However, exceptions were noted in isolates of specific *emm* types, including *emm*64, supporting the existence of differences in the CovRS regulon of distinct GAS lineages (Sugareva *et al*., 2010). Although both NADase and SLS have been shown to contribute to GAS cytotoxicity and virulence (Bricker *et al.*, 2002; Molloy *et al.*, 2011), their individual activities were not correlated with the type of infection caused by the isolates.

The association of missense alleles with invasiveness and with the activities of the studied virulence factors is harder to evaluate. Most of the alleles corresponding to amino acid changes in CovS and RopB occurred in multiple (often all) isolates of the same *emm* type, indicating that these are clonal mutations and probably do not represent the mechanism described for GAS in which mutations leading to a hypervirulent phenotype are acquired during infection (Engleberg *et al.*, 2001; Sumby *et al.*, 2006). Missense changes can have distinct effects on protein function depending on the nature of the amino acid replacement and the region where they occur. It is not possible to know from the current data if an altered phenotype in isolates carrying missense alleles is due to amino acid changes in these regulators or to some other characteristic of that

specific lineage. Possibly reflecting this, the prevalence of missense alleles in *ropB*, *covR*, and *covS* could not be associated with the phenotypes tested. When considering only the non-clonal missense alleles, an association with SpeB-negative isolates, but not with invasive disease, was observed in the case of *ropB*. Non-clonal missense alleles in *covRS* were not associated with any particular phenotype regarding the tested virulence factors, nor with disease presentation.

Null *covS* alleles are associated with invasive isolates and present phenotypes consistent with the ablation of this regulatory system. However, none of the phenotypes is itself associated with invasiveness, indicating that other factors under *covRS* regulation may be responsible for this higher virulence. In agreement, impairment of RopB function, known to result in SpeB downregulation, is not associated with invasiveness. The data presented supports the abrogation of CovS function as a mechanism contributing to the pathogenesis of invasive GAS infections, although this is not specifically associated with lineages identified as having enhanced invasiveness, but occurred in isolates of diverse *emm* types and STs. This mechanism is uncommon and the acquisition of such changes, despite being beneficial in the specific context of invasive disease, may incur an overall fitness cost for GAS, preventing their fixation in the population.

MATERIALS AND METHODS

Bacterial strains and culture conditions.

GAS isolates used in this study (104 recovered from pharyngeal exudates of patients with pharyngitis and 87 associated with invasive disease, recovered from normally sterile sites) are listed in Supplementary Table S1 online. The strains were randomly selected among a collection of 480 non-duplicate isolates recovered from human infections between 2000 and 2005 that have been previously characterized (Friães *et al.*, 2012), so as to include one third of the isolates of each *emm* type present in the collection with $n \geq 3$, as well as all isolates of *emm* types 1 and 64. Strains SF370 (CECT 5109, obtained from Colección Española de Cultivos Tipo) and MGAS5005 (BAA-947, obtained from American Type Culture Collection) were used as controls. Strains were grown at 37 °C in Todd Hewitt broth (THB) (BD, Sparks, MD, USA) or in Tryptone Soya Agar (Oxoid, Basingstoke, UK) supplemented with 5% defibrinated sheep blood.

Molecular typing.

In addition to the molecular characterisation of the isolates that had been previously performed (Friães *et al.*, 2012), multilocus sequence typing analysis was completed for all isolates.

Gene sequencing and analysis.

Genomic DNA of the isolates was extracted using cetyl trimethylammonium bromide (CTAB) (Ausubel, 1999). All PCR and sequencing primers used in this study are listed in Supplementary Table S3. The sequences obtained for each isolate were assembled and compared with the corresponding regions of the genome of strain SF370 (GenBank AE004092), considered to be the reference wild-type alleles.

Geneious R7 (Biomatters, Auckland, New Zealand) was used to search the GenBank database (accessed on April 7th 2014) using BLAST for all previously deposited sequences of *covR*, *covS*, and *ropB* of *S. pyogenes*, and align them with the trimmed sequences of the isolates analysed in this study, using the MUSCLE algorithm with default settings.

The subset of non-duplicate alleles without indels was used to calculate the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site (*K*a/*K*s), and the respective Z test for stabilizing selection, using MEGA5, with the Mega-Kumar method (Kimura 2-paramether). Values of $P < 0.05$ were considered statistically significant. Nucleotide diversity was calculated using DnaSP v.5.10.1.

Determination of protease activity.

Azocasein assay. SpeB protease activity in stationary phase GAS culture supernatants was determined using an adapted azocaseinolytic assay already modified for the 96-well plate format (Collin and Olsen, 2000; Cole *et al.*, 2010). Briefly, 24 h-cultures of each strain were diluted 1:10 in fresh THB and grown for 18 h, in 96-well plates. Bacteria-free supernatants were obtained by centrifugation at 3220 × g for 10 min, and transferred to a new microtitre plate. An equal volume of activation buffer (0.1 M sodium acetate [pH 5], 1 mM EDTA, 20 mM dithiothreitol) was added and the plate was incubated for 1 h at 40 °C. After activation, 2% azocasein (Sigma-Aldrich, St. Louis, MO, USA) (w/v in activation buffer) was added and incubated for 6 h at 40 °C. Samples were then precipitated with 2.5 volumes of 6% trichloroacetic acid and centrifuged at 15000 × g for 5 min. The optical density at 450 nm of the resulting supernatants was determined. The corresponding proteolytic activities were calculated using a calibration curve performed for each plate with known concentrations of proteinase K (Roche Diagnostics). Three independent assays were performed for each strain. The presence of proteolytic activity was considered positive when at least two of the three assays showed a value ≥0.0025 U.

Casein-plate assay. GAS expression of extracellular cysteine protease was determined by a plate assay as previously described (Ashbaugh et al., 1998). Briefly, single GAS colonies were stab-inoculated into plates of medium containing 0.5-strength Columbia broth (BD), 3% w/v skim milk (BD), and 1% w/v agar (Oxoid). Protease-expressing strains produced a translucent zone surrounding the site of inoculation after 24-h incubation at 37 °C. Three independent assays were performed for each strain. The presence of proteolytic activity was considered positive when at least two of the three assays presented a translucent zone of size similar to the one of strain SF370.

SpeB Western blot analysis.

For Western blot analysis, bacterial cultures were grown to late stationary phase (18 h) and centrifuged at 3200 × g for 15 min. Sterile-filtered supernatants were separated by electrophoresis on a 12% SDS-PAGE gel and transferred to a nitrocellulose membrane (Whatman, Dassel, Germany) using a Mini Trans-Blot (Bio-Rad, Hercules, CA, USA). Immunodetection was performed by chemiluminescence using monoclonal anti-SpeB antibody (Toxin Technology, Serasota, FL, USA) and goat anti-rabbit IgG conjugated with horseradish peroxidase (Santa Cruz Biotechnology, Dallas, TX, USA). Detection was performed with ECL detection reagents (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

Determination of NADase activity.

The NADase activity of each strain was determined by an endpoint titre method based on previously published assays (Bricker *et al.*, 2002). Briefly, bacteria-free supernatants of stationary-phase cultures of GAS strains were obtained as described for the azocasein assay and serially diluted (three-fold for the first dilution and two-fold for the second dilution onwards, up to 1/384) in β -NAD (Sigma-Aldrich) dissolved in PBS for a final concentration of 0.67 mM, in black 96-well plates (transparent bottom) (Greiner Bio-One, Frickenhausen, Germany). Wells containing sterile THB serially diluted in PBS and in β -NAD/PBS were used as controls, for each plate. After incubation for 1 h at 37°C in the dark, an equal volume of NaOH was added, for a final concentration of 2 N, and the plates were incubated for 1 h at room temperature, in the dark. Fluorescence was measured with excitation at 340 nm and emission detection at 460 nm. For each strain, the NADase activity was expressed as the inverse of the highest dilution prior to a greater or equal to two-fold increase in the fluorescence value. This method has a limit of detection of 3, and strains for which a two-fold increase was not observed and with values similar to the THB + β -NAD control were considered to have a NADase activity of ≤3. A minimum of three independent assays were performed for each strain, and the majority rule was used to determine the final NADase activity.

Determination of streptolysin (SLS) activity.

The streptolysin activity of each strain was determined by an endpoint titre method adapted from a previously published assay (Ravins *et al.*, 2000). Briefly, bacteria-free supernatants of stationary-phase cultures of GAS strains were obtained as described for the azocasein assay. The supernatants (or sterile THB for control of spontaneous haemolysis - blank) were serially diluted in PBS in microplate wells (three-fold for the first dilution and two-fold for the second dilution onwards, up to 1/192), and an equal volume of a 2.5% (v/v) suspension of defibrinated sheep erythrocytes was added. Two complete haemolysis controls were included in each plate, by incubating the erythrocytes suspension with 1% Triton X-100 (positive controls). The blank for these wells consisted only of PBS and the erythrocyte suspension. After incubation at 37 °C for 1 h, erythrocytes were pelleted and the absorbance at 570 nm of the supernatants was measured. For each well, the percentage of haemolysis relative to the positive control was calculated as follows:

$$\frac{Abs_{570}\,sample - Abs_{570}\,blank}{Abs_{570}\,positive\ control - Abs_{570}\,PBS\ blank} \times 100$$

The streptolysin activity was expressed as the inverse of the highest dilution prior to a ≥ two-fold decrease in the percentage of haemolysis. This method has a limit of detection of 3, and strains for which a two-fold decrease was not observed were considered to have a streptolysin activity of ≤ 3. A minimum of three independent assays were performed for each strain, and the majority rule was used to determine the final streptolysin activity. In four strains, two with functional *covRS* alleles (SF370 and SH0959A in Supplementary Table S1) and another two, of the same *emm* types, with null *covS* alleles and increased NADase activity (SH1025A and SH0421A in Supplementary Table S1), streptolysin assays were also performed in the presence of 33.3 μ g/ml of trypan blue (Sigma-Aldrich) or of 16.7 μ g/ ml cholesterol (Sigma-Aldrich). In all four strains, the results were not significantly changed by the presence of cholesterol, while trypan blue completely inhibited haemolysis, indicating that streptolysin O does not contribute to the haemolytic activity determined under these conditions, in agreement with previous reports (Ravins et al., 2000).

Statistical analysis.

The allelic diversity of the *covR*, *covS*, and *ropB* genes in the studied collection of GAS isolates was evaluated using the Simpson's index of diversity (SID) and corresponding 95% confidence intervals ($CI_{95\%}$) (Carrico *et al.*, 2006). The overall association between *covR*, *covS*, or *ropB* alleles and the *emm* type and ST was evaluated with the Adjusted Wallace coefficient with corresponding CI95% (Pinto *et al.*, 2008).

Unless otherwise specified, the statistical significance of pairwise associations was evaluated by calculating the respective odds ratios (when applicable) and the two-tailed Fisher's exact test, correcting the *P* values for multiple testing through the FDR linear procedure (Benjamini, 1995).

Additional Information:

Accession codes: All new *covRS* and *ropB* sequences identified in this study were deposited in GenBank (accession numbers KM985476 to KM985497 and KP101294 to KP101323).

SUPPLEMENTARY DATA

Table S1. Isolation source, molecular typing data, results of *covRS* and *ropB* sequencing, and results of SpeB, NADase, and SLS activity obtained for the 191 GAS isolates studied. http://dx.doi.org/10.6084/m9.figshare.1328335 (27/07/2018)

Table S2. Adjusted Wallace values (95% confidence intervals) of the *emm* type and the ST towards the *covR*, *covS*, and *ropB* alleles among the 191 GAS isolates analyzed in this study.

	covR	*covS*	*ropB*
emm type	0.978 (0.962-0.994)	0.834 (0.741-0.928)	0.806 (0.723-0.888)
ST	0.988 (0.974-1.000)	0.854 (0.753-0.955)	0.824 (0.735-0.913)

Table S3. Primers used for the amplification and sequencing of the *covRS* and *ropB* genes.

Primer Name	Sequence (5'→3')
covRS-1	TAACCTCGAAGAAAGTATTGTGG
covRS-4	CAACACGCTCAAAGGTAGTAAAG
covRS-5	GTTTGCCAGTCACTGAAAGG
covRS-6	GGATTTTCAGAGATATTAC
covRS-7	AATCAGTGTAAAGGCAGAG
ropB-F	GATAAAACTATCGCATCTGGC
ropB-R	CCTGGAGCTGTTGAGATAAAC

REFERENCES

Agarwal, S., P. Pancholi and V. Pancholi. 2011. Role of serine/threonine phosphatase (SP-STP) in *Streptococcus pyogenes* physiology and virulence. J Biol Chem. 286: 41368-41380.

Alam, F. M., C. E. Turner, K. Smith, S. Wiles and S. Sriskandan. 2013. Inactivation of the CovR/S Virulence Regulator Impairs Infection in an Improved Murine Model of *Streptococcus pyogenes* Naso-Pharyngeal Infection. PLoS One. 8: e61655.

Ashbaugh, C. D., H. B. Warren, V. J. Carey and M. R. Wessels. 1998. Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. J Clin Invest. 102: 550-560.

Ausubel, F. *et. al.* 1999. Greene Publishing Associates and Wiley-Interscience: J. Wiley. In Current Protocols (ed Chanda, V.).

Aziz, R. K., R. Kansal, B. J. Aronow, W. L. Taylor, S. L. Rowe, M. Kubal, G. S. Chhatwal, M. J. Walker and M. Kotb. 2010. Microevolution of group A streptococci in vivo: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. PLoS One. 5: e9798.

Benjamini, Y. and Y. Hochberg. 1995. Controling the False Discovery Rate: a pratical and powerful approach to multiple testing J. R. Stat. Soc. Ser. B Methodol. 57: 289-300.

Bricker, A. L., C. Cywes, C. D. Ashbaugh and M. R. Wessels. 2002. NAD+-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group A streptococci. Mol Microbiol. 44: 257-269.

Carrico, J. A., C. Silva-Costa, J. Melo-Cristino, F. R. Pinto, H. de Lencastre, J. S. Almeida and M. Ramirez. 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol. 44: 2524-2532.

Carroll, R. K. and J. M. Musser. 2011. From transcription to activation: how group A *Streptococcus*, the flesh-eating pathogen, regulates SpeB cysteine protease production. Mol Microbiol. 81: 588-601.

Carroll, R. K., S. A. Shelburne, 3rd, R. J. Olsen, B. Suber, P. Sahasrabhojane, M. Kumaraswami, S. B. Beres, P. R. Shea, A. R. Flores and J. M. Musser. 2011. Naturally occurring single amino acid replacements in a regulatory protein alter streptococcal gene expression and virulence in mice. J Clin Invest. 121: 1956-1968.

Chaussee, M. S., D. Gerlach, C. E. Yu and J. J. Ferretti. 1993. Inactivation of the streptococcal erythrogenic toxin B gene (speB) in *Streptococcus pyogenes*. Infect Immun. 61: 3719-3723.

Churchward, G. 2007. The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. Mol Microbiol. 64: 34-41.

Cole, J. N., M. A. Pence, M. von Kockritz-Blickwede, A. Hollands, R. L. Gallo, M. J. Walker and V. Nizet. 2010. M protein and hyaluronic acid capsule are essential for in vivo selection of covRS mutations characteristic of invasive serotype M1T1 group A *Streptococcus*. MBio. 1.

Collin, M. and A. Olsen. 2000. Generation of a mature streptococcal cysteine proteinase is dependent on cell wall-anchored M1 protein. Mol Microbiol. 36: 1306-1318.

Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 13: 470-511.

Dmitriev, A. V., E. J. McDowell and M. S. Chaussee. 2008. Inter- and intraserotypic variation in the *Streptococcus pyogenes* Rgg regulon. FEMS Microbiol Lett. 284: 43-51.

Engleberg, N. C., A. Heath, A. Miller, C. Rivera and V. J. DiRita. 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis. 183: 1043-1054.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2012. Group A streptococci clones associated with invasive infections and pharyngitis in Portugal present differences in emm types, superantigen gene content and antimicrobial resistance. BMC Microbiol. 12: 280.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2013. Superantigen gene complement of *Streptococcus pyogenes*-relationship with other typing methods and short-term stability. Eur J Clin Microbiol Infect Dis. 32: 115-125.

Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott and J. M. Musser. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. Proc Natl Acad Sci U S A. 99: 13855-13860.

Gusa, A. A. and J. R. Scott. 2005. The CovR response regulator of group A *Streptococcus* (GAS) acts directly to repress its own promoter. Mol Microbiol. 56: 1195-1207.

Hasegawa, T., A. Okamoto, T. Kamimura, I. Tatsuno, S. N. Hashikawa, M. Yabutani, M. Matsumoto, K. Yamada, M. Isaka, M. Minami and M. Ohta. 2010. Detection of invasive protein profile of *Streptococcus pyogenes* M1 isolates from pharyngitis patients. APMIS. 118: 167-178.

Hollands, A., R. K. Aziz, R. Kansal, M. Kotb, V. Nizet and M. J. Walker. 2008. A naturally occurring mutation in ropB suppresses SpeB expression and reduces M1T1 group A streptococcal systemic virulence. PLoS One. 3: e4102.

Horstmann, N., M. Saldana, P. Sahasrabhojane, H. Yao, X. Su, E. Thompson, A. Koller and S. A. Shelburne, 3rd. 2014. Dual-site phosphorylation of the control of

virulence regulator impacts group a streptococcal global gene expression and pathogenesis. PLoS Pathog. 10: e1004088.

Hurst, L. D. 2002. The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends Genet. 18: 486.

Ikebe, T., M. Ato, T. Matsumura, H. Hasegawa, T. Sata, K. Kobayashi and H. Watanabe. 2010. Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. PLoS Pathog. 6: e1000832.

Kansal, R. G., A. McGeer, D. E. Low, A. Norrby-Teglund and M. Kotb. 2000. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. Infect Immun. 68: 6362-6369.

Lin, J. N., L. L. Chang, C. H. Lai, H. H. Lin and Y. H. Chen. 2014. Association between polymorphisms in the csrRS two-component regulatory system and invasive group A streptococcal infection. Eur J Clin Microbiol Infect Dis. 33: 735-743.

Luca-Harari, B., J. Darenberg, S. Neal, T. Siljander, L. Strakova, A. Tanna, R. Creti, K. Ekelund, M. Koliou, P. T. Tassios, M. van der Linden, M. Straut, J. Vuopio-Varkila, A. Bouvet, A. Efstratiou, C. Schalen, B. Henriques-Normark and A. Jasir. 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 47: 1155-1165.

Maamary, P. G., M. L. Sanderson-Smith, R. K. Aziz, A. Hollands, J. N. Cole, F. C. McKay, J. D. McArthur, J. K. Kirk, A. J. Cork, R. J. Keefe, R. G. Kansal, H. Sun, W. L. Taylor, G. S. Chhatwal, D. Ginsburg, V. Nizet, M. Kotb and M. J. Walker. 2010. Parameters governing invasive disease propensity of non-M1 serotype group A streptococci. J Innate Immun. 2: 596-606.

Mayfield, J. A., Z. Liang, G. Agrahari, S. W. Lee, D. L. Donahue, V. A. Ploplis and F. J. Castellino. 2014. Mutations in the Control of Virulence Sensor Gene from *Streptococcus pyogenes* after Infection in Mice Lead to Clonal Bacterial Variants with Altered Gene Regulatory Activity and Virulence. PLoS One. 9: e100698.

Miller, A. A., N. C. Engleberg and V. J. DiRita. 2001. Repression of virulence genes by phosphorylation-dependent oligomerization of CsrR at target promoters in *S. pyogenes*. Mol Microbiol. 40: 976-990.

Molloy, E. M., P. D. Cotter, C. Hill, D. A. Mitchell and R. P. Ross. 2011. Streptolysin S-like virulence factors: the continuing *sagA*. Nat Rev Microbiol. 9: 670-681.

Neely, M. N., W. R. Lyon, D. L. Runft and M. Caparon. 2003. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. J Bacteriol. 185: 5166-5174.

Pinto, F. R., J. Melo-Cristino and M. Ramirez. 2008. A confidence interval for the wallace coefficient of concordance and its application to microbial typing methods. PLoS One. 3: e3696.

Ravins, M., J. Jaffe, E. Hanski, I. Shetzigovski, S. Natanson-Yaron and A. E. Moses. 2000. Characterization of a mouse-passaged, highly encapsulated variant of group A *Streptococcus* in in vitro and in vivo studies. J Infect Dis. 182: 1702-1711.

Riddle, D. J., D. E. Bessen and M. G. Caparon. 2010. Variation in *Streptococcus pyogenes* NAD+ glycohydrolase is associated with tissue tropism. J Bacteriol. 192: 3735-3746.

Shulman, S. T., R. R. Tanz, W. Kabat, K. Kabat, E. Cederlund, D. Patel, Z. Li, V. Sakota, J. B. Dale and B. Beall. 2004. Group A streptococcal pharyngitis serotype surveillance in North America, 2000-2002. Clin Infect Dis. 39: 325-332.

Sugareva, V., R. Arlt, T. Fiedler, C. Riani, A. Podbielski and B. Kreikemeyer. 2010. Serotype- and strain- dependent contribution of the sensor kinase CovS of the CovRS two-component system to *Streptococcus pyogenes* pathogenesis. BMC Microbiol. 10: 34.

Sumby, P., S. F. Porcella, A. G. Madrigal, K. D. Barbian, K. Virtaneva, S. M. Ricklefs, D. E. Sturdevant, M. R. Graham, J. Vuopio-Varkila, N. P. Hoe and J. M. Musser. 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 group A *Streptococcus* involved multiple horizontal gene transfer events. J Infect Dis. 192: 771-782.

Sumby, P., A. R. Whitney, E. A. Graviss, F. R. DeLeo and J. M. Musser. 2006. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog. 2: e5.

Trevino, J., N. Perez, E. Ramirez-Pena, Z. Liu, S. A. Shelburne, 3rd, J. M. Musser and P. Sumby. 2009. CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A *Streptococcus* virulence factor-encoding genes. Infect Immun. 77: 3141-3149.

Walker, M. J., A. Hollands, M. L. Sanderson-Smith, J. N. Cole, J. K. Kirk, A. Henningham, J. D. McArthur, K. Dinkla, R. K. Aziz, R. G. Kansal, A. J. Simpson, J. T. Buchanan, G. S. Chhatwal, M. Kotb and V. Nizet. 2007. DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat Med. 13: 981-985.

Yoshino, M., S. Y. Murayama, K. Sunaoshi, T. Wajima, M. Takahashi, J. Masaki, I. Kurokawa and K. Ubukata. 2010. Nonhemolytic *Streptococcus pyogenes* isolates that lack large regions of the sag operon mediating streptolysin S production. J Clin Microbiol. 48: 635-638.

CHAPTER III

STREPTOCOCCUS PYOGENES CAUSING SKIN AND SOFT TISSUE INFECTIONS HAVE INCREASED PREVALENCE OF THE RECENTLY EMERGED *EMM*89 CLADE 3 AND ARE NOT ASSOCIATED WITH ABROGATION OF COVRS

This chapter is a reproduction of the manuscript submitted to Frontiers in Microbiology in June of 2018.

The supplementary information referred throughout the text can be consulted in this chapter before the section of references.

I was involved in the conception and design of the study. I performed all the experimental work, although the majority of the phenotypic and molecular characterization of SSTI was already described in my Master Thesis. I was also responsible for the majority of statistical analysis. I wrote the first draft of the manuscript and I was involved during all the revisions.

Streptococcus pyogenes Causing Skin and Soft Tissue Infections Have Increased Prevalence of the Recently Emerged *emm*89 Clade 3 and are not Associated with abrogation of CovRS

Catarina Pato, José Melo-Cristino, Mario Ramirez, Ana Friães and the Portuguese Group for the Study of Streptococcal Infections

Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

SUMMARY

Although skin and soft tissue infections (SSTI) are the most common focal infections associated with invasive disease caused by *Streptococcus pyogenes* (Lancefield Group A streptococci - GAS), there is scarce information on the characteristics of isolates recovered from SSTI in temperate-climate regions. In this study, 320 GAS isolated from SSTI in Portugal were characterized by multiple typing methods and tested for antimicrobial susceptibility and SpeB activity, and compared to 313 contemporary isolates from invasive infections (iGAS). The *covRS* and *ropB* genes of isolates with no detectable SpeB activity were sequenced. The antimicrobial susceptibility profile was similar to that of iGAS, presenting a decreasing trend in macrolide resistance. However, the clonal composition of SSTI was significantly different from that of iGAS. Overall, iGAS were associated with *emm*1, *emm*3 and *emm*64, while SSTI were associated with *emm*89, the dominant *emm* type among SSTI (19%). Within *emm*89, SSTI were only significantly associated with isolates lacking the *hasABC* locus, suggesting that the recently emerged *emm*89 clade 3 may have an increased potential to cause SSTI. Reflecting these associations between *emm* type and disease presentation, there were also differences in the distribution of *emm* clusters, sequence types, and superantigen gene profiles between SSTI and iGAS. According to the predicted ability of each *emm* cluster to interact with host proteins, iGAS were associated with the ability to bind fibrinogen and albumin, whereas SSTI isolates were associated with the ability to bind C4BP and IgG. SpeB activity was absent in 79 isolates (25%), in line with the proportion previously observed among iGAS. Null *covS*

and *ropB* alleles (predicted to eliminate protein function) were detected in 10 (3%) and 12 (4%) isolates, respectively, corresponding to an underrepresentation of mutations impairing CovRS function in SSTI relative to iGAS. Overall, these results indicate that the isolates responsible for SSTI are genetically distinct from those recovered from normally sterile sites, supporting a role for mutations impairing CovRS activity specifically in invasive infection and suggesting that this role relies on a differential regulation of other virulence factors besides SpeB.

INTRODUCTION

Streptococcus pyogenes (group A streptococcus, GAS) is responsible for a variety of human infections ranging from mild and frequent diseases, such as pharyngitis and cutaneous infections, to more severe and rare invasive infections including sepsis, necrotizing fasciitis and streptococcal toxic shock syndrome (Walker *et al.*, 2014). In 2005, the estimated global incidence of invasive GAS infections (iGAS) was 663,000 new cases, resulting in 163,000 deaths, while at least 111 million children under 15 years suffered from pyoderma, mostly in developing countries, and pharyngitis incidence was estimated at over 616 million cases (Carapetis *et al.*, 2005). Even though mild infections are usually self-limited, they may play a crucial role in transmission. Furthermore, the nasopharyngeal mucosa and the skin can also be asymptomatically colonized representing primary reservoirs of GAS (Cunningham, 2000).

The gold-standard typing methodology of GAS is *emm* typing, which relies on the variability of the amino acid sequence of the N-terminal portion of *S. pyogenes* major virulence factor: the M protein (McMillan *et al.*, 2013). The sequence of the 5' variable region of the *emm* gene encoding the M protein determines the *emm* type, of which there are more than 250 distinct variants (https://www.cdc.gov/streplab/m-proteingene-typing.html). However, *emm* typing is based on the sequence of only approximately 10-15% of the complete *emm* gene. Recently a new classification was proposed based on *emm* clusters established by phylogenetic analysis of the entire sequence of the *emm* gene of 175 different *emm* types (Sanderson-Smith *et al.*, 2014). Each cluster contains isolates with closely related M proteins that share binding motifs to host proteins and other structural properties. Since isolates with the same *emm* type encode nearly identical M proteins, the *emm* cluster can be inferred from the *emm* type.

Associations between certain *emm* types and specific disease presentations have been established. Of particular importance is the association of iGAS with a contemporary *emm*1 clone, frequently designated as M1T1, which has persisted for decades as the major cause of invasive disease in most developed countries (Friães *et al.*, 2012; Luca-Harari *et al.*, 2009; O'Loughlin *et al.*, 2007; Aziz and Kotb, 2008). Recently, the emergence of a specific *emm*89 clade (clade 3) that rapidly outcompeted the previously circulating *emm*89 strains was reported in multiple countries and associated with an increase in the prevalence of *emm*89 among GAS infections (Turner

et al., 2015; Zhu *et al.*, 2015; Friães *et al.*, 2015a). Isolates from clade 3 are characterized by the absence of the *hasABC* locus encoding the hyaluronic acid capsule of GAS, and by a variant *nga-ifs-slo* locus, similar to the one present in M1T1 strains, which is associated with increased expression of NAD-glycohydrolase (NADase) and streptolysin O (SLO) (Turner *et al.*, 2015; Zhu *et al.*, 2015).

Despite the success of *emm* typing, some studies have suggested it is not enough to identify GAS clones and that it must be complemented with other typing methods such as multilocus sequence typing (MLST), superantigen (SAg) gene profiling (Carrico *et al.*, 2006; Friães *et al.*, 2013b), and, more recently, whole genome sequencing (Carrico *et al.*, 2013). However, variability in key virulence factors and regulators within clones defined by these typing methods may have important consequences for the virulence of a particular isolate.

A mouse model of skin and soft tissue infection (SSTI) showed that mutations in the *covRS* two component system were a key step for the transition from a localized to a systemic infection (Sumby *et al.*, 2006). Consistently, several studies reported *covRS* mutations in isolates recovered from human infections (Engleberg *et al.*, 2001; Hasegawa *et al.*, 2010; Ikebe *et al.*, 2010; Lin *et al.*, 2014; Friães *et al.*, 2015b). In these isolates the downregulation of SpeB expression, a potent extracellular cysteine protease, is thought to be fundamental towards the switch to a hipervirulent phenotype (Aziz *et al.*, 2004; Kansal *et al.*, 2010). Transcription of *speB* is also under direct control of RopB (Carroll and Musser, 2011), and naturally occurring mutations in *ropB* were also shown to impair SpeB production (Hollands *et al.*, 2008; Carroll *et al.*, 2011). We reported previously that mutations resulting in the truncation of CovS, which presumably impaired its function, were significantly overrepresented among iGAS in Portugal. However these were only present in 10% of invasive isolates, not explaining why most isolates caused invasive infections. Additionally, among all studied isolates, which included invasive and pharyngeal isolates, 20% had no detectable SpeB activity but no significant association was detected between the presence or absence of SpeB activity and the type of infection (Friães *et al.*, 2015b).

In order to cause a wide spectrum of disease, GAS has to be able to adapt to different environments in the host and despite decades of research there is still no consensus regarding which molecular or phenotypic properties are responsible for an enhanced invasive potential of certain lineages. Although the nasopharyngeal mucosa is usually considered as the main source of isolates causing iGAS in developed countries

(Fiorentino *et al.*, 1997), SSTI are commonly reported as the predominant foci associated with invasive disease (Lamagni *et al.*, 2008). This raises the possibility that strains adapted to infect the skin have an increased ability to invade and survive in deeper tissues. However, there is scarce data about the characteristics of GAS isolates responsible for SSTI, especially in developed, temperate climate regions. Most of the studies from these regions report the characteristics of SSTI isolates together with GAS from other non-invasive sites (mostly pharyngeal swabs) when comparing invasive and non-invasive disease (Rivera *et al.*, 2006; Ekelund *et al.*, 2005; Descheemaeker *et al.*, 2000). A few others specify the molecular characteristics of the SSTI isolates subset, but are limited in the number of isolates or are restricted to short time periods (Tamayo *et al.*, 2014; Vahakuopus *et al.*, 2012; Mijac *et al.*, 2010; Kittang *et al.*, 2008). In this study we characterized 320 isolates from SSTI recovered in Portugal during 2003-2009 for their susceptibility to a panel of antimicrobials, *emm* type, SAg profile, and MLST. The genes conferring resistance to selected antimicrobials were also investigated. In addition, all 320 isolates were tested for SpeB activity and in those without detectable activity we sequenced the *covRS* and *ropB* genes to document any mutations. The SSTI isolates presented substantial differences relative to a collection of iGAS isolates recovered in the same period in Portugal regarding both the clonal composition and the prevalence of mutations impairing CovRS function.

METHODS

Bacterial isolates

For this study, 24 hospital laboratories distributed throughout Portugal were asked to send us, on a voluntary basis, all GAS isolated from SSTI between January 2003 and December 2009. The study was approved by the Institutional Review Board of the Centro Académico de Medicina de Lisboa. These were considered surveillance activities and were exempt from informed consent. All methods were performed in accordance with the relevant guidelines and regulations. The data and isolates were de-identified so that these were irretrievably unlinked to an identifiable person. Participation was low in the first two years (n = 17 in 2003, n = 8 in 2004), but subsequently increased (average n = 59 / year, range 48-77) (Dataset available at 10.6084/m9.figshare.6736313). Overall, 320 non-duplicate GAS isolates from SSTI were recovered and included in the study: 306 from skin and soft tissue exudates (pus) and 14 from skin and soft tissue biopsies. Identification of isolates was performed by colony morphology, β-hemolysis on blood agar, and the presence of the characteristic Lancefield group A antigen (OXOID, Basingstoke, UK). In addition, 313 non-duplicate iGAS isolates recovered in the same hospitals and time period, which had been partially characterized previously (Friães *et al.*, 2007; Friães *et al.*, 2013a), were also included. Strain SF370 was obtained from Colección Española de Cultivos Tipo (CECT5109). Strains were grown at 37ºC in Todd Hewitt broth (THB) (BD, Sparks, MD, USA) or in Tryptone Soya Agar (Oxoid, Basingstoke, UK) supplemented with 5% defibrinated sheep blood.

Antimicrobial susceptibility testing and genetic determinants

Susceptibility tests were performed by disk diffusion according to the guidelines of the Clinical and Laboratory Standards Institute (CLSI) for penicillin, vancomycin, erythromycin, levofloxacin, tetracycline, chloramphenicol, clindamycin and linezolid (Oxoid, Basingstoke, UK). E-test strips (BioMérieux, Marcy l'Etoile, France) were used for MIC determination in cases of intermediate susceptibility and to confirm resistance when ≤ 5 isolates were resistant to a particular antimicrobial. Determination of macrolide resistant phenotype was performed as previously described (Melo-Cristino and Fernandes, 1999). A multiplex PCR reaction for *erm*(B), *erm*(A) and *mef* genes was

used on macrolide resistant isolates to identify the resistance conferring genes (Figueira-Coelho *et al.*, 2004). The *mef* positive isolates were further analyzed in order to distinguish between *mef*(A) and *mef*(E) (Silva-Costa *et al.*, 2008). The tetracycline resistance genotype was determined for resistant isolates by a multiplex PCR for the genes *tet*(K), *tet*(L), *tet*(M) and *tet*(O) (Trzcinski *et al.*, 2000).

.

Molecular typing

The *emm* typing was performed according to the protocols and recommendations of the Center for Disease Control and Prevention (CDC) (https://www.cdc.gov/streplab/protocol-emm-type.html).

The presence of 11 SAg genes (*speA, speC, speG, speH, speI, speJ, speK, speL, speM, smeZ* and *ssa*) was tested by two multiplex PCR reactions using the amplification of *speB* and *speF* gene fragments as positive controls (Friães *et al.*, 2013b). Absence of a PCR product for *speB* and *speF* was confirmed by Southern blotting as previously described (Friães *et al.*, 2013b).

MLST analysis was performed for all SSTI isolates (n = 320), as well as for the iGAS isolates that had not been previously typed by MLST (n = 152/313) (Friães *et al.*, 2012; Friães *et al.*, 2007; Friães *et al.*, 2013a). Allele and sequence type (ST) identification was done using the *S. pyogenes* MLST database (https://pubmlst.org/spyogenes/).

The sequence of the *covRS* and *ropB* loci of all isolates that presented no proteolytic activity (see below) was determined as previously described (Friães *et al.*, 2015b). For each isolate, the sequences were assembled and compared with the corresponding regions of the genome of strain SF370 (GenBankAE004092), considered as the reference wild-type alleles. Isolates were considered to carry null alleles if the changes found were predicted to result in absence of a functional protein due to nonsense mutations or frameshifts. All new *covRS* and *ropB* sequences identified in this study were deposited in GenBank (accession numbers MH537795-MH537849).

Determination of proteolytic activity and SpeB expression

All isolates were screened for detectable SpeB activity using a plate assay, as previously described (Friães et al., 2015b). Briefly, single GAS colonies were stab-inoculated into fresh plates of medium containing 0.5-strength Columbia broth (BD),

3% w/v skim milk (BD, Sparks, MD, USA), and 1% w/v agar (Oxoid). A strain was considered to show proteolytic activity when it presented a translucent zone of size similar to the one of strain SF370 after 24h incubation at 37ºC, in three independent assays. On the contrary, a strain was considered not to show proteolytic activity if it did not produce a translucent halo in any of the three assays. For strains in which the results of the three assays were inconclusive or not consistent, detection of SpeB by Western blot was performed as described previously (Friães *et al.*, 2015b) and this was considered the final result.

Statistical analysis

Simpson's index of Diversity (SID) with respective 95% confidence intervals ($CI_{95\%}$) was used for the analysis of the typing methodologies and to evaluate the allelic diversity of the *covR*, *covS* and *ropB* genes found in this study (Carrico et al., 2006). Two–tailed Fisher's exact test and odds ratios were used to identify significant pairwise associations. P-values for multiple tests were corrected using the FDR linear procedure (Benjamini, 1995). The Cochran-Armitage test was used to evaluate trends. A *P*-value < 0.05 was considered significant for all tests.

RESULTS

Antimicrobial susceptibility and genotypic determinants of resistance

All 320 SSTI isolates analyzed (Dataset) were susceptible to penicillin, vancomycin, and linezolid. Two isolates were non-susceptible to levofloxacin (MIC = 6 mg/mL and 3 mg/mL, both *emm*89). One was resistant to chloramphenicol (*emm*147).

Macrolide resistance was detected in 33 SSTI isolates (10%) of 12 different *emm* types (Dataset), with a significant decreasing trend during the years of the study ($p < 0.001$) (Supplementary Figure 1). The majority of these isolates presented the $cMLS_B$ phenotype and carried the *erm*(B) gene ($n = 22$). The remaining macrolide resistant isolates ($n = 11$) exhibited the M phenotype and carried the *mef*(A) gene, with the exception of one isolate that was positive for the *mef*(E) gene.

Tetracycline resistance was detected in 47 SSTI isolates (15%) comprising 25 *emm* types (Dataset) with no significant temporal trend (Supplementary Figure 1). The majority of these isolates ($n = 42$) presented only the *tet*(M) gene, while three isolates carried both *tet*(M) and *tet*(L) and two isolates presented solely the *tet*(O) gene. Resistance to tetracycline and macrolides was simultaneously detected in 12 isolates, of which 9 presented the $cMLS_B$ phenotype.

emm typing

A total of 51 different *emm* types were identified among SSTI isolates (Table 1). Isolates with *emm*89 were the most frequently recovered ($n = 62$, 19%), followed by *emm*1 ($n = 55$, 17%), together accounting for 36% of all SSTI isolates (Figure 1). Considering *emm* types with ≥ 10 isolates, no significant changes in time were detected during 2003-2009 after FDR correction (data not shown).

The *emm* types were distributed into 15 *emm* clusters (Table 1 and Figure 1). The majority of SSTI isolates belong to *emm* clusters from clade X ($n = 189$, 59%). No significant changes in prevalence of any cluster were detected in the study years (data not shown).

Table 1. Simpson's Index of Diversity (SID) and respective 95% confidence intervals for *emm* type, *emm* cluster, superantigen (SAg) profile, and sequence type (ST) among skin and soft tissue infections (SSTI) and invasive infections (iGAS) in Portugal during 2003-2009.

		***emm* type**	***emm* cluster**	**SAg profile**	**ST**
SSTI (*n* = 320)	No. partitions	51	15[a]	50	70
	SID ($CI_{95\%}$)	0.916 (0.899-0.933)	0.838 (0.813-0.862)	0.934 (0.922-0.946)	0.945 (0.933-0.958)
iGAS (*n* = 313)	No. partitions	37	16[a]	37	52
	SID ($CI_{95\%}$)	0.891 (0.870-0.912)	0.852 (0.832-0.873)	0.908 (0.891-0.926)	0.911 (0.890-0.933)

[a] Isolates with no *emm* cluster (SSTI: *n* = 3, *emm* type 127, 147 and 167; iGAS: n = 3, *emm* types 196, 199 and stG1750) were not considered for SID determination.

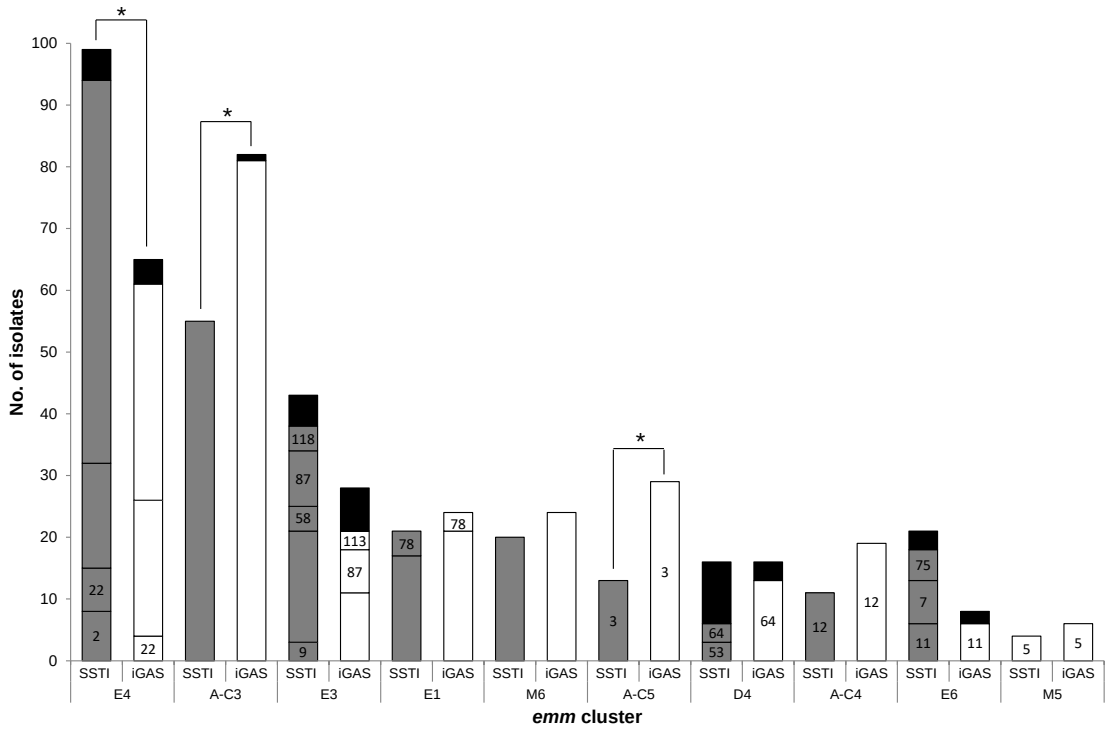


Figure 1. Distribution of isolates recovered from skin and soft tissue infections (SSTI) and invasive infections (iGAS) according to *emm* cluster and *emm* type. Only *emm* clusters with a total of ≥ 10 isolates in both infection types are represented. Numbers inside the bars represent the *emm* types included in each cluster. Black bars include *emm* types with < 3 isolates [SSTI, E4: *emm*77 (*n* = 2), *emm*73, *emm*102, and *emm*109 (each *n* = 1); E3: *emm*209 (*n* = 2), *emm*82, *emm*103, and *emm*168 (each *n* = 1); D4: *emm*70, and *emm*223 (each *n* = 2), *emm*33, *emm*43, *emm*80, *emm*83, *emm*93, and *emm*225 (each *n* = 1); E6: *emm*65, *emm*81, and *emm*94 (each *n* = 1). iGAS, E4: *emm*77 (*n* = 2), *emm*2, and *emm*84 (each *n* = 1); A-C3: *emm*227 (*n* = 1); E3: *emm*9, *emm*58, and *emm*103 (each *n* = 2), *emm*118 (*n* = 1); D4: *emm*53 (*n* = 2), *emm*43 (*n* = 1); E6: *emm*75, and *emm*81 (each *n* = 1)]. *p ≤ 0.01.

When comparing with iGAS isolates recovered in the same period, significant differences were detected in the overall distribution of *emm* types (p = 0.007), as well as in the prevalence of specific *emm* types (Supplementary Figure 2), although the SIDs were not significantly different (Table 1, p = 0.064). Twenty-three *emm* types accounting for 27 isolates were identified exclusively among SSTI, while seven *emm* types accounting for nine isolates were found only in iGAS. While *emm*1, *emm*3 and *emm*64 were significantly overrepresented among iGAS (p = 0.009, p = 0.010 and p = 0.011, respectively), *emm*89 was significantly overrepresented among SSTI (p = 0.006). However, when stratifying the *emm*89 isolates as to the presence or absence of the *hasABC* locus, only the isolates lacking the locus were associated with SSTI (p = 0.006). All these differences were still significant after FDR correction.

These differences of *emm* types were reflected in the distribution of *emm* clusters among SSTI and iGAS isolates (Figure 1). While iGAS isolates were significantly associated with *emm* clusters A-C3 (comprising almost exclusively *emm*1 isolates) and A-C5 (comprising exclusively *emm*3 isolates) (p = 0.007 and p = 0.010, respectively), SSTI isolates were significantly associated with *emm* cluster E4 (dominated by *emm*89) (p = 0.004).

The *emm* cluster typing system can be used for predicting the ability of the strains to bind different host proteins based on the binding properties of the respective M proteins (Supplementary Table 1) (Sanderson-Smith *et al.*, 2014). When comparing SSTI and iGAS isolates, the ability to bind C4BP and IgG was associated with SSTI isolates (p < 0.001 and p = 0.008, respectively), whereas the invasive isolates were associated with the ability to bind fibrinogen and albumin (both p < 0.001), all significant after FDR.

MLST

The studied SSTI isolates comprised 70 different STs, presenting a higher diversity than the invasive isolates recovered in the same period (p = 0.006) (Table 1 and Table 2). A total of 20 new STs were found among the SSTI and iGAS isolates typed in this study (ST817-ST836), including two new alleles for *gtr* (108; 109), three for *murI* (105-107), two for *mutS* (88; 89), two for *recP* (125-126), two for *xpt* (102; 103), and three for *yqil* (119-121).

Table 2. Distribution of sequence types (ST) among *emm* types identified in skin and soft tissue infections (SSTI) and invasive infections (iGAS) in Portugal during 2003-2009.

***emm* type**	**ST (no. of isolates)**	
	SSTI (*n* = 320)	**iGAS (*n* = 313)**
1	28(52); 643(1); 830(2)	28(77); 618(1); 643(3)
89	101(37); 408(15); 824(10)	101(18); 408(14); 824(3)
6	382(13); 411(7)	382(15); 411(9)
3	15(8); 315(5)	15(19); 315(4); 406(6)
28	52(17)	52(20); 458(1); 821(1)
4	38(2); 39(15)	39(18); 771(1); 823(2)
12	36(10); 467(1)	36(18); 551(1)
44	25(13); 178(2); 429(1); 555(2)	25(8); 178(1); 555(2)
64	164(3)	124(1); 164(12)
87	62(9)	62(7)
11	403(6)	403(5); 562(1)
22	46(6); 389(1)	46(4)
5	99(4)	99(6)
Others[a]	2(1); 3(2); 5(1);10 (1); 24(2); 50(2); 55(8); 60(1); 63(2); 75(3); 89(1);120(1); 130(3); 150(3); 161(7); 164 (1);166 (1); 167(1); 184(1); 200(1); 253(3); 331(2); 340(2); 341(1); 409(1); 565 (2); 569(1); 573(1); 642(3); 701(1); 718(1); 754 (1); 817(1); 819 (1); 820 (1); 822(1); 826(3); 825(1); 827(1); 828(1); 829(1); 831(1); 833(1); 834(1); 835(1); 836(1)	3(1); 11(2); 28(2); 50(1); 53(1); 55(1); 63(2); 75 (2); 95(1); 99(1); 120(1); 150(1); 184(1); 201(2); 258(1); 402(1); 409 (3); 410(2); 619(1); 679(3); 769(1); 816 (1); 818(1), 832(1); 833(1)

[a] **"Others"** include *emm*-types with n<10 isolates considering both SSTI and invasive isolates. Those include *emm*-types: 2; 9; 18; 19; 33; 43; 48; 50; 53; 58; 65; 68 70; 71; 73; 74; 75; 76; 77; 78; 80; 81; 82; 83; 84; 90; 93; 94; 102; 103; 109; 113; 118; 209; 223 127; 147; 164; 167; 168; 179; 196; 199; 225; 227 and stG1750.

The overall distribution of STs differed significantly between SSTI and iGAS ($p < 0.001$). Some STs were associated with certain infection types, in agreement with the associations of the corresponding dominant *emm* types with infection type (*emm*1 for ST28, *emm*3 for ST15, *emm*64 for ST164, and *emm*89 for ST101). However, these differences were not statistically supported after FDR correction (Supplementary Figure 3).

No significant differences were detected when comparing the ST diversity between isolates with the same *emm* type recovered from each type of infection (Table 2).

SAg profiling

Overall, chromosomally encoded *smeZ* and *speG* genes were the most frequently detected among SSTI isolates, being present in 307 (96%) and 301 (94%) isolates, respectively (Table 3 and Supplementary Figure 4(A)). PCR-amplification of *speB* and *speF* was not possible for one isolate, and the absence of both genes was confirmed by Southern blot (data not shown). The presence of the genes *speA* and *speK* was significantly associated with invasive isolates ($p = 0.001$ and 0.017, respectively), while *speL* and *speM* were significantly associated with isolates from SSTI ($p = 0.002$ and $p < 0.001$, respectively) (Supplementary Figure 4(A)).

A total of 50 distinct SAg profiles were identified among the SSTI isolates (Table 1 and Table 3). The SAg profiles of SSTI isolates were more diverse than those of invasive isolates ($p = 0.016$) (Table 1). Reflecting the association between *emm* type and infection, SAg profiles 8 and 10, which correspond mostly to *emm*3 and *emm*1, respectively, were associated with iGAS ($p = 0.005$ and 0.003, respectively), while SAg profile 29, mostly found among *emm*89, was associated with SSTI ($p = 0.010$) (Supplementary Figure 4(B)).

Table 3. Superantigen (SAg) profiles identified in isolates recovered from skin and soft tissue infections (SSTI) in Portugal during 2003-2009.

SAg profile[a]	*speA*	*speC*	*speG*	*speH*	*speI*	*speJ*	*speK*	*speL*	*speM*	*ssa*	*smeZ*	*emm* type (No. of isolates)
2	+	+	+	-	-	-	+	-	-	-	+	6(17); 74(1)
3	+	+	+	-	-	+	-	-	-	-	+	1(11); 87(1)
4	+	+	+	-	-	-	-	+	+	-	+	18(3); 22(1)
5	+	+	+	-	-	-	-	-	-	-	+	5(2); 11(1); 18(1); 122(1)
8	+	-	+	-	-	-	+	-	-	+	+	3(11)
9	+	-	+	-	-	-	+	-	-	-	+	6(1)
10	+	-	+	-	-	+	-	-	-	-	+	1(43); 71(1)
11	+	-	+	-	-	-	-	-	-	-	+	43(1); 82(1)
12	-	+	+	+	+	+	-	-	-	+	+	44(2)
13	-	+	+	+	+	-	-	-	-	+	+	22(1)
15	-	+	+	+	-	+	+	-	-	-	+	28(1)
16	-	+	+	+	+	-	-	-	-	-	+	11(4); 16(6); 48(4)
18	-	+	+	-	-	+	+	-	-	+	+	87(1)
19	-	+	+	-	-	-	+	-	-	+	+	22(2)
20	-	+	+	-	-	+	-	-	-	+	+	84(4)
21	-	+	+	-	-	-	-	-	-	+	+	22(2)
23	-	+	-	-	-	-	-	-	-	+	+	4(16)
24	-	+	+	-	-	+	+	-	-	-	+	28(8)
26	-	+	+	-	-	-	+	-	-	-	+	89(3)
27	-	+	+	-	-	+	-	-	-	-	+	28(8); 68(1); 70(2); 89(15); 90(1)
28	-	+	+	-	-	-	-	+	+	-	+	75(1)
29	-	+	+	-	-	-	-	-	-	-	+	5(2); 11(1); 48(3); 50(1); 77(1); 78(4); 89(35); 118(1); 127(1)
31	-	+	+	-	-	-	-	+	+	-	-	2(7)

32	-	-	+	+	+	+	-	-	-	+	+	44(12)
33	-	-	+	+	+	-	-	-	-	-	+	12(5); 73(1); 76(2)
35	-	-	+	+	-	-	-	-	-	-	+	75(1); 94(1)
38	-	-	+	-	-	+	-	-	-	+	+	44(1);87(1); 223(1)
39	-	-	+	-	-	-	-	+	+	+	+	53(1); 75(2)
40	-	-	+	-	-	-	-	-	-	+	+	90(3); 22(1)
41	-	-	-	-	-	-	-	-	-	+	+	4(1)
43	-	-	+	-	-	-	+	-	-	-	+	65(1); 209(1)
44	-	-	+	-	-	+	-	-	-	-	+	102(1); 103(1); 118(2); 223(1)
45	-	-	+	-	-	-	-	+	+	-	+	80(1)
46	-	-	+	-	-	-	-	-	-	-	+	53(2); 58(1); 64(3); 75(1); 89(9); 118(1); 147(1); 167(1); 168(1)
47	-	-	-	-	-	-	-	-	-	-	+	77(1)
48	-	-	+	-	-	-	-	+	+	-	-	2(1)
51	+	+	+	+	+	-	+	-	-	-	+	6(2)
52	+	+	+	-	-	-	+	-	-	+	+	22(1)
53	+	-	+	-	-	-	-	-	-	+	+	3(2)
54	-	+	+	+	+	+	-	-	-	-	+	87(1)
56	-	-	+	+	-	+	-	-	-	-	-	44(3)
60	+	-	+	+	+	+	-	-	-	-	+	71(2)
61	+	-	+	-	-	+	+	-	-	-	+	1(1)
64	-	+	+	+	+	-	-	+	+	-	+	93(1)
66	-	+	+	-	-	-	-	-	+	-	+	58(3)
67	-	-	+	+	-	+	-	-	-	+	+	33(1)
68	-	-	+	+	+	-	-	-	-	-	-	209(1)
69	-	-	+	+	-	-	-	-	-	-	-	81(1)
70	-	-	+	-	-	+	+	-	-	-	+	87(1)
71	-	+	+	-	-	+	-	+	+	-	+	83(1)

[a] The numbering of the SAg profiles follows the one adapted previously (Friães *et al.*, 2013b; Friães *et al.*, 2013a; Silva-Costa *et al.*, 2014)

SpeB protease activity and sequence of *covRS* and *ropB*

We previously showed that the total proteolytic activity could be used as a proxy for SpeB activity and that all isolates with null alleles, either in *covS* or in *ropB*, lacked SpeB activity (Friães *et al.*, 2015b).

Among the 320 SSTI isolates, 79 (25%) had no detectable SpeB activity (Dataset). These isolates represented 22 different *emm* types and 28 STs.

Among the isolates with no SpeB activity, 15 distinct *covR* alleles were identified, resulting in only two different amino acid sequences, while 34 and 47 different alleles were detected for *covS* and *ropB*, corresponding to 22 and 37 distinct amino acid sequences, respectively (Figure 2). In two isolates no PCR product was amplified using the *ropB* specific primers, indicating a possible deletion involving the *ropB* gene, as previously described in other GAS isolates (Friães *et al.*, 2013b). In one of these isolates, the absence of PCR amplification of the *speB* and *speF* genes, both located in the same locus than *ropB*, supports the occurrence of a large deletion encompassing the entire *ropB* gene, as well as *speB* and *speF*.

No null alleles were detected in *covR*, while null *covS* alleles were detected in 10 isolates. Since none of the SpeB-positive isolates is expected to have null *covS* alleles, these would correspond to 3% of all isolates recovered from SSTI. These null *covS* alleles were present in isolates of six different *emm* types and seven STs (Dataset), and there were no associations with specific *emm* types or STs. Two isolates presented in-frame indels (45 bp deletion and 3 bp insertion). Since both isolates also have a null *ropB* allele which could explain the downregulation of SpeB, the consequences of these indels in CovS protein function are not clear and therefore they were not considered null alleles.

A total of 12 isolates presented null *ropB* alleles (including the two isolates with a *ropB* deletion), representing 4% of all SSTI isolates, and 15% of the isolates with no detectable SpeB.

Overall, a total of 22 SSTI isolates (29%) presented a null allele in at least one of the three genes. Forty-four SpeB-negative SSTI isolates carry amino acid changes in *covRS* or *ropB* whose impact on SpeB expression is difficult to predict (Friães *et al.*, 2015b). The remaining 13 isolates with no detectable SpeB showed the same amino acid sequence as the reference strain for all sequenced genes (*emm*71, $n = 1$; *emm*80, $n = 1$; *emm*87, $n = 1$; *emm*89, $n = 8$; *emm*109, $n = 1$; *emm*122, $n = 1$). In these isolates the absence of SpeB probably resulted from changes other than the impairment of CovRS or RopB function, possibly due to mutations in one of the at least 19 other genes whose products have been shown to influence the production of an enzimatically active SpeB (Carroll and Musser, 2011). No *emm* type was associated with the absence of detectable SpeB activity and the only ST (ST382 representing 65% of all *emm*6 isolates) associated with a lack of SpeB activity ($p = 0.004$) presented no null alleles at any of the sequenced loci. However, isolates of this ST harbor multiple amino acid substitutions in both *covS* and *ropB* genes that could contribute to the speB-negative phenotype.

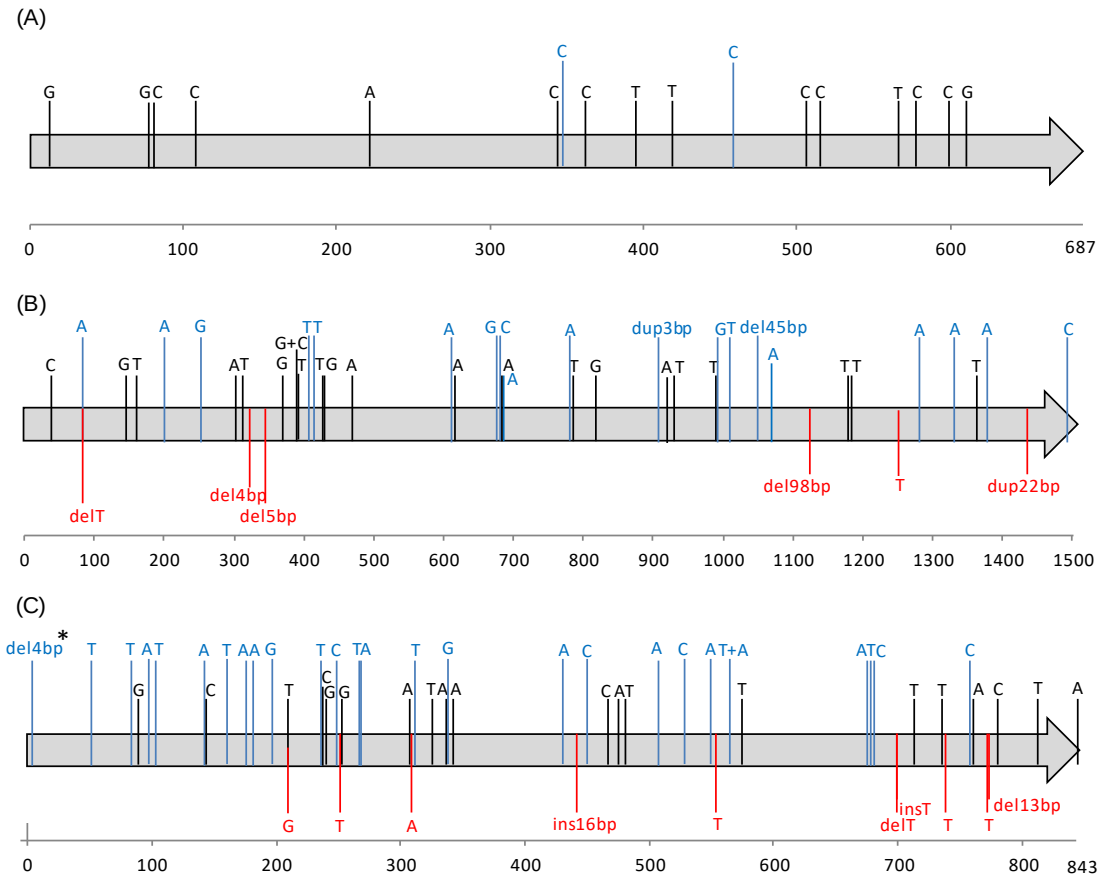


Figure 2. Nucleotide alterations identified in the *covR* (A), *covS* (B), and *ropB* (C) alleles among skin and soft tissue isolates with no detectable SpeB (*n* = 79), relative to the alleles present in strain SF370 (AE004092). Each gene is indicated by a grey arrow with the respective numbering scale below. Each sequence variation is indicated in the respective nucleotide position by a letter corresponding to the variant nucleotide, or the sequence or number of base-pairs inserted (ins), deleted (del), or duplicated (dup). Synonymous changes are represented in black; missense mutations and in-frame indels are represented in blue. Changes predicted to result in null alleles, including nonsense mutations and indels that generate frameshifts are represented in red. One exception is represented by "*" corresponding to an isolate with an early deletion of 4 bp in *ropB* which is more probable to result in the alteration of the start codon rather than a premature stop codon and therefore was not considered as a null mutation.

DISCUSSION

Antimicrobial resistance among SSTI isolates was not significantly different from that found among contemporary iGAS and pharyngitis isolates, with the decreasing resistance to macrolides in SSTI mirroring declines in resistance previously described among isolates causing these other infections (Friães *et al.*, 2013a; Silva-Costa *et al.*, 2015).

A high diversity of *emm* types was found among the studied isolates. Although most of the *emm* types were identified among both iGAS and SSTI isolates, there were clear differences in the overall *emm* and clonal distribution between the two infections. The *emm* types 1 and 64 associated to iGAS were also previously found to be overrepresented among iGAS relative to pharyngitis (Friães *et al.*, 2012), confirming the enhanced capacity of these lineages to cause invasive disease. The *emm*1 isolates have remained a major cause of iGAS in Portugal in recent years (Friães *et al.*, 2013a), in line with the worldwide dissemination of the M1T1 clone (Luca-Harari *et al.*, 2009; O'Loughlin *et al.*, 2007; Aziz and Kotb, 2008). On the other hand, *emm*64 decreased markedly as a cause of iGAS in Portugal in recent years (Friães *et al.*, 2013a) and was not associated with iGAS elsewhere, indicating that its importance as a cause of invasive disease was short-lived. The association between *emm*3 and iGAS is not unexpected, since *emm*3 has been reported as a major cause of invasive disease in Portugal and other countries, although the comparison with pharyngitis isolates did not identify this *emm* type (Luca-Harari *et al.*, 2009; Friães *et al.*, 2013a; Beres *et al.*, 2004).

The association between *emm*89 and SSTI has been previously reported in Northern Spain, but with no information regarding the specific clades involved (Tamayo *et al.*, 2014). In Portugal, only *emm*89 isolates lacking the *hasABC* locus, presumed to belong to the recently emerged clade 3 (Zhu *et al.*, 2015; Friães *et al.*, 2015a), were significantly associated with SSTI. This observation is in line with an increase in the prevalence of *emm*89 among SSTI in Portugal, but not in iGAS or pharyngitis, associated with the emergence of this acapsular clade (Friães *et al.*, 2015a). Our data thus raises the possibility that the genome remodeling underlying the emergence of *emm*89 clade 3 may have led to a particular propensity to cause SSTI, although this clade also quickly replaced the previously circulating *emm*89 clades in all infection types. This association with SSTI could be due to an increased capacity to

colonize the skin, as well as to an improved ability to overcome the major host defense mechanisms present in skin and soft tissue or to produce cytotoxic effects at these sites. Increased transmissibility and persistence has been suggested for clade 3-associated strains based on enhanced capacity to adhere to uncoated plastic (Turner *et al.*, 2015), while the increased expression of NADase and SLO by clade 3 has been associated with virulence in a mouse model of necrotizing fasciitis (Zhu *et al.*, 2015). However, further investigation is needed to clarify the role of these phenotypes in the specific association of *emm*89 clade 3 with SSTI.

These associations between *emm* type and disease presentation were reflected on the different prevalence of the respective STs and SAg profiles between iGAS and SSTI, in agreement with the high congruence previously observed between these three typing methods (Carrico *et al.*, 2006; Friães *et al.*, 2013b). The observed association of *emm* clusters A-C3 and A-C5 with iGAS, as well as that of cluster E4 with SSTI also reflects the association of the respective *emm* types with the types of infection.

According to the predicted ability of the different M proteins to interact with host factors (Sanderson-Smith *et al.*, 2014), invasive isolates would be associated with the ability to bind fibrinogen and albumin. The ability to bind fibrinogen was proposed as a mechanism to decrease complementation deposition resulting from the activation of the classical pathway (Carlsson *et al.*, 2005). A similar function is believed to be performed by C4BP (Carlsson *et al.*, 2003), whose binding was associated with SSTI isolates. The reasons why complement inhibition could be achieved through fibrinogen in iGAS and C4BP recruitment in SSTI remain elusive. Most *emm* clusters include proteins able to bind albumin, with the exception of cluster E4 (Sanderson-Smith *et al.*, 2014). The association of cluster E4 with SSTI resulted in the overrepresentation of albumin binding among iGAS. Although albumin binding was shown to mask epitopes in the C-repeated region of the M protein (Sandin *et al.*, 2006), deletion of this region did not impair virulence in a mouse intraperitoneal infection model (Waldemarsson *et al.*, 2009), not providing clues as to why this could be important in the context of iGAS. The ability to bind to human IgG was associated with SSTI isolates. Immunoglobulin binding was shown to hinder opsonophagocytosis, even in the absence of a specific immune response (Carlsson *et al.*, 2003). One can imagine such defense mechanism could be useful to bacteria in the context of both SSTI and iGAS and the reasons for the observed differences deserve further scrutiny. Only 5% of the SSTI isolates were predicted to bind plasminogen, although this was recently shown to facilitate

keratinocyte invasion (Siemens *et al.*, 2011) and could thus be beneficial in the context of SSTI.

The acquisition of mutations in the two-component regulatory system CovRS is regarded as an important mechanism promoting the transition to an invasive phenotype among GAS isolates. Current data suggests that upon infection with a wild type strain, *covRS* mutants arise at the focal infection site, from which mixed populations are recovered, and are subsequently selected for during transition to deeper, normally sterile sites (Sumby *et al.*, 2006; Mayfield *et al.*, 2014). In our previous work, we observed that null *covS* alleles were significantly overrepresented among iGAS (10%) relative to pharyngeal isolates (2%) (Friães *et al.*, 2015b). The data reported here reveals an underrepresentation of *covS* null alleles in SSTI (3%) relative to iGAS isolates ($p = 0.009$), but no significant difference relative to pharyngeal isolates (2%). Since our isolates were obtained from a single colony from each patient, there was a sampling of possibly mixed populations, which could introduce a bias or at least underestimate the ongoing selection of *covRS* impaired variants. Still we do not believe this compromises our conclusions since the colonies were randomly picked without particular care to select any of the variants. Our results therefore indicate that mutations impairing CovRS are a hallmark of iGAS isolates, although they still represent a minority of iGAS (Friães *et al.*, 2015b). The lower prevalence of *covS* null alleles among pharyngeal and SSTI isolates supports the importance of a functional CovRS in the initial stages of non-invasive infection, both in skin and in the upper respiratory tract (Hollands *et al.*, 2010; Alam *et al.*, 2013). As reported for pharyngeal and invasive isolates (Friães *et al.*, 2015b), *covRS* and *ropB* mutations occurred in SSTI isolates of diverse lineages and were not a particular characteristic of any specific clone. In contrast to *covS*, the proportion of *ropB* null alleles among SSTI isolates (4%) was similar to that found among both pharyngeal and invasive isolates (3%), indicating that mutations abrogating RopB activity occur in a low fraction of GAS isolates, regardless of the type of infection they cause.

The abrogation of SpeB activity is usually regarded as one of the most important features of CovRS mutants contributing to the invasive phenotype, since this abrogation spares multiple virulence factors involved in invasive disease pathogenesis and evasion of host immunity (Carroll and Musser, 2011). The proportion of isolates lacking SpeB among our SSTI collection (25%) was similar to that found previously among iGAS (24%) and higher, but not significantly, than that found among pharyngitis isolates in

Portugal (16%, p = 0.081) (Friães *et al.*, 2015b). Only one ST was associated with a lack of SpeB expression (ST382) and both this ST and the *emm* type it expresses (*emm*6) were evenly distributed in both SSTI and iGAS. These results suggest that, in addition to SpeB downregulation, the differential regulation of other virulence genes induced by impairment of CovRS is also critical for an enhanced invasive capacity of the isolates. In agreement with this suggestion, a mouse model of SSTI showed that *covRS* mutants produced larger necrotic regions than *speB* mutants (Engleberg *et al.*, 2004).

A limitation of this study is the lack of information regarding the specific infection caused by the isolates, as well as the clinical evolution and outcome of each infection. SSTI include a wide range of superficial and deep infections which vary greatly in severity. It was therefore not possible to evaluate possible correlations between the clones identified in SSTI and the severity of the respective infections. This could bias the comparison between SSTI and iGAS, leading to a possible underestimation of the differences between the two types of infection. However, even in these conditions we still identified significant differences between the two subsets of isolates regarding both clonal composition and presence of null *covS* alleles, which are in agreement with our previous observations from the comparison between pharyngeal and iGAS isolates (Friães *et al.*, 2012; Friães *et al.*, 2015b).

GAS isolates causing SSTI in Portugal are a temporally stable and genetically diverse population. Still, the 30 valent M protein-based vaccine, which was shown to evoke cross-opsonic antibodies against non-vaccine serotypes (Dale *et al.*, 2011), could potentially cover up to 95% of SSTI in Portugal. Although SSTI are the main primary foci associated with invasive disease (Lamagni *et al.*, 2008), SSTI isolates, inasmuch as pharyngeal isolates (Friães *et al.*, 2012), present a clearly different clonal composition from contemporary iGAS isolates. It is less clear how the differences in the presumed interaction with host proteins and exotoxin profiles brought about by those clonal differences could explain the distinct disease presentations. Despite the different prevalence of multiple *emm* types between SSTI and iGAS, within each *emm* type the same MLST defined lineages and the same SAg profiles could be found in both infection types. This indicates that any intra-*emm* type genetic differences between the two populations must be explored at a more detailed level, such as by whole-genome sequencing. Our current data therefore indicates that among the GAS clones causing infection in the Portuguese population, some have an increased capacity to invade

deeper tissues and cause severe infections, while others seem to be particularly successful at establishing SSTI or pharyngeal infections. The proportion of SSTI isolates with no detectable SpeB activity was modest and similar to that found previously in Portugal among iGAS and pharyngitis isolates, indicating that the selective pressure to eliminate SpeB is also not a primary factor in SSTI. However, we confirmed an association between mutations abrogating CovRS function and iGAS, suggesting that increased expression of other virulence factors, such as the hyaluronic acid capsule, streptolysins or NAD-glycohydrolase, rather than SpeB downregulation, may be under selection in the context of iGAS.

ACKNOWLEDGMENTS

Members of the Portuguese Group for the Study of Streptococcal Infections are: Teresa Vaz, Marília Gião, Rui Ferreira (Centro Hospitalar do Barlavento Algarvio); Ana Cristina Silva, Hermínia Costa, Maria Fátima Silva, Maria Amélia Afonso (Centro Hospitalar de Entre Douro e Vouga); Ana Domingos, Gina Marrão, José Grossinho (Centro Hospitalar de Leiria); Paulo Lopes, Angelina Lameirão, Gabriela Abreu; Aurélia Selaru (Centro Hospitalar de Vila Nova de Gaia/Espinho); Hermínia Marques, Margarida Tomáz, Paula Mota (Centro Hospitalar do Alto Ave); Maria Helena Ramos, Ana Paula Castro (Centro Hospitalar do Porto); Fernando Fonseca (Centro Hospitalar da Póvoa do Varzim/Vila do Conde); Nuno Canhoto, Teresa Afonso (Hospital Central do Funchal); Teresa Pina, Helena Peres, Odete Chantre, João Marques, Cristina Marcelo, Isabel Peres, Isabel Lourenço, Margarida Pinto (Centro Hospitalar de Lisboa Central); Lurdes Monteiro, Luís Marques Lito (Centro Hospitalar Lisboa Norte); Cristina Toscano, Maria Ana Pessanha (Centro Hospitalar Lisboa Ocidental); Elmano Ramalheira, Raquel Diaz, Sónia Ferreira, Inês Cravo Roxo (Centro Hospitalar do Baixo Vouga); Ana Paula Castro (Hospital de Vila Real); Graça Ribeiro, Rui Tomé, Celeste Pontes, Luísa Boaventura, Catarina Chaves, Teresa Reis (Hospitais da Universidade de Coimbra); Ana Buschy Fonseca (Hospital de Cascais); Manuela Ribeiro, Helena Gonçalves (Hospital de São João, Porto); Alberta Faustino, Adelaide Alves, Maria Cármen Iglesias (Hospital de Braga); Ilse Fontes, Paulo Martinho (Hospital de Santa Luzia, Elvas); Maria Luísa Gonçalves, Olga Neto (Hospital dos SAMS, Lisboa); Luísa Sancho (Hospital Dr. Fernando da Fonseca, Amadora/Sintra); Adriana Coutinho (Hospital do Espírito Santo, Évora); José Diogo, Ana Rodrigues (Hospital Garcia de Orta, Almada); Maria Antónia Read, Valquíria Alves, Margarida Monteiro (Hospital Pedro Hispano, Matosinhos); Rosa Bento (Unidade Local de Saúde do Baixo Alentejo, Beja).

Funding

This work was partially supported by LISBOA-01-0145-FEDER-007391, project cofunded by FEDER, through POR Lisboa 2020 - Programa Operacional Regional de Lisboa, PORTUGAL 2020, and Fundação para a Ciência e a Tecnologia. CP was

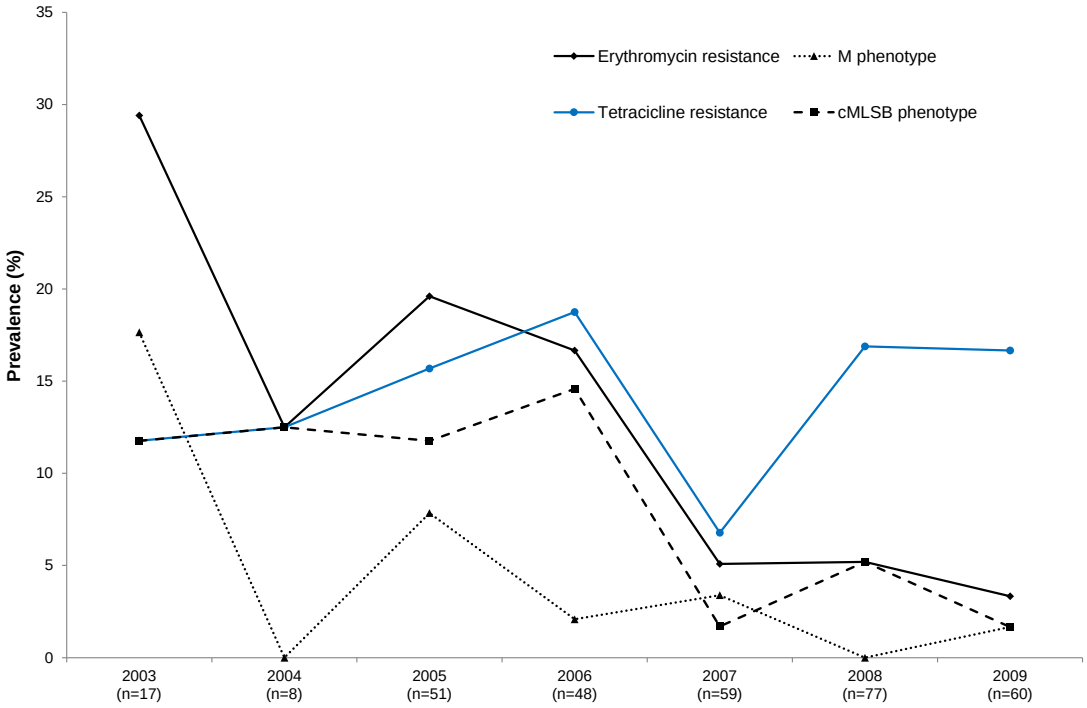
supported by a grant from Fundação para a Ciência e Tecnologia (SFRH/BD/81329/2011).

SUPPLEMENTARY DATA

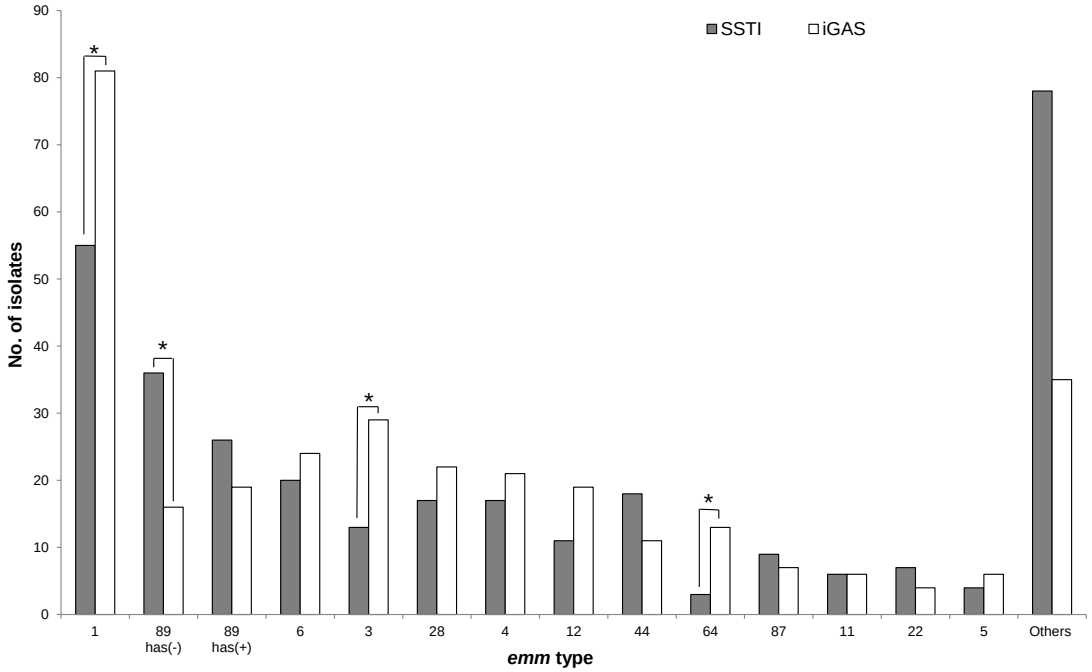
Supplementary Table 1- Host protein binding profile prediction for each cluster (Sanderson-Smith *et al.*, 2014) and distribution of isolates according the type of infection.

***emm* cluster**	**SSTI/iGAS[a]**	**Plasminogen**	**IgA**	**IgG**	**Fibrinogen**	**Albumin**	**C4BP**
E4	99/65	No	UC[b]	Yes	No	No	UC[b]
A-C3	55/82	No	No	Yes	Yes	Yes	No
E3	43/28	No	UC[b]	Yes	No	Yes	Yes
E1	21/24	No	Yes	Yes	No	Yes	Yes
M6	20/24	No	No	No	No	UC[b]	UC[b]
A-C5	13/29	No	No	No	Yes	Yes	No
D4	16/16	Yes	UC[b]	No	No	UC[b]	No
A-C4	11/19	No	No	No	Yes	Yes	No
E6	21/8	No	Yes	Yes	No	Yes	Yes
M5	4/6	No	No	No	Yes,	UC[b]	UC[b]
E2	5/3	No	No	Yes	No	Yes	No
M18	4/2	No	No	No	No	UC[b]	UC[b]
D2	3/0	No	No	No	No	UC[b]	UC[b]
M74	1/1	No	No	No	Yes	UC[b]	UC[b]
M19	0/1	No	No	No	Yes	Yes	No
M122	1/0	No	No	No	No	UC[b]	UC[b]
M164	0/1	No	No	No	No	UC[b]	UC[b]
M179	0/1	No	No	No	Yes	UC[b]	UC[b]

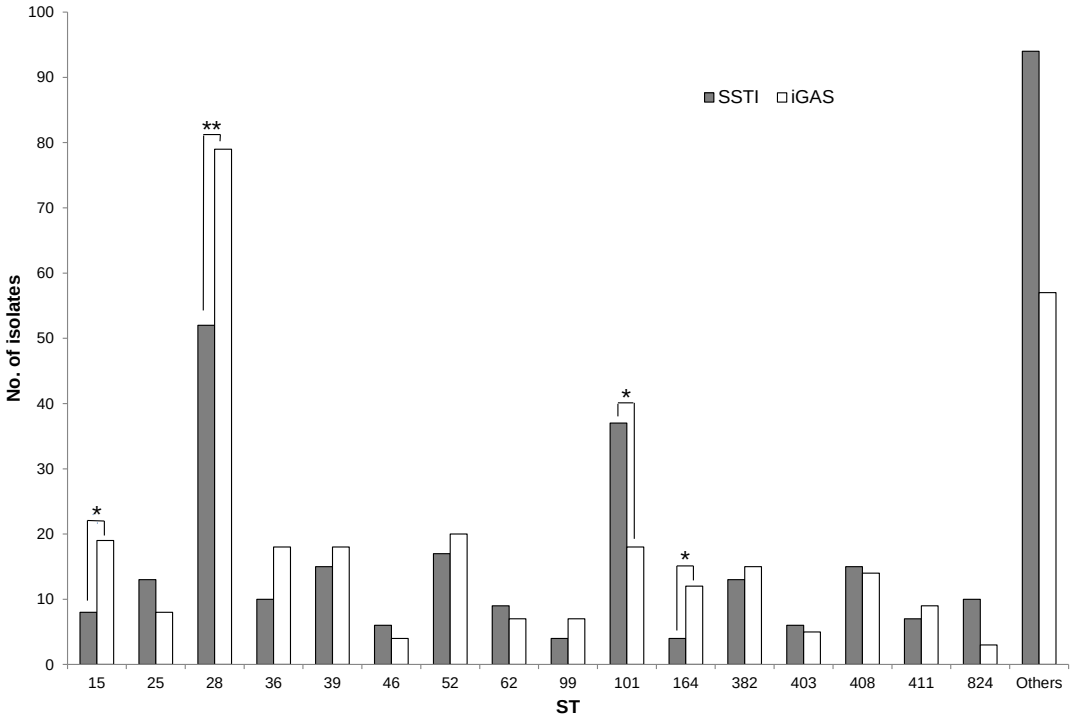
[a] Results from iGAS *emm* typing were reported previously (Friães *et al.*, 2007; Friães *et al.*, 2012; Friães *et al.*, 2013a).
[b] When < 80% isolates were tested for each *emm* cluster the property is classified as uncertain (UC).



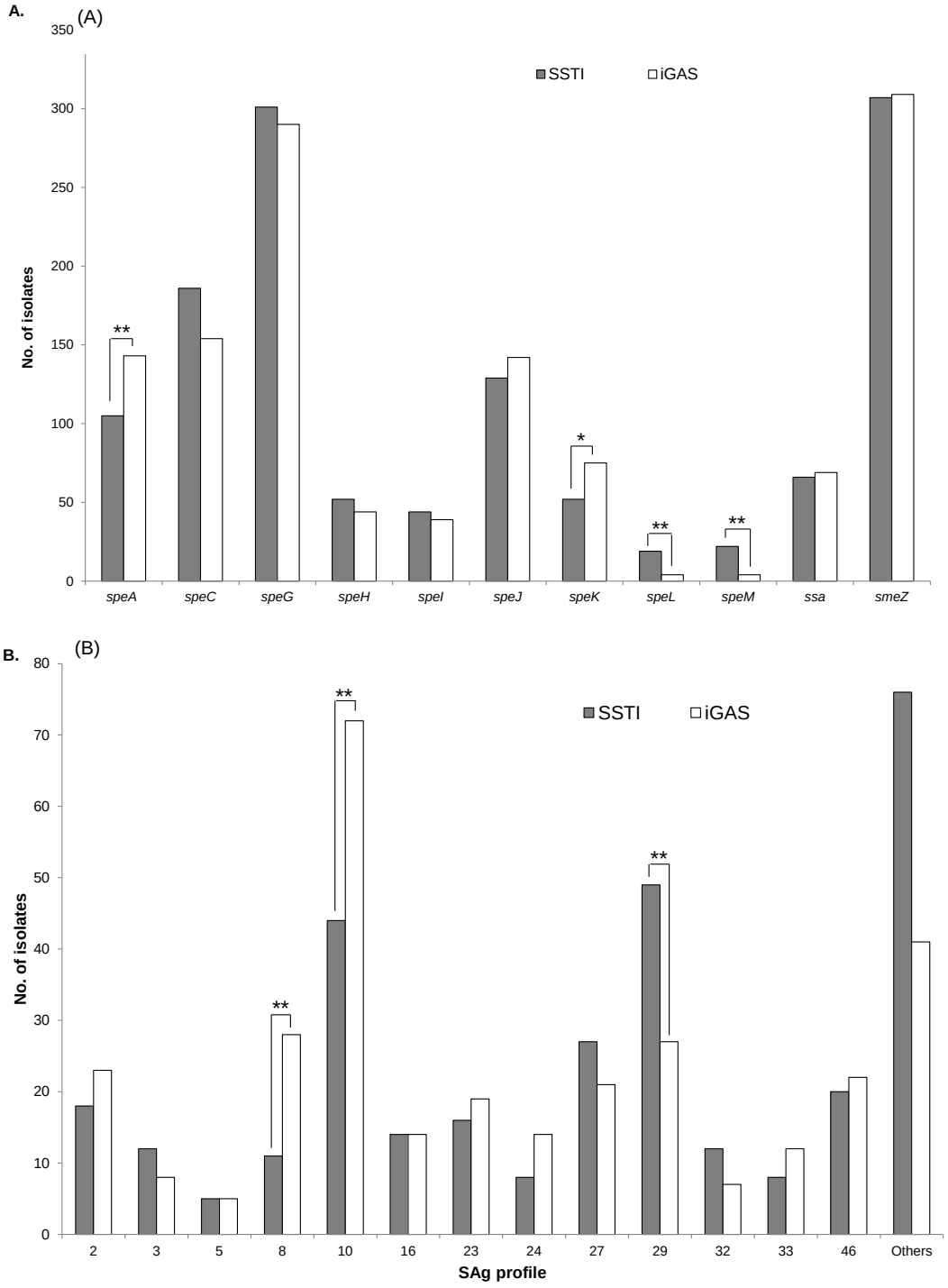
Supplementary Figure 1. Prevalence of tetracycline resistance, erythromycin resistance, and macrolide resistance phenotypes among GAS isolates recovered from skin and soft tissue infections (SSTI) in Portugal during 2003-2009. M, resistance to erythromycin and susceptibility to clindamycin; $cMLS_B$, constitutive resistance to erythromycin and clindamycin.



Supplementary Figure 2. Distribution of *emm* types among isolates recovered from skin and soft tissue infections (SSTI) and invasive infections (iGAS) in Portugal during 2003-2009. "Others" include *emm* types with a total of <10 isolates in both infection types [SSTI: *emm*2 (*n* = 8), *emm*48 (*n* = 7), *emm*75 (*n* = 5), *emm*18, *emm*58, *emm*78, and *emm*118 (each *n* = 4), *emm*9, *emm*53, and *emm*71 (each *n* = 3), *emm*70, *emm*76, *emm*77, *emm*209, and *emm*223 (each *n* = 2), *emm*33, *emm*43, *emm*50, *emm*65, *emm*68, *emm*73, *emm*74, *emm*80, *emm*81, *emm*82, *emm*83, *emm*90, *emm*93, *emm*94, *emm*102, *emm*103, *emm*109, *emm*122, *emm*127, *emm*147, *emm*167, *emm*168, and *emm*225 (each *n* = 1); iGAS: *emm*78 and *emm*113 (each *n* = 3), *emm*9, *emm*18, *emm*53, *emm*58, *emm*77, *emm*90, and *emm*103 (each *n* = 2), *emm*2, *emm*19, *emm*43, *emm*74, *emm*75, *emm*76, *emm*81, *emm*84, *emm*118, *emm*164, *emm*179, *emm*196, *emm*199, *emm*227, and stG1750 (each *n* = 1)].*$p \leq 0.01$.



Supplementary Figure 3. Distribution of sequence types (ST) among skin and soft tissue infections (SSTI) and invasive infections (iGAS) in Portugal during 2003-2009. "Others" include STs with a total of <10 isolates in both infection types [SSTI: ST55 (*n* = 8), ST161 (*n* = 7), ST315 (*n* = 5), ST75, ST130, ST150, ST253, ST642, and ST826 (each *n* = 3), ST3, ST24, ST38, ST50, ST63, ST178, ST331, ST340, ST555, ST565, and ST830 (each *n* = 2), ST2, ST5, ST10, ST60, ST89, ST120, ST166, ST167, ST184, ST200, ST341, ST389, ST409, ST429, ST467, ST569, ST573, ST643, ST701, ST718, ST754, ST817, ST819, ST820, ST822, ST825, ST827, ST828, ST829, ST831, ST833, ST834, ST835, and ST836 (each *n* = 1); iGAS: ST406 (*n* = 6), ST315 (*n* = 4), ST409, ST643 and ST679 (each *n* = 3), ST11, ST63, ST75, ST201, ST410, ST555, and ST823 (each *n* = 2), ST3, ST50, ST53, ST55, ST95, ST120, ST124, ST150, ST178, ST184, ST258, ST402, ST458, ST551, ST562, ST618, ST619, ST769, ST771, ST816, ST818, ST821, ST832, and ST833 (each *n* = 1)]. *$p < 0.05$ **$p < 0.01$. These p-values did not support statistical significance after FDR correction.



Supplementary Figure 4. Distribution of individual superantigen (SAg) genes (A) and SAg profiles (B) among isolates recovered from skin and soft tissue infections (SSTI) and invasive infections (iGAS) in Portugal during 2003-2009. "Others" include SAg profiles with a total of <10 isolates in both infection types [SSTI: SAg profiles 31 (*n* = 7), 44 (*n* = 5), 4, 20, and 40 (each *n* = 4), 26, 38, 39, 56, and 66 (each *n* = 3), 11, 12, 19, 21, 35, 43, 47, 51, 53, and 60 (each *n* = 2), 9, 13, 15, 18, 28, 41, 45, 48, 52, 54, 61, 64, 67, 68, 69, 70, and 71 (each *n* = 1); iGAS: SAg profiles 20 (*n* = 5), 44 (*n* = 4), 26 (*n* = 3), 17, 19, 30, 38, 45, 52, 54, and 56 (each *n* = 2), 1, 11, 12, 31, 34, 40, 43, 49, 50, 51, 53, 55, and 57 (each *n* = 1)]. *p < 0.05 **p < 0.01.

REFERENCES

Alam, F. M., C. E. Turner, K. Smith, S. Wiles and S. Sriskandan. 2013. Inactivation of the CovR/S Virulence Regulator Impairs Infection in an Improved Murine Model of *Streptococcus pyogenes* Naso-Pharyngeal Infection. PLoS One. 8: e61655.

Aziz, R. K. and M. Kotb. 2008. Rise and persistence of global M1T1 clone of *Streptococcus pyogenes*. Emerg Infect Dis. 14: 1511-1517.

Aziz, R. K., M. J. Pabst, A. Jeng, R. Kansal, D. E. Low, V. Nizet and M. Kotb. 2004. Invasive M1T1 group A *Streptococcus* undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. Mol Microbiol. 51: 123-134.

Benjamini, Y. 1995. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing J R Sat Soc Ser B Methodol. 57: 298-300.

Beres, S. B., G. L. Sylva, D. E. Sturdevant, C. N. Granville, M. Liu, S. M. Ricklefs, A. R. Whitney, L. D. Parkins, N. P. Hoe, G. J. Adams, D. E. Low, F. R. DeLeo, A. McGeer and J. M. Musser. 2004. Genome-wide molecular dissection of serotype M3 group A *Streptococcus* strains causing two epidemics of invasive infections. Proc Natl Acad Sci U S A. 101: 11833-11838.

Carapetis, J. R., A. C. Steer, E. K. Mulholland and M. Weber. 2005. The global burden of group A streptococcal diseases. Lancet Infect Dis. 5: 685-694.

Carlsson, F., K. Berggard, M. Stalhammar-Carlemalm and G. Lindahl. 2003. Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. J Exp Med. 198: 1057-1068.

Carlsson, F., C. Sandin and G. Lindahl. 2005. Human fibrinogen bound to *Streptococcus pyogenes* M protein inhibits complement deposition via the classical pathway. Mol Microbiol. 56: 28-39.

Carrico, J. A., A. J. Sabat, A. W. Friedrich and M. Ramirez. 2013. Bioinformatics in bacterial molecular epidemiology and public health: databases, tools and the next-generation sequencing revolution. Euro Surveill. 18: 20382.

Carrico, J. A., C. Silva-Costa, J. Melo-Cristino, F. R. Pinto, H. de Lencastre, J. S. Almeida and M. Ramirez. 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol. 44: 2524-2532.

Carroll, R. K. and J. M. Musser. 2011. From transcription to activation: how group A streptococcus, the flesh-eating pathogen, regulates SpeB cysteine protease production. Mol Microbiol. 81: 588-601.

Carroll, R. K., S. A. Shelburne, 3rd, R. J. Olsen, B. Suber, P. Sahasrabhojane, M. Kumaraswami, S. B. Beres, P. R. Shea, A. R. Flores and J. M. Musser. 2011. Naturally occurring single amino acid replacements in a regulatory protein alter

streptococcal gene expression and virulence in mice. J Clin Invest. 121: 1956-1968.

Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 13: 470-511.

Dale, J. B., T. A. Penfound, E. Y. Chiang and W. J. Walton. 2011. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine. 29: 8175-8178.

Descheemaeker, P., F. Van Loock, M. Hauchecorne, P. Vandamme and H. Goossens. 2000. Molecular characterisation of group A streptococci from invasive and non-invasive disease episodes in Belgium during 1993-1994. J Med Microbiol. 49: 467-471.

Ekelund, K., J. Darenberg, A. Norrby-Teglund, S. Hoffmann, D. Bang, P. Skinhoj and H. B. Konradsen. 2005. Variations in emm type among group A streptococcal isolates causing invasive or noninvasive infections in a nationwide study. J Clin Microbiol. 43: 3101-3109.

Engleberg, N. C., A. Heath, A. Miller, C. Rivera and V. J. DiRita. 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis. 183: 1043-1054.

Engleberg, N. C., A. Heath, K. Vardaman and V. J. DiRita. 2004. Contribution of CsrR-regulated virulence factors to the progress and outcome of murine skin infections by *Streptococcus pyogenes*. Infect Immun. 72: 623-628.

Figueira-Coelho, J., M. Ramirez, M. J. Salgado and J. Melo-Cristino. 2004. *Streptococcus agalactiae* in a large Portuguese teaching hospital: antimicrobial susceptibility, serotype distribution, and clonal analysis of macrolide-resistant isolates. Microb Drug Resist. 10: 31-36.

Fiorentino, T. R., B. Beall, P. Mshar and D. E. Bessen. 1997. A genetic-based evaluation of the principal tissue reservoir for group A streptococci isolated from normally sterile sites. J Infect Dis. 176: 177-182.

Friães, A., J. P. Lopes, J. Melo-Cristino and M. Ramirez. 2013a. Changes in *Streptococcus pyogenes* causing invasive disease in Portugal: evidence for superantigen gene loss and acquisition. Int J Med Microbiol. 303: 505-513.

Friães, A., M. P. Machado, C. Pato, J. Carrico, J. Melo-Cristino and M. Ramirez. 2015a. Emergence of the Same Successful Clade among Distinct Populations of *emm89 Streptococcus pyogenes* in Multiple Geographic Regions. MBio. 6: e01780-01715.

Friães, A., C. Pato, J. Melo-Cristino and M. Ramirez. 2015b. Consequences of the variability of the CovRS and RopB regulators among *Streptococcus pyogenes* causing human infections. Sci Rep. 5: 12057.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2012. Group A streptococci clones associated with invasive infections and pharyngitis in

Portugal present differences in emm types, superantigen gene content and antimicrobial resistance. BMC Microbiol. 12: 280.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2013b. Superantigen gene complement of *Streptococcus pyogenes*-relationship with other typing methods and short-term stability. Eur J Clin Microbiol Infect Dis. 32: 115-125.

Friães, A., M. Ramirez and J. Melo-Cristino. 2007. Nonoutbreak surveillance of group A streptococci causing invasive disease in Portugal identified internationally disseminated clones among members of a genetically heterogeneous population. J Clin Microbiol. 45: 2044-2047.

Hasegawa, T., A. Okamoto, T. Kamimura, I. Tatsuno, S. N. Hashikawa, M. Yabutani, M. Matsumoto, K. Yamada, M. Isaka, M. Minami and M. Ohta. 2010. Detection of invasive protein profile of *Streptococcus pyogenes* M1 isolates from pharyngitis patients. APMIS. 118: 167-178.

Hollands, A., R. K. Aziz, R. Kansal, M. Kotb, V. Nizet and M. J. Walker. 2008. A naturally occurring mutation in ropB suppresses SpeB expression and reduces M1T1 group A streptococcal systemic virulence. PLoS One. 3: e4102.

Hollands, A., M. A. Pence, A. M. Timmer, S. R. Osvath, L. Turnbull, C. B. Whitchurch, M. J. Walker and V. Nizet. 2010. Genetic switch to hypervirulence reduces colonization phenotypes of the globally disseminated group A streptococcus M1T1 clone. J Infect Dis. 202: 11-19.

Ikebe, T., M. Ato, T. Matsumura, H. Hasegawa, T. Sata, K. Kobayashi and H. Watanabe. 2010. Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. PLoS Pathog. 6: e1000832.

Kansal, R. G., V. Datta, R. K. Aziz, N. F. Abdeltawab, S. Rowe and M. Kotb. 2010. Dissection of the molecular basis for hypervirulence of an in vivo-selected phenotype of the widely disseminated M1T1 strain of group A Streptococcus bacteria. J Infect Dis. 201: 855-865.

Kittang, B. R., N. Langeland and H. Mylvaganam. 2008. Distribution of emm types and subtypes among noninvasive group A, C and G streptococcal isolates in western Norway. APMIS. 116: 457-464.

Lamagni, T. L., J. Darenberg, B. Luca-Harari, T. Siljander, A. Efstratiou, B. Henriques-Normark, J. Vuopio-Varkila, A. Bouvet, R. Creti, K. Ekelund, M. Koliou, R. R. Reinert, A. Stathi, L. Strakova, V. Ungureanu, C. Schalen and A. Jasir. 2008. Epidemiology of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 46: 2359-2367.

Lin, J. N., L. L. Chang, C. H. Lai, H. H. Lin and Y. H. Chen. 2014. Association between polymorphisms in the csrRS two-component regulatory system and invasive group A streptococcal infection. Eur J Clin Microbiol Infect Dis. 33: 735-743.

Luca-Harari, B., J. Darenberg, S. Neal, T. Siljander, L. Strakova, A. Tanna, R. Creti, K. Ekelund, M. Koliou, P. T. Tassios, M. van der Linden, M. Straut, J. Vuopio-Varkila, A. Bouvet, A. Efstratiou, C. Schalen, B. Henriques-Normark and A. Jasir. 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 47: 1155-1165.

Mayfield, J. A., Z. Liang, G. Agrahari, S. W. Lee, D. L. Donahue, V. A. Ploplis and F. J. Castellino. 2014. Mutations in the Control of Virulence Sensor Gene from *Streptococcus pyogenes* after Infection in Mice Lead to Clonal Bacterial Variants with Altered Gene Regulatory Activity and Virulence. PLoS One. 9: e100698.

McMillan, D. J., P. A. Dreze, T. Vu, D. E. Bessen, J. Guglielmini, A. C. Steer, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2013. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. Clin Microbiol Infect. 19: E222-229.

Melo-Cristino, J. and M. L. Fernandes. 1999. Streptococcus pyogenes isolated in Portugal: macrolide resistance phenotypes and correlation with T types. Portuguese Surveillance Group for the Study of Respiratory Pathogens. Microb Drug Resist. 5: 219-225.

Mijac, V., L. Ranin, M. Markovic, C. Heeg, R. R. Reinert and N. Opavski. 2010. Distribution of *emm* types among group A streptococcal isolates from Serbia. Clin Microbiol Infect. 16: 295-298.

O'Loughlin, R. E., A. Roberson, P. R. Cieslak, R. Lynfield, K. Gershman, A. Craig, B. A. Albanese, M. M. Farley, N. L. Barrett, N. L. Spina, B. Beall, L. H. Harrison, A. Reingold and C. Van Beneden. 2007. The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. Clin Infect Dis. 45: 853-862.

Rivera, A., M. Rebollo, E. Miro, M. Mateo, F. Navarro, M. Gurgui, B. Mirelis and P. Coll. 2006. Superantigen gene profile, emm type and antibiotic resistance genes among group A streptococcal isolates from Barcelona, Spain. J Med Microbiol. 55: 1115-1123.

Sanderson-Smith, M., D. M. De Oliveira, J. Guglielmini, D. J. McMillan, T. Vu, J. K. Holien, A. Henningham, A. C. Steer, D. E. Bessen, J. B. Dale, N. Curtis, B. W. Beall, M. J. Walker, M. W. Parker, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2014. A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. J Infect Dis. 210: 1325-1338.

Sandin, C., F. Carlsson and G. Lindahl. 2006. Binding of human plasma proteins to *Streptococcus pyogenes* M protein determines the location of opsonic and non-opsonic epitopes. Mol Microbiol. 59: 20-30.

Siemens, N., N. Patenge, J. Otto, T. Fiedler and B. Kreikemeyer. 2011. *Streptococcus pyogenes* M49 plasminogen/plasmin binding facilitates keratinocyte invasion via integrin-integrin-linked kinase (ILK) pathways and protects from macrophage killing. J Biol Chem. 286: 21612-21622.

Silva-Costa, C., J. A. Carrico, M. Ramirez and J. Melo-Cristino. 2014. Scarlet fever is caused by a limited number of *Streptococcus pyogenes* lineages and is associated with the exotoxin genes ssa, speA and speC. Pediatr Infect Dis J. 33: 306-310.

Silva-Costa, C., F. R. Pinto, M. Ramirez and J. Melo-Cristino. 2008. Decrease in macrolide resistance and clonal instability among *Streptococcus pyogenes* in Portugal. Clin Microbiol Infect. 14: 1152-1159.

Silva-Costa, C., M. Ramirez and J. Melo-Cristino. 2015. Declining macrolide resistance in *Streptococcus pyogenes* in Portugal (2007-13) was accompanied by continuous clonal changes. J Antimicrob Chemother. 70: 2729-2733.

Sumby, P., A. R. Whitney, E. A. Graviss, F. R. DeLeo and J. M. Musser. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog. 2: e5.

Tamayo, E., M. Montes, J. M. Garcia-Arenzana and E. Perez-Trallero. 2014. *Streptococcus pyogenes emm*-types in northern Spain; population dynamics over a 7-year period. J Infect. 68: 50-57.

Trzcinski, K., B. S. Cooper, W. Hryniewicz and C. G. Dowson. 2000. Expression of resistance to tetracyclines in strains of methicillin-resistant *Staphylococcus aureus*. J Antimicrob Chemother. 45: 763-770.

Turner, C. E., J. Abbott, T. Lamagni, M. T. Holden, S. David, M. D. Jones, L. Game, A. Efstratiou and S. Sriskandan. 2015. Emergence of a New Highly Successful Acapsular Group A *Streptococcus* Clade of Genotype *emm89* in the United Kingdom. MBio. 6: e00622.

Vahakuopus, S., R. Vuento, T. Siljander, J. Syrjanen and J. Vuopio. 2012. Distribution of emm types in invasive and non-invasive group A and G streptococci. Eur J Clin Microbiol Infect Dis. 31: 1251-1256.

Waldemarsson, J., M. Stalhammar-Carlemalm, C. Sandin, F. J. Castellino and G. Lindahl. 2009. Functional dissection of *Streptococcus pyogenes* M5 protein: the hypervariable region is essential for virulence. PLoS One. 4: e7279.

Walker, M. J., T. C. Barnett, J. D. McArthur, J. N. Cole, C. M. Gillen, A. Henningham, K. S. Sriprakash, M. L. Sanderson-Smith and V. Nizet. 2014. Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. Clin Microbiol Rev. 27: 264-301.

Zhu, L., R. J. Olsen, W. Nasser, I. de la Riva Morales and J. M. Musser. 2015. Trading Capsule for Increased Cytotoxin Production: Contribution to Virulence of a Newly Emerged Clade of *emm89 Streptococcus pyogenes*. MBio. 6: e01378-01315.

CHAPTER IV

GENERAL DISCUSSION

Streptoccocus pyogenes is responsible for a wide variety of human infections ranging from mild superficial infections of the respiratory tract and skin, such as pharyngitis and impetigo, to potentially fatal invasive infection such as necrotizing fasciitis and streptococcal toxic shock syndrome (Cunningham, 2000). In order to cause this wide spectrum of diseases, GAS has to be able to adapt to the different environments in the host and despite decades of research, there is still no consensus regarding what are the genotypic or phenotypic properties responsible for an enhanced invasive potential of certain isolates. Moreover, most studies in attempting to find the answer to this question compare subsets of invasive isolates with subsets of non-invasive isolates recovered from several non-invasive sites, including from skin and more frequently from pharyngeal swabs. However, since the characteristics of SSTI isolates are reported together with GAS from other non-invasive sites, there is a lack of data regarding the characteristics of GAS isolates specifically responsible for SSTI.

In the present thesis, we sequenced the *covRS* and *ropB* genes of 191 GAS isolates, including isolates from pharyngitis and invasive infections with a high diversity of *emm* types and STs that were previously characterized (Friães *et al.*, 2012). We also screened all isolates for the presence of SpeB activity and quantified their NADase and SLS activity.

In the second part, we characterized a total of 320 isolates from SSTI recovered in Portugal and performed the comparison with invasive isolates recovered during the same period which had been characterized elsewhere (Friães *et al.*, 2007; Friães *et al.*, 2013a). All SSTI isolates were tested for SpeB activity and for those without detectable SpeB activity we also determined the sequence of the *covRS* and *ropB* genes.

Importance of CovRS and RopB spontaneous mutations among pharyngitis, invasive and SSTI isolates.

In 2006, Sumby *et al.*, using a microarray analysis of 9 M1 isolates including six from patients with pharyngitis and three from invasive disease, found mainly two different transcriptomes named pharyngeal transcriptome profile (PTP) and invasive transcriptome profile (ITP). They also observed that isolates with and ITP were recovered from mice infected subcutaneously with PTP isolates and that these isolates

were more able to resist killing by human polymorphonuclear leukocytes. Moreover using complete genome sequencing, the only difference between an ITP isolate from its PTP precursor was a 7bp frameshift mutation in the *covS* gene, which results in the production of a truncated 202-amino acid CovS (Sumby *et al.*, 2006). The operon where *covS* is found encodes the two-component regulatory system CovRS, which is estimated to directly or indirectly influence the expression of 10-15% of the GAS genome (Graham *et al.*, 2002; Sumby *et al.*, 2006). In agreement with these results, others studies have also reported the presence of mutations impairing *covRS* function in isolates recovered from murine infections as well in patient isolates recovered mostly from invasive infections (Engleberg *et al.*, 2001; Ikebe *et al.*, 2010; Mayfield *et al.*, 2014; Kansal *et al.*, 2010). Some of these studies, using mainly M1T1 isolates, also indicate that isolates with these mutations show an enhanced virulence in murine models of infection (Engleberg *et al.*, 2001; Ikebe *et al.*, 2010; Mayfield *et al.*, 2014). Taken together, the previous results led to the hypothesis that spontaneous mutations in the *covRS* operon during SSTI results in a phenotypic switch that could explain the increased capacity of certain isolates to invade deeper tissues by changing their regulatory networks.

However, contrasting results have been reported, namely regarding which are the consequences of spontaneous CovRS mutations on mouse models of infection and more specifically on the expression of several virulence factors, such as SpeB and streptolysins. We reported that *covS* null alleles were significantly associated with an increased activity of both NADase and SLS, and with the absence of SpeB production. These results are in agreement with previous studies reporting that the acquisition of null *covS* alleles actually results in the downregulation of SpeB and upregulation of NADase and SLS (Hasegawa *et al.*, 2010; Aziz *et al.*, 2010). However, we found a few exceptions in specific *emm* types, namely two *emm*64 isolates with null *covS* alleles that, in contrast with other isolates with null *covS* alleles, did not express higher levels of NADase activity nor of SLS activity. We suggest that these results could be due to intrinsic characteristics of this lineage, such as alterations in the *nga* gene, its promoter, or other regulatory pathways. Moreover two isolates with null *covS* alleles belonging to *emm* type 6 and 44, respectively, also did not exhibit an increased SLS activity suggesting that these isolates also show a potential different CovRS regulon in which the *sag* operon is not controlled by CovRS for example. The existence of mutations in the *sag* operon of these isolates is also a possibility. The use of isolates with distinct

emm types (and hence representing different genetic lineages), could be one reason, besides to the type and localization of the mutation, for the conflicting data between studies regarding the impact of *covRS* mutations in virulence. In agreement, with their diversity of phenotypic behaviours, it was shown that different *in vivo* induced *covR* mutations can result in markedly different transcriptomes (Trevino *et al.*, 2009). Additionally, some studies highlight the interaction of CovRS with other transcriptional regulators, resulting in complex regulation patterns that can vary between different *emm* types and even between strains of the same *emm* type (Aziz *et al.*, 2010; Sugareva *et al.*, 2010; Horstmann *et al.*, 2014).

Naturally occurring mutations in the standalone transcriptional regulator RopB have also been reported, and similarly to CovRS, with contrasting results regarding their impact in virulence in mouse models and in the expression of several virulence factors. The exception is SpeB, for which a functional RopB is considered an absolute requirement (Carroll *et al.*, 2011; Ikebe *et al.*, 2010; Hollands *et al.*, 2008; Neely *et al.*, 2003). In our studies, all isolates with null *ropB* alleles had no detectable SpeB activity, in agreement with its role as a direct regulator of SpeB (Neely *et al*., 2003). Additionally, null *ropB* alleles were not associated with altered NADase or SLS activity, suggesting that RopB does not contribute significantly to the expression of the *nga* and *sag* operons. These results disagree with previous studies that reported an enhanced expression of several virulence factors attributed to *ropB* mutations (Carroll *et al.*, 2011; Ikebe *et al.*, 2010). However, it was shown that RopB regulation can be highly variable even among strains of the same *emm* type, what can explain these differences (Dmitriev *et al.*, 2008).

Mutations in *covRS* have been previously identified in clinical isolates. However while some studies suggest that *covRS* and *ropB* mutations occur more frequently among isolates from invasive infections, others found them in a similar proportion among non-invasive isolates (Hasegawa *et al.*, 2010; Ikebe *et al.*, 2010; Lin *et al.*, 2014). In addition, some studies that used mainly murine models suggest that isolates of the highly invasive M1T1 clone are more prone to acquire *covRS* mutations than strains of other *emm* types (Mayfield *et al.*, 2014; Maamary *et al.*, 2010). Considering our results, null *covS* alleles were found to have a significant association with invasive infections in comparison with both pharyngitis and SSTI isolates suggesting that these mutations can contribute for a higher invasive potential of these isolates. Even so, these mutations were present in only 10% of invasive isolates and were not associated to any

specific *emm* type, not even with those previously associated with invasive infections, but occurred in isolates of diverse lineages which is in agreement with reports from Japan and Taiwan (Ikebe *et al.*, 2010; Lin *et al.*, 2014).

Regarding null *ropB* alleles, they were not differently distributed among isolates recovered from any of the type of infection, suggesting they are not related with the higher ability of some isolates to cause infection.

In our isolates, no null *covR* mutations were detected in any of the infections, which is in agreement with the hypothesis that during the infection process, the acquisition of mutations that impair signalling through CovS and keep CovR functional and possibly responsive to phosphorylation by other kinases, such as SP-STK, may contribute for the progression to invasive infection (Trevino *et al.*, 2009; Agarwal *et al.*, 2011).

The loss of SpeB activity due to spontaneous *covRS* mutations have been considered the main event for explaining the increased virulence observed in some studies. This led to the hypothesis that M1T1 isolates are under selective pressure to decrease SpeB protease activity during infection. It was reported that the absence of SpeB allows the sequestration of human plasmin protease activity on the GAS surface, which may contribute to GAS systemic dissemination. The absence of SpeB also prevents the degradation of several GAS virulence factors, including M1 protein and Sda1, contributing for survival to neutrophil killing and degradation of NETs (Cole *et al.*, 2006; Walker *et al.*, 2007; Aziz *et al.*, 2004). In agreement, an inverse relationship between disease severity and expression of SpeB among M1T1 isolates recovered from invasive infections was reported. Among isolates recover from SSTS, 41% produced little or no SpeB compared to only 14% of isolates recovered from non-severe cases. Moreover, the protease activity among those isolates that expressed SpeB was significantly lower for STSS isolates than for isolates from non-severe cases (Kansal *et al.*, 2000). In our results, the majority of isolates produce detectable SpeB activity. Moreover, the proportion of isolates lacking SpeB expression among our SSTI collection (25%) was similar to that found among iGAS (24%) and higher but not significantly to that found among pharyngitis isolates (16%). Additionally, the absence of SpeB was not associated with invasive infection neither with *emm* types previously associated with invasive infection such as *emm*1 and *emm*64. These results are in agreement with a recent study that observed that among 6775 isolates, 84.3% of the strains have a wild-type SpeB phenotype. Furthermore, they also did not detect any

difference in the frequency of the SpeB-deficient phenotype between strains recovered from invasive infections and those recovered from pharyngitis (Olsen *et al.*, 2015). The authors suggest that, contrary to the idea that the downregulation of SpeB promotes virulence, SpeB expression is actually a key contributor to pathogenesis. To support their results, the authors cite some of the following studies: it was observed that systemic infection in murine models of infection GAS requires an intact RopB and an efficient SpeB production (Hollands *et al.*, 2008); consistently, isolates with *ropB* polymorphisms that lack SpeB activity killed significantly fewer mice and cause smaller lesions with less tissue destruction (Olsen *et al.*, 2012; Carroll *et al.*, 2011); *speB* mutants are less virulent in mouse models in comparison of their wt strains (Lukomski *et al.*, 1997; Kuo *et al.*, 1998; Lukomski *et al.*, 1999); SpeB decreases biofilm formation *in vitro* what was associated with larger lesions in mice (Connolly *et al.*, 2011); SpeB inactivates the human cathelicidin LL-37 *in vitro*; the SpeB activity was also shown to contribute to escape from IgG mediated phagocytosis increasing GAS ability to survive in immune blood (Eriksson and Norgren, 2003); acute-phase serum of patients with GAS bacteremia had significantly lower neutralizing ability against SpeB than did sera from patients with uncomplicated tonsillitis (Norrby-Teglund *et al.*, 1994); similarly, neutralization of SpeB was significantly lower in acute-phase sera of patients with STSS than in sera from persons with bacteremia or erysipelas (Eriksson *et al.*, 1999); immunization of mice with SpeB or treatment with a protease inhibitor confers protection against lethal challenge or reduced the size of the skin lesions (Bjorck *et al.*, 1989; Kapur *et al.*, 1994; Ulrich, 2008; Connolly *et al.*, 2011).

Despite isolates with *covRS* mutations possessing an apparent selective advantage in invasive infection due to improved neutrophil resistance and propensity for bacterial dissemination, we concluded through Ka/Ks determination that *covS*, *covR* and *ropB* genes are under stabilizing selection. This means that selection favours alleles that do not change the amino acid sequence. One explanation for why isolates with these mutations have not become prevalent in the community is the presence of an overall fitness cost specific to nasopharyngeal infection or colonization that can affect the initial stages of infections that then progress to more serious manifestations (Alam *et al.*, 2013). It was reported that Δ*covRS* is detrimental to long term infection of the nasopharynx and to the transmission (Alam *et al.*, 2013). Consistently, *covS* mutants showed a reduced capacity to bind to epithelial cell layers as a consequence of increased capsule expression and also a reduced capacity to bind fibronectin and to form biofilms

on plastic and epithelial cell layers. A defect in skin adherence of the *covS* mutant strain was also demonstrated in a murine model in *vivo*, suggesting that the fitness cost could also be incurred in skin additionally to the nasopharynx (Hollands *et al.*, 2010). This theory of an associated fitness cost is also supported by a reported inverse correlation between the ability to adhere to host cells and GAS virulence (Miyoshi-Akiyama *et al.*, 2009). Furthermore, it was shown that GAS isolates with mutations in *covS* lack competitiveness in human saliva relative to the wt (Trevino *et al*., 2009). This phenotype could also be related with the levels of SpeB production, since it was demonstrated that SpeB contributes to persistence and growth in human saliva *in vitro* in M1 GAS (Shelburne *et al.*, 2005). Moreover, the presence of SpeB is required for the establishment of localized skin infections (Cole *et al*., 2006). It was reported that in a murine skin and soft tissue infection model the inoculation of a mixture of the wild-type strain and its *covRS* mutant resulted in more necrotic lesions than did either strain at twice the inoculum, which suggests that these variants may exhibit pathogenic synergy (Engleberg *et al*., 2001), highlighting the complex interactions between these two phenotypes *in vivo*.

Even though we have found a significant association of null *covS* alleles with invasive isolates we did not observe any further association between the presence/absence of detectable SpeB activity, the activity of NADase or SLS with invasiveness. These results suggest that other factors under CovRS regulation may also be critical for an enhanced invasive capacity of the isolates. Actually, in agreement with our suggestion, inoculation of CovRS mutants in a mouse model of SSTI yielded larger necrotic regions that SpeB mutants (Engleberg *et al.*, 2004), but the nature of these possible enhancers of pathogenicity remains undefined.

It has to be taken into consideration that our isolates were obtained from a single colony from each patient. According to Sumby *et al*., regardless of the transcriptional profile of the original infecting strain (ITP or PTP), all GAS isolated from the spleens of infected animals were mucoid (ITP_*covS*mut) and, although all GAS isolated from the skin lesions of mice infected with the ITP strain were also mucoid, colonies grown from the skin lesions of PTP-infected mice were a mixture of nonmucoid parental and mucoid variant strains at approximately 1:1 ratio (Sumby *et al*., 2006). Considering that a similar mechanism could occur in SSTI in humans, we could possibly have introduced a bias by sampling a colony of a possibly mixed population. Even so, since we picked

each colony without particular care to select any of the variants we do not believe that this fact compromises our conclusions.

Moreover, 4.8% of isolates among pharyngitis and invasive isolates with no SpeB activity had wt alleles in all 3 genes. Among skin and soft tissue isolates, this rate was higher but not significantly (15.8%). The generation of the final SpeB product is a complex, multistage process, dependent on at least 21 gene products, therefore a mutation in any one of them could lead to isolates with no detectable SpeB activity (Olsen *et al*., 2015). The higher value among SSTI could be due to the higher diversity of the sample. While among pharyngitis and invasive isolates, the sample was chosen to include one third of the isolates of each *emm* type present in the collection with n≥3 as well all *emm* types 1 and 64, all isolates from SSTI infection recovered during 2003 and 2009 were screened for SpeB activity. It could also be related with the methodology that we used for screening of the SpeB activity. While among pharyngitis and invasive isolates we used two methods (azocasein and casein assay) combined with WB, for SSTI only the casein assay was used in combination with WB. Transcription and posttranscriptional processing of SpeB is influenced by various environmental signals (Olsen *et al*., 2015), including host factors that are not well defined and can contribute to significant differences between *in vivo* and *in vitro* observations. This is highlighted by Sumby *et al*., for the case of isolate MGAS5005 (ITP), that we also used as our negative control, that despite minimal SpeB expression *in vitro* (Sumby *et al*., 2006), it was reported that this strain produces extensive amounts of SpeB when grown *ex vivo* in human saliva (Shelburne *et al*., 2005).

Characterization of SSTI isolates.

Most of the information regarding GAS isolates responsible for SSTI isolates are from studies that compare isolates from invasive and non-invasive isolates (Montes *et al.*, 2011; Ekelund *et al.*, 2005; Rivera *et al.*, 2006; Lintges *et al.*, 2010; Wajima *et al.*, 2008). Generally, invasive isolates are described as isolates collected from sterile sites such as blood, cerebrospinal fluid, ascitic fluid and synovial fluid (Friães *et al*., 2013a). All the rest are considered non-invasive isolates including isolates recovered from pharyngitis, otitis, vaginitis and also from SSTI. Therefore, it is difficult to distinguish the composition of isolates from each source separately. Even within SSTI caused by GAS there are several entities, ranging from mild to severe. In our study, isolates were

recovered from purulent exudates and unfortunately, we do not have information regarding the specific syndrome caused by the isolates, the clinical evolution, and outcome of each infection. We consider this to be a major limitation of our studies because it prevents us from evaluating further correlations between specific clones in SSTI and the severity of their respective infections.

Macrolide resistance was detected in 10.3% of SSTI isolates and showing a decreasing trend similar to that found among contemporary invasive and pharyngitis isolates in Portugal (Silva-Costa *et al.*, 2015a). Moreover, these values are very similar to those reported for invasive isolates in Portugal between 2006 and 2009 (8%), (Friães *et al*., 2013a), suggesting similar values of macrolide resistance rates among isolates responsible different types of infection. Penicillin is the antimicrobial of choice for the treatment of GAS infections since naturally occurring resistant isolates were never reported (Walker *et al.*, 2014). Even so, in cases of patients with previously documented sensitivity reactions to β-lactams, macrolides are an important alternative that could be used. Moreover, in cases of severe invasive infections the use of a combination of clindamycin with β-lactam is recommended (Stevens *et al.*, 2005; Allen and Moore, 2010). In agreement with the results from Portugal, a decreasing trend in macrolide resistance was reported in several countries particularly in Europe, namely in the southern countries, such as Spain, Greece, and Italy (Silva-Costa *et al*., 2015a). Fluctuations in macrolide resistance rates are commonly associated with macrolide consumption. However, in some countries, including Portugal and Spain, the reported decreasing trend is not associated to lower macrolide consumption, suggesting that others factors that need to be further investigated may influence macrolide resistance (Silva-Costa *et al*., 2015a).

Among our 33 SSTI macrolide-resistant isolates, the majority presented the $cMLS_B$ phenotype (n=22), while 11 isolates presented the M phenotype. Due to these lower numbers and also due to the difference in the total number of isolates between the years of the study it is hard to evaluate the fluctuations of phenotypes. Even so, with the use of Cochran-Armitage statistical test to both phenotypes considering the 33 macrolide-resistant isolates as total, both phenotypes show a significant decreasing trend even with FDR, which was more significant for M isolates (p=0.001 for M isolates and P=0.006 for $cMLS_B$). Macrolide resistance was detected in isolates comprising 12 different *emm* types. Isolates displaying a $cMLS_B$ phenotype were distributed among 5 different *emm* types [SID (CI95%)=0.662 (0.491-0.834)], while

isolates with an M phenotype were distributed among 8 different *emm* types [SID (CI95%)=0.927(0.833-1.000)]. Additionally, all isolates with $cMLS_B$ phenotype were positive for the presence of the *ermB* gene while all isolates with an M phenotype were positive for the presence of the *mefA* gene, with the exception of one isolate that presented the *mefE* gene. These results are also in agreement with the most common genetic elements found in M and $cMLS_B$ isolates among pharyngeal and invasive isolates in Portugal (Friães *et al.*, 2013a; Silva-Costa *et al.*, 2015b).

Tetracycline resistance was found in 15% among SSTI isolates with no significant temporal trend. This result is also in agreement with the resistant rates described in invasive isolates between 2006-2009 (10.5%) (Friães *et al*., 2013a).

The comparison of our SSTI isolates with invasive isolates recovered from the same period previously characterized in our lab (Friães *et al.*, 2007; Friães *et al.*, 2013a), revealed significant differences regarding the *emm* type composition. While *emm*1, *emm*3, and *emm*64 isolates were significantly associated with invasive infections, *emm*89 isolates were significantly associated with SSTI infections and also the most prevalent among these. In a previous study from our lab with isolates recovered during 2000-2005, *emm* types 1 and 64 were considered as markers of invasiveness relative to pharyngeal isolates (Friães *et al*., 2012). Worldwide, *emm*1 isolates are the most common isolates recovered from invasive infections and have been widely associated with resurgence of invasive infections since the 1980s (Aziz and Kotb, 2008; Steer *et al.*, 2009; Smeesters *et al.*, 2009; Luca-Harari *et al.*, 2009; Henriet *et al.*, 2010; Imohl *et al.*, 2010; Olafsdottir *et al.*, 2014) including in Portugal (Friães *et al.*, 2007; Friães *et al.*, 2013a). On the contrary, the number of *emm*64 isolates recovered from invasive infections has decreased markedly in Portugal in the period 2006-2009, indicating that this clone did not persist (Friães *et al*., 2013a). The *emm*3 isolates are also frequently associated to invasive infections, both in a large European study, as well in Portugal, where it was the third most frequent *emm* type recovered from invasive infection (Luca-Harari *et al.*, 2009; Friães *et al.*, 2013a).

Regarding the *emm*89 isolates, a previous association with SSTI has been reported in Spain (Tamayo *et al.*, 2014). These isolates have also been reported associated with invasive infections (Darenberg *et al.*, 2007). Considering the distinct clades among *emm*89 isolates, only those lacking the *hasABC* locus were associated with SSTI. These isolates are presumed to belong to the recently emerged acapsular clade 3 that additionally to the absence of capsule also present an *nga-ifs-slo* locus variant

associated with the increased expression of NADase and SLO (Friães *et al.*, 2015; Zhu *et al.*, 2015). We also observed that the increase in the prevalence of *emm*89 among SSTI isolates was associated with the emergence of this clade, in contrast to isolates recovered from pharyngitis and invasive infections (Friães *et al*., 2015). Even so, this clade quickly replaced the previously circulating *emm*89 clades in all infection types. These results lead us to suggest that this clade has a higher propensity to cause SSTI. This could be due to several mechanisms such as, increased capacity to colonize the skin, an improved ability to overcome the major host defence mechanisms or the ability to produce cytotoxic effects at these sites. In fact, it was reported that clade 3 associated strains have an enhanced ability to adhere to plastic suggesting that these isolates may have an advantage in environmental persistence and transmission (Turner *et al.*, 2015). On the contrary, it was also reported that there was increased virulence in an acapsular strain with the variant 3 of the *nga* promoter in a mouse model of necrotizing fasciitis (Zhu *et al*., 2015). Therefore, additional surveillance and further investigation is crucial to define the relevance of this new *emm*89 acapsular clade in different types of infection.

In 2014, *emm*-clustering was described as a new typing method based on the portion of *emm* genes encoding the entire surface-exposed region of the M proteins (Sanderson-Smith *et al.*, 2014). According to the authors, the M protein types assigned to an individual *emm*-cluster have high sequence similarity sharing functional properties, including the ability to bind to an array of host proteins. The authors also suggest that *emm*-clusters and *emm*-patterns can be directly inferred from the *emm* type because *emm*-type is considered to be predictive of the whole protein sequence (McMillan *et al.*, 2013). Therefore, it is possible through *emm*-typing to predict the interaction of M protein with host proteins. We found that among our isolates, the *emm*-clusters and the possible interactions with host proteins reflected the differences between *emm*-types. While invasive isolates were significantly associated with *emm*-clusters A-C3 and A-C5, SSTI isolates were significantly associated to the *emm*-cluster E4. These differences can be explained by *emm*-clusters A-C3 and A-C5 grouping exclusively isolates of *emm* type 1 and 3, respectively, and *emm*-cluster E4 being dominated by *emm89* which is associated with SSTI.

After the attribution of a potential "binding profile" to each *emm*-cluster we observed that the ability to bind fibrinogen and albumin were associated to invasive isolates whereas the ability to bind C4BP and IgG were significantly associated to SSTI

isolates. Curiously, binding of protein M to either fibrinogen or C4BP were shown to reduce opsonophagocytosis due to inhibition of complement deposition on the bacterial surface (Carlsson *et al.*, 2003; Carlsson *et al.*, 2005). Thus it was proposed that all M proteins may share the ability to bind to fibrinogen or C4BP (Carlsson *et al*., 2005). However, the reasons why complement inhibition is achieved through fibrinogen in iGAS and C4BP recruitment in SSTI remain elusive. One possible reason is the released M1 protein-fibrinogen complexes due to the action of SpeB. These complexes were shown to interact with β_2 integrins on the surface of neutrophils, resulting in the production of heparin-binding protein (HBP), which has been suggested to be a potent inflammatory mediator that induce vascular leakage which could promote invasiveness (Herwald *et al.,* 2004). However, additionally to the M protein, other proteins also present on the GAS surface can bind fibrinogen (Courtney *et al.,* 2006). Moreover, binding to fibrinogen does not always inhibit the activation of the classical pathway and it has been related to the presence of other virulence factors which could also contribute to the invasive potential of the isolates (Dale *et al.*, 1996; Carlsson *et al.*, 2005; Horstmann *et al.*, 1992; Kotarsky *et al.*, 2000).

The association of cluster E4 with SSTI resulted in the overrepresentation of albumin binding among iGAS because most *emm*-clusters include proteins able to bind albumin, with the exception of cluster E4, which was associated to SSTI (Sanderson-Smith *et al*., 2014) It was suggested that the binding of fibrinogen and albumin to B- and C-repeats, respectively, inhibit antibody binding under physiological conditions (Sandin *et al.*, 2006). However, mutant strains lacking the conserved C-repeats were only slightly attenuated in a mouse intraperitoneal infection model (Waldemarsson *et al.*, 2009). Therefore, the importance of the ability of iGAS to bind albumin remains to be explored.

The binding ability to human immunoglobulin IgG was found to be significant among SSTI isolates. Immunoglobulin binding was shown to hinder opsonophagocytosis even in the absence of a specific immune response (Carlsson *et al*., 2003). However, the potential additionally benefits contributing to its overrepresentations among SSTI remain to be clarified.

Only 5% of the SSTI isolates were predicted to bind plasminogen and no differences were observed compared with invasive isolates. Several roles have been proposed regarding de deposition of plasminogen on the cell surface which has been related with streptokinase and SpeB and therefore with *covRS* mutations. So,

considering that *covRS* mutations contribute to invasiveness, it could be expected that more invasive isolates would bind to plasminogen in comparison to SSTI. However, several other cell surface proteins can bind plasminogen and we are only predicting the binding ability of the M protein. Moreover, plasminogen binding was also shown to facilitate keratinocyte invasion (Siemens *et al.*, 2011) which could be important in the context of SSTI.

Regarding the distribution of isolates among *emm*-patterns, while SSTI isolates were associated with *emm*-pattern E, invasive isolates were associated with *emm*-pattern A-C (both $p < 0.001$). We would expect that our SSTI isolates, which were recovered from skin, would belong mainly to *emm*-pattern D ("skin specialist"). Instead, the majority presented an *emm*-pattern E, considered "the generalist" (Bessen *et al.*, 1997). Possible explanations for this result include: that the GAS population responsible for SSTI has actually changed since the time the associations between *emm*-patterns and disease presentation were proposed or it could reflect differences between impetigo and other skin infections, given that the *emm*-pattern associations were established using only impetigo isolates.

Taking into account the antigens present in the new 30-valent M protein-based vaccine, 86% (n=274) of the SSTI infections and 89% (n=279) of invasive infections could have been prevented. Considering also the non-vaccine *emm* types against which it was demonstrated that the 30-valent vaccine evokes cross-opsonic antibodies, the potential coverage rises to 95% among SSTI isolates and invasive isolates (Sanderson-Smith *et al.*, 2014; Dale *et al.*, 2011; Dale *et al.*, 2013). This vaccine would be particularly important in developing countries, in order to prevent the non-supputarive complications of GAS, such as ARF and RHD. In developed countries, given the universal susceptibility of GAS to penicillin, this vaccine would be important to reduce the frequency of invasive infections that continue to cause significant morbidly and mortality despite appropriate and effective antimicrobial therapy. However, these are rare events so vaccination may be questioned in economic grounds.

The screening of SAgs genes revealed significant differences between isolates from both sources. While the presence of SAg gene *speA* and *speK* was significantly associated with invasive isolates, *speL* and *speM* were significantly associated with isolates from SSTI. Consistent with our results, several studies reported the association of *speA* with iGAS (Friães *et al.*, 2012; Lintges *et al.*, 2010; Creti *et al.*, 2005; Montes *et al.*, 2011). However, in the studies where a comparison was performed, the majority of

non-invasive isolates are often recovered from pharyngitis. The fact that we observed an association of *speA* with iGAS relative to SSTI supports a role for SpeA specifically in invasive infection. The association of *speK* (formerly speL) with invasive disease was mainly due to the presence of *speK* in the majority of *emm3* isolates. This association was previously reported (for the protein formerly known as SpeL) (Ikebe *et al.*, 2002). The significant association of *speL and speM* with non-invasive isolates was also reported in other studies (Friães *et al.*, 2012; Haukness *et al.*, 2002).

Currently, there are evidences that *emm* typing is not sufficient to identify GAS clones and that it must be complemented with other typing methods such as MLST and SAg profiling for further discrimination of genetic lineages within each *emm* type (Maripuu *et al.*, 2008; Carrico *et al.*, 2006; Friães *et al.*, 2013b; Lintges *et al.*, 2010; Meisal *et al.*, 2010). Not surprisingly, due to the close link between *emm* type and SAg profile, the significant differences found between invasive and SSTI isolates reflect the differences found in *emm*-types. Therefore, SAg profiles 8 and 10, which correspond mostly to *emm3* and *emm1*, respectively, were associated with invasive infection, while SAg profile 29, mostly found among *emm89*, was associated with SSTI infection.

In contrast to previous results regarding SAg profiles in invasive isolates in Portugal (Friães *et al*., 2013a), in this study, no diversification of SAg profiles was observed in any of the most frequent *emm* types among SSTI isolates. These results suggest that isolates responsible for skin and soft tissue infections are probably under a weaker selective pressure than invasive isolates. Similarly, regarding MLST, no diversification in STs was observed in invasive nor in SSTI isolates.

Despite the significant differences of prevalence of multiples *emm* types between SSTI and iGAS, we did not detect any significant distinct STs or SAgs profiles within each of the most prevalent *emm* types according to the type of infection, nor were there significant differences in diversity within isolates of each *emm* type between SSTI and iGAS, that could explain the ability to cause different infections.

Concluding remarks and future work.

Overall we found an association of null *covS* alleles with invasive isolates when comparing both with isolates recovered from pharyngitis and from SSTI. We also found, with few exceptions, that these isolates showed no SpeB activity and an upregulation of NADase and SLS, which supports the abrogation of CovS function as a

mechanism that favors the progression to invasive disease. Even so, these mutations were only found in a minority of invasive isolates (10%), which could partially be explained by an associated fitness cost resulting in impaired colonization and transmission. Additionally, SpeB activity was detected in most of the isolates tested and their proportion was similar in different types of infection, suggesting that others virulence factors under *covRS* regulation may be responsible for the higher virulence of iGAS isolates. Therefore, further studies evaluating and comparing the expression of virulence factors under the control of *covRS,* from invasive and non-invasive isolates, as well as of their activity will be key to explore which features confer a higher invasive potential to an isolate. Regarding isolates responsible for SSTI in Portugal they are a genetically diverse population and although this is the main primary focus associated to invasive disease, significant differences in the prevalence of the *emm* types were detected when compared with contemporary isolates recovered from invasive infection. However, considering isolates with the same *emm* type recovered from both sites we found the same MLST defined lineages and the same SAg profiles, not providing any clues about what distinguishes these isolates. Therefore more methods, such as whole genome sequence, must be used in order to detect genetic differences intra-*emm* type that could explain the tissue tropism of each isolate.

REFERENCES

Agarwal, S., P. Pancholi and V. Pancholi. 2011. Role of serine/threonine phosphatase (SP-STP) in *Streptococcus pyogenes* physiology and virulence. J Biol Chem. 286: 41368-41380.

Alam, F. M., C. E. Turner, K. Smith, S. Wiles and S. Sriskandan. 2013. Inactivation of the CovR/S Virulence Regulator Impairs Infection in an Improved Murine Model of *Streptococcus pyogenes* Naso-Pharyngeal Infection. PLoS One. 8: e61655.

Allen, U. and D. Moore. 2010. Invasive group A streptococcal disease: Management and chemoprophylaxis. Can J Infect Dis Med Microbiol. 21: 115-118.

Aziz, R. K., R. Kansal, B. J. Aronow, W. L. Taylor, S. L. Rowe, M. Kubal, G. S. Chhatwal, M. J. Walker and M. Kotb. 2010. Microevolution of group A streptococci in vivo: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. PLoS One. 5: e9798.

Aziz, R. K. and M. Kotb. 2008. Rise and persistence of global M1T1 clone of *Streptococcus pyogenes*. Emerg Infect Dis. 14: 1511-1517.

Aziz, R. K., M. J. Pabst, A. Jeng, R. Kansal, D. E. Low, V. Nizet and M. Kotb. 2004. Invasive M1T1 group A *Streptococcus* undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. Mol Microbiol. 51: 123-134.

Bessen, D. E., T. R. Fiorentino and S. K. Hollingshead. 1997. Molecular markers for throat and skin isolates of group A streptococci. Adv Exp Med Biol. 418: 537-543.

Bjorck, L., P. Akesson, M. Bohus, J. Trojnar, M. Abrahamson, I. Olafsson and A. Grubb. 1989. Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. Nature. 337: 385-386.

Carlsson, F., K. Berggard, M. Stalhammar-Carlemalm and G. Lindahl. 2003. Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. J Exp Med. 198: 1057-1068.

Carlsson, F., C. Sandin and G. Lindahl. 2005. Human fibrinogen bound to *Streptococcus pyogenes* M protein inhibits complement deposition via the classical pathway. Mol Microbiol. 56: 28-39.

Carrico, J. A., C. Silva-Costa, J. Melo-Cristino, F. R. Pinto, H. de Lencastre, J. S. Almeida and M. Ramirez. 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol. 44: 2524-2532.

Carroll, R. K., S. A. Shelburne, 3rd, R. J. Olsen, B. Suber, P. Sahasrabhojane, M. Kumaraswami, S. B. Beres, P. R. Shea, A. R. Flores and J. M. Musser. 2011. Naturally occurring single amino acid replacements in a regulatory protein alter

streptococcal gene expression and virulence in mice. J Clin Invest. 121: 1956-1968.

Cole, J. N., J. D. McArthur, F. C. McKay, M. L. Sanderson-Smith, A. J. Cork, M. Ranson, M. Rohde, A. Itzek, H. Sun, D. Ginsburg, M. Kotb, V. Nizet, G. S. Chhatwal and M. J. Walker. 2006. Trigger for group A streptococcal M1T1 invasive disease. FASEB J. 20: 1745-1747.

Connolly, K. L., A. L. Roberts, R. C. Holder and S. D. Reid. 2011. Dispersal of Group A streptococcal biofilms by the cysteine protease SpeB leads to increased disease severity in a murine model. PLoS One. 6: e18984.

Courtney, H. S., D. L. Hasty and J. B. Dale. 2006. Anti-phagocytic mechanisms of *Streptococcus pyogenes*: binding of fibrinogen to M-related protein. Mol Microbiol. 59: 936-947.

Creti, R., G. Gherardi, M. Imperi, C. von Hunolstein, L. Baldassarri, M. Pataracchia, G. Alfarone, F. Cardona, G. Dicuonzo and G. Orefici. 2005. Association of group A streptococcal emm types with virulence traits and macrolide-resistance genes is independent of the source of isolation. J Med Microbiol. 54: 913-917.

Cunningham, M. W. 2000. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 13: 470-511.

Dale, J. B., T. A. Penfound, E. Y. Chiang and W. J. Walton. 2011. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine. 29: 8175-8178.

Dale, J. B., T. A. Penfound, B. Tamboura, S. O. Sow, J. P. Nataro, M. Tapia and K. L. Kotloff. 2013. Potential coverage of a multivalent M protein-based group A streptococcal vaccine. Vaccine. 31: 1576-1581.

Dale, J. B., R. G. Washburn, M. B. Marques and M. R. Wessels. 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. Infect Immun. 64: 1495-1501.

Darenberg, J., B. Luca-Harari, A. Jasir, A. Sandgren, H. Pettersson, C. Schalen, M. Norgren, V. Romanus, A. Norrby-Teglund and B. H. Normark. 2007. Molecular and clinical characteristics of invasive group A streptococcal infection in Sweden. Clin Infect Dis. 45: 450-458.

Dmitriev, A. V., E. J. McDowell and M. S. Chaussee. 2008. Inter- and intraserotypic variation in the *Streptococcus pyogenes* Rgg regulon. FEMS Microbiol Lett. 284: 43-51.

Ekelund, K., J. Darenberg, A. Norrby-Teglund, S. Hoffmann, D. Bang, P. Skinhoj and H. B. Konradsen. 2005. Variations in emm type among group A streptococcal isolates causing invasive or noninvasive infections in a nationwide study. J Clin Microbiol. 43: 3101-3109.

Engleberg, N. C., A. Heath, A. Miller, C. Rivera and V. J. DiRita. 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis. 183: 1043-1054.

Engleberg, N. C., A. Heath, K. Vardaman and V. J. DiRita. 2004. Contribution of CsrR-regulated virulence factors to the progress and outcome of murine skin infections by *Streptococcus pyogenes*. Infect Immun. 72: 623-628.

Eriksson, A. and M. Norgren. 2003. Cleavage of antigen-bound immunoglobulin G by SpeB contributes to streptococcal persistence in opsonizing blood. Infect Immun. 71: 211-217.

Eriksson, B. K., J. Andersson, S. E. Holm and M. Norgren. 1999. Invasive group A streptococcal infections: T1M1 isolates expressing pyrogenic exotoxins A and B in combination with selective lack of toxin-neutralizing antibodies are associated with increased risk of streptococcal toxic shock syndrome. J Infect Dis. 180: 410-418.

Friães, A., J. P. Lopes, J. Melo-Cristino and M. Ramirez. 2013a. Changes in *Streptococcus pyogenes* causing invasive disease in Portugal: evidence for superantigen gene loss and acquisition. Int J Med Microbiol. 303: 505-513.

Friães, A., M. P. Machado, C. Pato, J. Carrico, J. Melo-Cristino and M. Ramirez. 2015. Emergence of the Same Successful Clade among Distinct Populations of emm89 *Streptococcus pyogenes* in Multiple Geographic Regions. MBio. 6: e01780-01715.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2012. Group A streptococci clones associated with invasive infections and pharyngitis in Portugal present differences in emm types, superantigen gene content and antimicrobial resistance. BMC Microbiol. 12: 280.

Friães, A., F. R. Pinto, C. Silva-Costa, M. Ramirez and J. Melo-Cristino. 2013b. Superantigen gene complement of *Streptococcus pyogenes*-relationship with other typing methods and short-term stability. Eur J Clin Microbiol Infect Dis. 32: 115-125.

Friães, A., M. Ramirez and J. Melo-Cristino. 2007. Nonoutbreak surveillance of group A streptococci causing invasive disease in Portugal identified internationally disseminated clones among members of a genetically heterogeneous population. J Clin Microbiol. 45: 2044-2047.

Graham, M. R., L. M. Smoot, C. A. Migliaccio, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, M. J. Federle, G. J. Adams, J. R. Scott and J. M. Musser. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. Proc Natl Acad Sci U S A. 99: 13855-13860.

Hasegawa, T., A. Okamoto, T. Kamimura, I. Tatsuno, S. N. Hashikawa, M. Yabutani, M. Matsumoto, K. Yamada, M. Isaka, M. Minami and M. Ohta. 2010. Detection of invasive protein profile of *Streptococcus pyogenes* M1 isolates from pharyngitis patients. APMIS. 118: 167-178.

Haukness, H. A., R. R. Tanz, R. B. Thomson, Jr., D. K. Pierry, E. L. Kaplan, B. Beall, D. Johnson, N. P. Hoe, J. M. Musser and S. T. Shulman. 2002. The heterogeneity of endemic community pediatric group a streptococcal pharyngeal isolates and their relationship to invasive isolates. J Infect Dis. 185: 915-920.

Henriet, S., F. Kaguelidou, P. Bidet, M. Lorrot, A. De Lauzanne, S. Dauger, F. Angoulvant, J. C. Mercier, C. Alberti, E. Bingen and A. Faye. 2010. Invasive group A streptococcal infection in children: clinical manifestations and molecular characterization in a French pediatric tertiary care center. Eur J Clin Microbiol Infect Dis. 29: 341-346.

Herwald, H., H. Cramer, M. Morgelin, W. Russell, U. Sollenberg, A. Norrby-Teglund, H. Flodgaard, L. Lindbom and L. Bjorck. 2004. M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. Cell. 116: 367-379.

Hollands, A., R. K. Aziz, R. Kansal, M. Kotb, V. Nizet and M. J. Walker. 2008. A naturally occurring mutation in ropB suppresses SpeB expression and reduces M1T1 group A streptococcal systemic virulence. PLoS One. 3: e4102.

Hollands, A., M. A. Pence, A. M. Timmer, S. R. Osvath, L. Turnbull, C. B. Whitchurch, M. J. Walker and V. Nizet. 2010. Genetic switch to hypervirulence reduces colonization phenotypes of the globally disseminated group A streptococcus M1T1 clone. J Infect Dis. 202: 11-19.

Horstmann, N., M. Saldana, P. Sahasrabhojane, H. Yao, X. Su, E. Thompson, A. Koller and S. A. Shelburne, 3rd. 2014. Dual-site phosphorylation of the control of virulence regulator impacts group a streptococcal global gene expression and pathogenesis. PLoS Pathog. 10: e1004088.

Horstmann, R. D., H. J. Sievertsen, M. Leippe and V. A. Fischetti. 1992. Role of fibrinogen in complement inhibition by streptococcal M protein. Infect Immun. 60: 5036-5041.

Ikebe, T., M. Ato, T. Matsumura, H. Hasegawa, T. Sata, K. Kobayashi and H. Watanabe. 2010. Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. PLoS Pathog. 6: e1000832.

Ikebe, T., A. Wada, Y. Inagaki, K. Sugama, R. Suzuki, D. Tanaka, A. Tamaru, Y. Fujinaga, Y. Abe, Y. Shimizu and H. Watanabe. 2002. Dissemination of the phage-associated novel superantigen gene speL in recent invasive and noninvasive *Streptococcus pyogenes* M3/T3 isolates in Japan. Infect Immun. 70: 3227-3233.

Imohl, M., R. R. Reinert, C. Ocklenburg and M. van der Linden. 2010. Epidemiology of invasive *Streptococcus pyogenes* disease in Germany during 2003-2007. FEMS Immunol Med Microbiol. 58: 389-396.

Kansal, R. G., V. Datta, R. K. Aziz, N. F. Abdeltawab, S. Rowe and M. Kotb. 2010. Dissection of the molecular basis for hypervirulence of an in vivo-selected phenotype of the widely disseminated M1T1 strain of group A *Streptococcus* bacteria. J Infect Dis. 201: 855-865.

Kansal, R. G., A. McGeer, D. E. Low, A. Norrby-Teglund and M. Kotb. 2000. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. Infect Immun. 68: 6362-6369.

Kapur, V., J. T. Maffei, R. S. Greer, L. L. Li, G. J. Adams and J. M. Musser. 1994. Vaccination with streptococcal extracellular cysteine protease (interleukin-1 beta convertase) protects mice against challenge with heterologous group A streptococci. Microb Pathog. 16: 443-450.

Kotarsky, H., A. Thern, G. Lindahl and U. Sjobring. 2000. Strain-specific restriction of the antiphagocytic property of group A streptococcal M proteins. Infect Immun. 68: 107-112.

Kuo, C. F., J. J. Wu, K. Y. Lin, P. J. Tsai, S. C. Lee, Y. T. Jin, H. Y. Lei and Y. S. Lin. 1998. Role of streptococcal pyrogenic exotoxin B in the mouse model of group A streptococcal infection. Infect Immun. 66: 3931-3935.

Lin, J. N., L. L. Chang, C. H. Lai, H. H. Lin and Y. H. Chen. 2014. Association between polymorphisms in the csrRS two-component regulatory system and invasive group A streptococcal infection. Eur J Clin Microbiol Infect Dis. 33: 735-743.

Lintges, M., M. van der Linden, R. D. Hilgers, S. Arlt, A. Al-Lahham, R. R. Reinert, S. Plucken and L. Rink. 2010. Superantigen genes are more important than the emm type for the invasiveness of group A *Streptococcus* infection. J Infect Dis. 202: 20-28.

Luca-Harari, B., J. Darenberg, S. Neal, T. Siljander, L. Strakova, A. Tanna, R. Creti, K. Ekelund, M. Koliou, P. T. Tassios, M. van der Linden, M. Straut, J. Vuopio-Varkila, A. Bouvet, A. Efstratiou, C. Schalen, B. Henriques-Normark and A. Jasir. 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. J Clin Microbiol. 47: 1155-1165.

Lukomski, S., C. A. Montgomery, J. Rurangirwa, R. S. Geske, J. P. Barrish, G. J. Adams and J. M. Musser. 1999. Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis of invasive skin infection and dissemination in mice. Infect Immun. 67: 1779-1788.

Lukomski, S., S. Sreevatsan, C. Amberg, W. Reichardt, M. Woischnik, A. Podbielski and J. M. Musser. 1997. Inactivation of *Streptococcus pyogenes* extracellular cysteine protease significantly decreases mouse lethality of serotype M3 and M49 strains. J Clin Invest. 99: 2574-2580.

Maamary, P. G., M. L. Sanderson-Smith, R. K. Aziz, A. Hollands, J. N. Cole, F. C. McKay, J. D. McArthur, J. K. Kirk, A. J. Cork, R. J. Keefe, R. G. Kansal, H. Sun, W. L. Taylor, G. S. Chhatwal, D. Ginsburg, V. Nizet, M. Kotb and M. J. Walker. 2010. Parameters governing invasive disease propensity of non-M1 serotype group A streptococci. J Innate Immun. 2: 596-606.

Maripuu, L., A. Eriksson and M. Norgren. 2008. Superantigen gene profile diversity among clinical group A streptococcal isolates. FEMS Immunol Med Microbiol. 54: 236-244.

Mayfield, J. A., Z. Liang, G. Agrahari, S. W. Lee, D. L. Donahue, V. A. Ploplis and F. J. Castellino. 2014. Mutations in the Control of Virulence Sensor Gene from *Streptococcus pyogenes* after Infection in Mice Lead to Clonal Bacterial

Variants with Altered Gene Regulatory Activity and Virulence. PLoS One. 9: e100698.

McMillan, D. J., P. A. Dreze, T. Vu, D. E. Bessen, J. Guglielmini, A. C. Steer, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2013. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. Clin Microbiol Infect. 19: E222-229.

Meisal, R., I. K. Andreasson, E. A. Hoiby, I. S. Aaberge, T. E. Michaelsen and D. A. Caugant. 2010. *Streptococcus pyogenes* isolates causing severe infections in Norway in 2006 to 2007: emm types, multilocus sequence types, and superantigen profiles. J Clin Microbiol. 48: 842-851.

Miyoshi-Akiyama, T., J. Zhao, T. Uchiyama, J. Yagi and T. Kirikae. 2009. Positive correlation between low adhesion of group A *Streptococcus* to mammalian cells and virulence in a mouse model. FEMS Microbiol Lett. 293: 107-114.

Montes, M., C. Ardanuy, E. Tamayo, A. Domenech, J. Linares and E. Perez-Trallero. 2011. Epidemiological and molecular analysis of *Streptococcus pyogenes* isolates causing invasive disease in Spain (1998-2009): comparison with non-invasive isolates. Eur J Clin Microbiol Infect Dis. 30: 1295-1302.

Neely, M. N., W. R. Lyon, D. L. Runft and M. Caparon. 2003. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. J Bacteriol. 185: 5166-5174.

Norrby-Teglund, A., K. Pauksens, S. E. Holm and M. Norgren. 1994. Relation between low capacity of human sera to inhibit streptococcal mitogens and serious manifestation of disease. J Infect Dis. 170: 585-591.

Olafsdottir, L. B., H. Erlendsdottir, J. Melo-Cristino, D. M. Weinberger, M. Ramirez, K. G. Kristinsson and M. Gottfredsson. 2014. Invasive infections due to *Streptococcus pyogenes*: seasonal variation of severity and clinical characteristics, Iceland, 1975 to 2012. Euro Surveill. 19: 5-14.

Olsen, R. J., D. R. Laucirica, M. E. Watkins, M. L. Feske, J. R. Garcia-Bustillos, C. Vu, C. Cantu, S. A. Shelburne, 3rd, N. Fittipaldi, M. Kumaraswami, P. R. Shea, A. R. Flores, S. B. Beres, M. Lovgren, G. J. Tyrrell, A. Efstratiou, D. E. Low, C. A. Van Beneden and J. M. Musser. 2012. Polymorphisms in regulator of protease B (RopB) alter disease phenotype and strain virulence of serotype M3 group A *Streptococcus*. J Infect Dis. 205: 1719-1729.

Olsen, R. J., A. Raghuram, C. Cantu, M. H. Hartman, F. E. Jimenez, S. Lee, A. Ngo, K. A. Rice, D. Saddington, H. Spillman, C. Valson, A. R. Flores, S. B. Beres, S. W. Long, W. Nasser and J. M. Musser. 2015. The Majority of 9,729 Group A *Streptococcus* Strains Causing Disease Secrete SpeB Cysteine Protease: Pathogenesis Implications. Infect Immun. 83: 4750-4758.

Rivera, A., M. Rebollo, E. Miro, M. Mateo, F. Navarro, M. Gurgui, B. Mirelis and P. Coll. 2006. Superantigen gene profile, emm type and antibiotic resistance genes among group A streptococcal isolates from Barcelona, Spain. J Med Microbiol. 55: 1115-1123.

Sanderson-Smith, M., D. M. De Oliveira, J. Guglielmini, D. J. McMillan, T. Vu, J. K. Holien, A. Henningham, A. C. Steer, D. E. Bessen, J. B. Dale, N. Curtis, B. W. Beall, M. J. Walker, M. W. Parker, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash and P. R. Smeesters. 2014. A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. J Infect Dis. 210: 1325-1338.

Sandin, C., F. Carlsson and G. Lindahl. 2006. Binding of human plasma proteins to Streptococcus pyogenes M protein determines the location of opsonic and non-opsonic epitopes. Mol Microbiol. 59: 20-30.

Shelburne, S. A., 3rd, C. Granville, M. Tokuyama, I. Sitkiewicz, P. Patel and J. M. Musser. 2005. Growth characteristics of and virulence factor production by group A *Streptococcus* during cultivation in human saliva. Infect Immun. 73: 4723-4731.

Siemens, N., N. Patenge, J. Otto, T. Fiedler and B. Kreikemeyer. 2011. *Streptococcus pyogenes* M49 plasminogen/plasmin binding facilitates keratinocyte invasion via integrin-integrin-linked kinase (ILK) pathways and protects from macrophage killing. J Biol Chem. 286: 21612-21622.

Silva-Costa, C., A. Friães, M. Ramirez and J. Melo-Cristino. 2015a. Macrolide-resistant Streptococcus pyogenes: prevalence and treatment strategies. Expert Rev Anti Infect Ther. 13: 615-628.

Silva-Costa, C., M. Ramirez and J. Melo-Cristino. 2015b. Declining macrolide resistance in *Streptococcus pyogenes* in Portugal (2007-13) was accompanied by continuous clonal changes. J Antimicrob Chemother. 70: 2729-2733.

Smeesters, P. R., D. J. McMillan, K. S. Sriprakash and M. M. Georgousakis. 2009. Differences among group A *Streptococcus* epidemiological landscapes: consequences for M protein-based vaccines? Expert Rev Vaccines. 8: 1705-1720.

Steer, A. C., I. Law, L. Matatolu, B. W. Beall and J. R. Carapetis. 2009. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. Lancet Infect Dis. 9: 611-616.

Stevens, D. L., A. L. Bisno, H. F. Chambers, E. D. Everett, P. Dellinger, E. J. Goldstein, S. L. Gorbach, J. V. Hirschmann, E. L. Kaplan, J. G. Montoya and J. C. Wade. 2005. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. Clin Infect Dis. 41: 1373-1406.

Sugareva, V., R. Arlt, T. Fiedler, C. Riani, A. Podbielski and B. Kreikemeyer. 2010. Serotype- and strain- dependent contribution of the sensor kinase CovS of the CovRS two-component system to *Streptococcus pyogenes* pathogenesis. BMC Microbiol. 10: 34.

Sumby, P., A. R. Whitney, E. A. Graviss, F. R. DeLeo and J. M. Musser. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog. 2: e5.

Tamayo, E., M. Montes, J. M. Garcia-Arenzana and E. Perez-Trallero. 2014. *Streptococcus pyogenes emm*-types in northern Spain; population dynamics over a 7-year period. J Infect. 68: 50-57.

Trevino, J., N. Perez, E. Ramirez-Pena, Z. Liu, S. A. Shelburne, 3rd, J. M. Musser and P. Sumby. 2009. CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A *Streptococcus* virulence factor-encoding genes. Infect Immun. 77: 3141-3149.

Turner, C. E., J. Abbott, T. Lamagni, M. T. Holden, S. David, M. D. Jones, L. Game, A. Efstratiou and S. Sriskandan. 2015. Emergence of a New Highly Successful Acapsular Group A *Streptococcus* Clade of Genotype *emm89* in the United Kingdom. MBio. 6: e00622.

Ulrich, R. G. 2008. Vaccine based on a ubiquitous cysteinyl protease and streptococcal pyrogenic exotoxin A protects against *Streptococcus pyogenes* sepsis and toxic shock. J Immune Based Ther Vaccines. 6: 8.

Wajima, T., S. Y. Murayama, K. Sunaoshi, E. Nakayama, K. Sunakawa and K. Ubukata. 2008. Distribution of emm type and antibiotic susceptibility of group A streptococci causing invasive and noninvasive disease. J Med Microbiol. 57: 1383-1388.

Waldemarsson, J., M. Stalhammar-Carlemalm, C. Sandin, F. J. Castellino and G. Lindahl. 2009. Functional dissection of *Streptococcus pyogenes* M5 protein: the hypervariable region is essential for virulence. PLoS One. 4: e7279.

Walker, M. J., T. C. Barnett, J. D. McArthur, J. N. Cole, C. M. Gillen, A. Henningham, K. S. Sriprakash, M. L. Sanderson-Smith and V. Nizet. 2014. Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. Clin Microbiol Rev. 27: 264-301.

Walker, M. J., A. Hollands, M. L. Sanderson-Smith, J. N. Cole, J. K. Kirk, A. Henningham, J. D. McArthur, K. Dinkla, R. K. Aziz, R. G. Kansal, A. J. Simpson, J. T. Buchanan, G. S. Chhatwal, M. Kotb and V. Nizet. 2007. DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat Med. 13: 981-985.

Zhu, L., R. J. Olsen, W. Nasser, I. de la Riva Morales and J. M. Musser. 2015. Trading Capsule for Increased Cytotoxin Production: Contribution to Virulence of a Newly Emerged Clade of emm89 *Streptococcus pyogenes*. MBio. 6: e01378-01315.

APPENDIX I

SCIENTIFIC REPORTS

OPEN

Consequences of the variability of the CovRS and RopB regulators among *Streptococcus pyogenes* causing human infections

Received: 20 March 2015
Accepted: 05 June 2015
Published: 15 July 2015

Ana Friães*, Catarina Pato*, José Melo-Cristino & Mario Ramirez

To evaluate the importance of *covRS* and *ropB* mutations in invasive disease caused by Group A Streptococci (GAS), we determined the sequence of the *covRS* and *ropB* genes of 191 isolates from invasive infections and pharyngitis, comprising a diverse set of *emm* types and multilocus sequence types. The production of SpeB and the activity of NAD glycohydrolase (NADase) and streptolysin S (SLS) were evaluated. The results support the acquisition of null *covS* alleles (predicted to eliminate protein function), resulting in downregulation of SpeB and upregulation of NADase and SLS, as a mechanism possibly contributing to higher invasiveness. Among the isolates tested, this mechanism was found to be uncommon (10% of invasive isolates) and was not more prevalent among clones with enhanced invasiveness (including M1T1) but occurred in diverse genetic backgrounds. In lineages such as *emm*64, these changes did not result in upregulation of NADase and SLS, highlighting the diversity of regulatory pathways in GAS. Despite abrogating SpeB production, null alleles in *ropB* were not associated with invasive infection. The *covRS* and *ropB* genes are under stabilising selection and no expansion of isolates carrying null alleles has been observed, suggesting that the presence of these regulators is important for overall fitness.

Streptococcus pyogenes (Group A Streptococci, GAS) is a human pathogen that can asymptomatically colonize the oropharynx, but is also responsible for a variety of human diseases, ranging from uncomplicated superficial infections of the respiratory tract and skin, such as pharyngitis and impetigo, to severe invasive infections associated with high morbidity and mortality, like necrotizing fasciitis and streptococcal toxic shock syndrome (STSS)[1].

The recognition of the ability of strains belonging to the highly invasive M1T1 clone to acquire mutations in the *covRS* genes during skin and soft tissue infection in mice suggested that GAS could increase its capacity to invade deeper tissues by altering its regulatory networks in order to produce a switch to an invasive transcriptome profile[2,3]. The adjacent and cotranscribed genes *covR* and *covS* encode the two-component regulatory system CovRS (also known as CsrRS). The DNA-binding response regulator CovR acts mostly as a repressor of transcription upon phosphorylation by the sensor kinase/phosphatase CovS, which responds to stress factors, such as elevated temperatures, high saline concentrations, and decreased pH[4]. CovRS is estimated to directly or indirectly influence the expression of 10–15% of the GAS genome. Although it is also involved in the regulation of metabolic genes, it is mostly known for controlling the expression of a large number of genes encoding factors that promote GAS virulence and evasion of the host immune system[5,6]. Most studies indicate that mutations impairing CovRS function result in the upregulation of genes encoding the hyaluronic acid capsule, streptolysins S and O, streptokinase, DNases, the interleukin-protease SpyCEP, and NAD-glycohydrolase (NADase), among others, and

Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.R. (email: ramirez@fm.ul.pt)

Gene	No. alleles (this study)	No. amino acid sequences (this study)	No. alleles with indels	No. alleles with nonsense mutations	Nucleotide diversity	*Ka*/*Ks*	*P* value
covR	54 (19)	32 (5)	1	2	0.0058	0.19	0.0032
covS	121 (43)	83 (26)	28[a]	11	0.0043	0.14	0.0029
ropB	78 (37[c])	56 (21[c])	7[b]	2	0.0055	0.23	0.0015

Table 1. Type of amino acid alteration, nucleotide diversity, and *Ka*/*Ks* values associated with the *covR*, *covS*, and *ropB* alleles described in this study and those previously reported in GenBank. [a]Three of the *covS* indels are in frame. [b]Two of the *ropB* indels are in frame, but one of these deletes a considerable portion of the protein (119 residues). [c]Excluding the isolate with a complete *ropB* deletion.

in the downregulation of proteins like the extracellular cysteine protease SpeB and the protein-G-related α2-macroglobulin-binding protein GRAB[7,8]. However, contrasting results have been reported regarding the effect of *covRS* mutations in some of those virulence factors, which may be partly explained by the interaction of CovRS with other transcriptional regulators, resulting in complex regulation patterns that can vary between different strains[8,9].

Mutations in CovRS and in the stand-alone transcriptional regulator RopB have been identified in isolates recovered from human infections[7,10–12]. RopB (also known as Rgg) is encoded by the *ropB* (*rgg*) gene, which is located 940 bp away from *speB*, in the opposite DNA strand, and directly binds the promoter of the latter gene to activate its transcription[13]. In some strains, this regulator has been reported to affect the transcription of other virulence factors, including DNases, the hyaluronic acid capsule, NADase, streptokinase, streptolysins, and phage-encoded superantigens, among others[10,11].

The downregulation of SpeB, which is usually observed as a consequence of CovRS and RopB mutations, is considered to be a determining step in the transition to an invasive phenotype, since this potent protease degrades several extracellular GAS virulence factors that play an important role in the invasive process, including the M protein, the F1 protein, C5a peptidase, streptokinase, and SmeZ[14]. In agreement, SpeB production has been inversely correlated with disease severity, in both human infections and murine models[15,16]. SpeB is initially synthesised and secreted as an inactive 42-kDa zymogen, which is then autocatalytically processed through a series of intermediates to form the 28-kDa mature active protease. Each stage from transcription to mature SpeB involves tight regulatory controls, including post-transcriptional and post-translational mechanisms[14].

Some studies suggest that *covRS* and *ropB* mutations occur more frequently among isolates from invasive infections, while others found them in a similar proportion among non-invasive isolates[7,11,12]. Contrasting observations have also been reported regarding the impact of alterations in these regulators on virulence using animal models, as well as on the expression of virulence factors like SpeB and the streptolysins[2,3,5,7,8,11,17–20]. Given these discrepancies and the fact that most of the studies have focused on particular *emm* types, especially *emm*1 and *emm*3, the importance of the acquisition of these mutations for the overall invasiveness and virulence of GAS strains is still not completely understood.

In order to address these questions, we sequenced the *covRS* and *ropB* genes in a collection of 191 GAS isolates presenting a high diversity of *emm* types and multilocus sequence types (STs), including the most prevalent clones causing pharyngitis and invasive infections in Portugal and in a majority of other countries from temperate climate regions[21–23]. The genetic diversity of the genes was evaluated, and the respective alleles were correlated with the activity of SpeB and of two other virulence factors whose expression has been suggested to be under the influence of CovRS and RopB, at least in some strains, namely the NADase and the streptolysin S (SLS)[7,11,18,24].

Results

Genetic variation of the *covR*, *covS* and *ropB* genes. In the present work, the sequence of the *covR*, *covS*, and *ropB* genes was determined for a collection of 191 isolates (Supplementary Table S1) comprising one third of the isolates of each *emm* type present in a larger collection of strains recovered from pharyngitis and invasive infections in Portugal, which has been characterized elsewhere[21]. In addition, all isolates of *emm* types 1 and 64 were included in this study, since these two *emm* types were shown to be significantly associated with invasive disease.

The three genes presented a high allelic diversity in the studied GAS collection, with Simpson's index of diversity (SID) values close to the ones obtained for *emm* type and ST (Table 2). The allelic diversity of the *covR* genes was lower than that of *covS* ($P = 0.0140$) and of the ST ($P = 0.0137$). Overall, isolates sharing the same *emm* type or ST frequently shared the same *covR*, *covS*, and *ropB* alleles (Supplementary Table S2). This association was stronger for *covR*, in line with the lower allelic diversity of this gene. These results indicate that the *covR*, *covS*, and *ropB* alleles present in a given isolate were

Gene	All isolates ($n=191$)		Invasive isolates ($n=87$)		Pharyngitis isolates ($n=104$)	
	No. partitions	SID [$CI_{95\%}$]	No. partitions	SID [$CI_{95\%}$]	No. partitions	SID [$CI_{95\%}$]
covR	19	0.831 [0.793–0.870]	17	0.816 [0.752–0.880]	13	0.839 [0.794–0.885]
covS	43	0.893 [0.861–0.925]	31	0.870 [0.811–0.930]	25	0.905 [0.873–0.938]
ropB	38	0.850 [0.805–0.896]	23	0.806 [0.729–0.883]	29	0.879 [0.828–0.929]
emm type	26	0.877 [0.844–0.910]	20	0.833 [0.769–0.898]	20	0.900 [0.868–0.932]
ST	41	0.894 [0.861–0.927]	27	0.860 [0.798–0.922]	32	0.910 [0.876–0.945]

Table 2. Simpson's index of diversity (SID) and corresponding 95% confidence intervals ($CI_{95\%}$) of the alleles of *covR*, *covS*, and *ropB*, the *emm* types, and the STs identified among the 191 GAS isolates analysed in this study.

essentially clonal properties, although some level of intra-clonal diversity was observed, particularly for *covS* and *ropB*.

Together with the alleles previously deposited in GenBank, we identified a total of 54 alleles for *covR*, 121 for *covS*, and 78 for *ropB* (Table 1). One of the isolates in our study is devoid of the *ropB* gene due to a previously characterized deletion that also involves the *speB* gene[25]. Alterations relative to the nucleotide sequence of the three genes in strain SF370 are depicted in Fig. 1. For each of the genes, more than half of the alleles result in alterations in the amino acid sequence, with *ropB* presenting the highest proportion of distinct amino acid sequences relative to the total number of alleles. However, all the *K*a/*K*s ratios were significantly lower than 1 (Table 1), indicating that all three genes are under stabilising selection. *K*a/*K*s is the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site. This measure is often used to evaluate deviations from neutral evolution. A value of *K*a/*K*s $<<$ 1 indicates that the protein is under stabilising selection, i.e. selection favours alleles that do not change the amino acid sequence. In contrast, a value of *K*a/*K*s $>>$ 1 is strong evidence for positive selection, indicating that selection drove changes to the protein[26]. Nucleotide diversity was also similar for the three genes, with differences in the total number of alleles basically reflecting the different sizes of the genes (687 bp for *covR*, 1503 bp for *covS*, and 843 bp for *rop*B). Null alleles were defined as those presenting changes predicted to result in a complete absence of protein function – including nonsense mutations, indels that result in frameshifts, large in-frame indels ($\geq$357 bp), and complete gene deletions. In spite of similar diversity, the *covS* gene presented more null alleles than *covR* ($P=0.0004$).

Mutations in *covR*, *covS*, and *ropB*, and association with *emm* type and invasiveness. For *covR* and *ropB*, the reference alleles (those present in SF370) were the most common among the 191 GAS isolates analysed in this study [$n=67$ (35%), and $n=69$ (36%), respectively]. The SF370 *covS* allele was present in only five isolates, but 68 isolates (36%) presented only synonymous nucleotide changes (15 alleles). Since our collection was enriched in *emm*1 isolates, the most common *covS* allele (*covS*-02, 28%) included a missense mutation characteristic of the M1T1 clone, namely I332 V. Other missense mutations were also commonly identified in isolates of specific *emm* types, in both *covS* and *ropB* genes.

In addition to the coding sequence of the genes, the region upstream of *covR*, which includes the *covRS* promoter, was also analysed for all isolates, but no nucleotide changes were found in the identified -35 and -10 regions, in the transcription start site, nor in the consensus sequences that have been identified as binding targets for CovR auto-repression[27].

The overall distribution of *emm* types differed significantly between the isolates carrying missense or null alleles of *covR*, *covS*, or *ropB* and those with no amino acid alterations ($P<0.0001$). Null alleles were not associated with any particular *emm* type, occurring in isolates of seven different *emm* types for *covS* (SID = 0.909 [$CI_{95\%}$ 0.822–0.996]) and of four different *emm* types in the case of *ropB* (SID = 0.867 [$CI_{95\%}$ 0.738–0.995]). Missense mutations in *covS* and *ropB* were significantly more prevalent in isolates of certain *emm* types due to the presence of clonal alleles in specific lineages, i.e. alleles carrying amino acid changes that occur in several isolates of that *emm* type (Supplementary Table S1).

According to the SID values, the allele diversity of the three genes was similar for the subsets of isolates associated with pharyngitis and with invasive infections (Table 2). The overall distribution of *covS* and *ropB* alleles was significantly different between the two subsets ($P<0.0001$), but no specific *covR*, *covS*, or *ropB* allele presented a significant association with either infection type after correcting for the false discovery rate (FDR), except for *ropB*-23, which was significantly associated with invasive infections ($P=0.0109$). This association is explained by the fact that *ropB*-23 was present in all *emm*64 isolates and not found in any other *emm* type (Supplementary Table S1), and *emm*64 was significantly associated with invasive infections in Portugal[21]. This allele presents only two synonymous nucleotide changes, which are not likely to be the cause of the high invasiveness of this lineage.

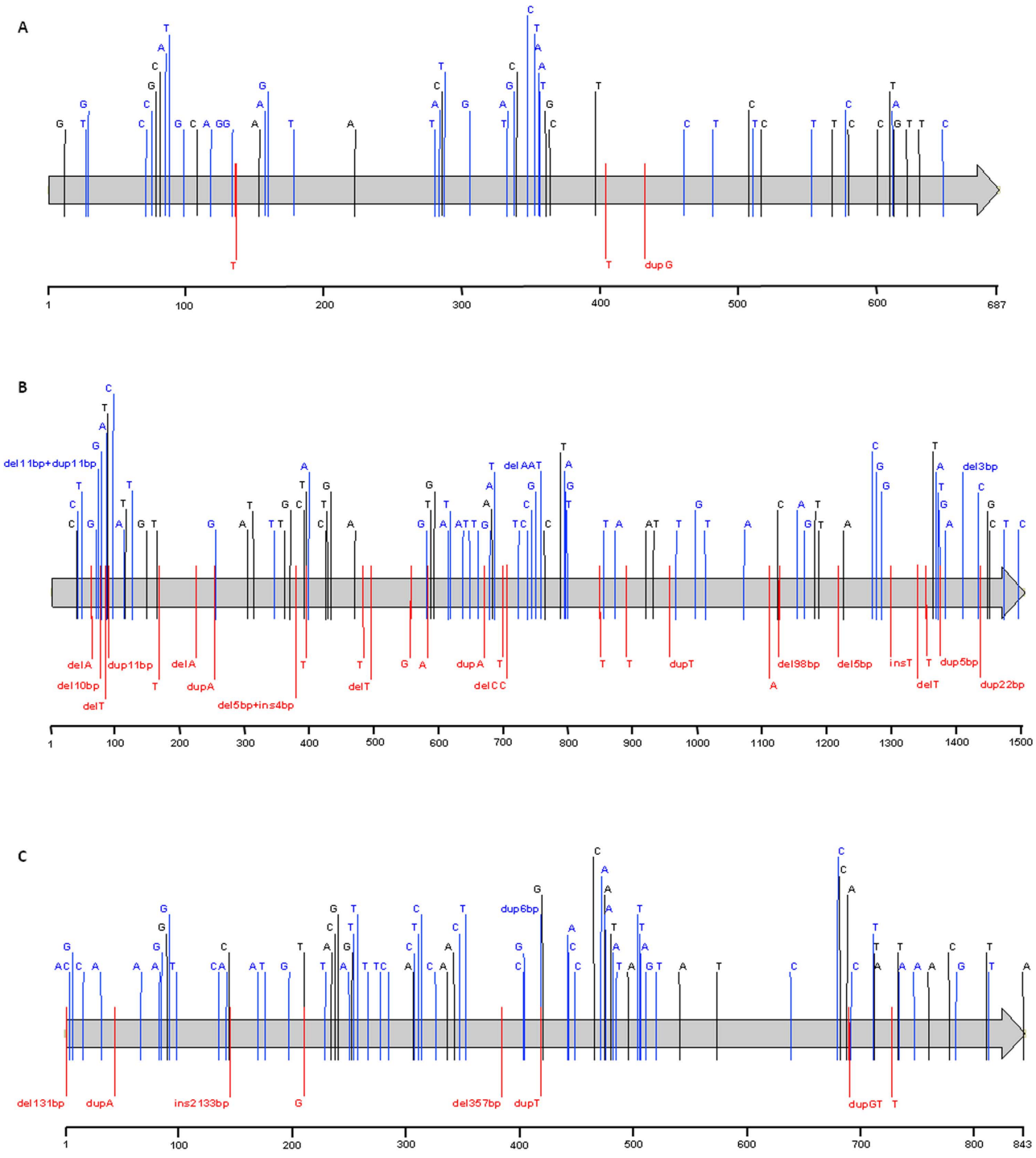


Figure 1. Nucleotide alterations identified in the *covR* (A), *covS* (B), and *ropB* (C) alleles reported in this study and previously deposited in GenBank, relative to the alleles present in strain SF370 (AE004092). Each gene is indicated by a grey arrow and a nucleotide numbering scale is represented below each one. Each sequence variation is indicated in the respective nucleotide position by a letter corresponding to the variant nucleotide, or the sequence or number of base-pairs inserted (ins), deleted (del), or duplicated (dup). Synonymous nucleotide changes are represented in black; missense mutations and short in-frame indels ($\leq$6 bp) are represented in blue; changes predicted to result in null alleles, including nonsense mutations, indels that generate frameshifts, and long in-frame indels ($\geq$357 bp) are represented in red.

	Allele type	Invasive ($n=87$)	Pharyngitis ($n=104$)
covR	Null	0	0
	Missense	2	2
covS	Null	9	2
	Missense	51	61
ropB	Null	3	3
	Missense	13	26
covRS/ropB	Null	12	5
	Missense	56	72

Table 3. Number of isolates from invasive infections and pharyngitis presenting null and missense alleles in the *covR*, *covS*, and *ropB* genes.

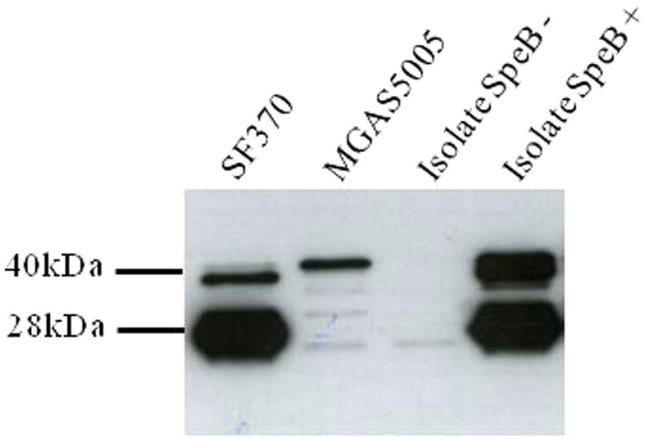


Figure 2. Representative Western blot result for detection of mature SpeB (28 kDa) expression by GAS isolates. The blot shown was cropped to the region of interest.

The presence of an altered amino acid sequence in general was not associated with infection type for any of the genes. However, null *covS* alleles were significantly overrepresented among the invasive isolates ($P=0.0247$) (Table 3), even though they were only present in nine of these isolates (10%).

SpeB production. In this study, the presence or absence of extracellular proteolytic activity was tested for all isolates using a spectrophotometric assay based on the degradation of azocasein by culture supernatants, as well as by a plate assay in which the strains were cultured in solid medium containing casein. To confirm that the proteolytic activity detected by these two methods was essentially due to SpeB, which has been described as the major extracellular protease of GAS[28], Western blot analysis using monoclonal anti-SpeB antibodies was performed in a subset of six isolates of *emm* types 1, 64 and 89. For each *emm* type, one protease-positive and one protease-negative isolate were randomly chosen. The presence of a 28-kDa band similar to the one detected for SF370 was considered as a positive result (Fig. 2). The Western blot results were concordant with those of the two proteolytic activity assays for this subset of isolates. The *emm*1 strains SF370 and MGAS5005 were used as controls in all assays. As expected, strain SF370, which encodes a functional CovRS, presented proteolytic activity in both assays and was positive for mature SpeB production by Western blotting. In contrast, almost undetectable amounts of mature SpeB and no proteolytic activity were observed for MGAS5005, which harbours a frameshift mutation in *covS* (Fig. 2 and Fig. 3). Taken together, these results confirm that SpeB is the major protease of GAS, allowing us to use the total proteolytic activity as a proxy for SpeB activity. Western blot was also performed for all isolates in which the azocasein and casein-plate results were discordant ($n=31$). In these cases, the Western blot result was considered as final.

Of a total of 191 isolates, 153 (80%) presented SpeB activity, while for the remaining 38 (20%) no SpeB activity could be detected (Fig. 3 and Supplementary Table S1). All isolates with null alleles, either in *covS* or in *ropB*, were associated with an absence of SpeB activity ($P<0.0001$ for both). In contrast, the prevalence of missense alleles among the SpeB-negative and SpeB-positive isolates was similar, except for non-clonal missense alleles in *ropB* (i.e. missense alleles that were not identified in other isolates of the same *emm* type in this study or in the GenBank database), which were associated with the absence of

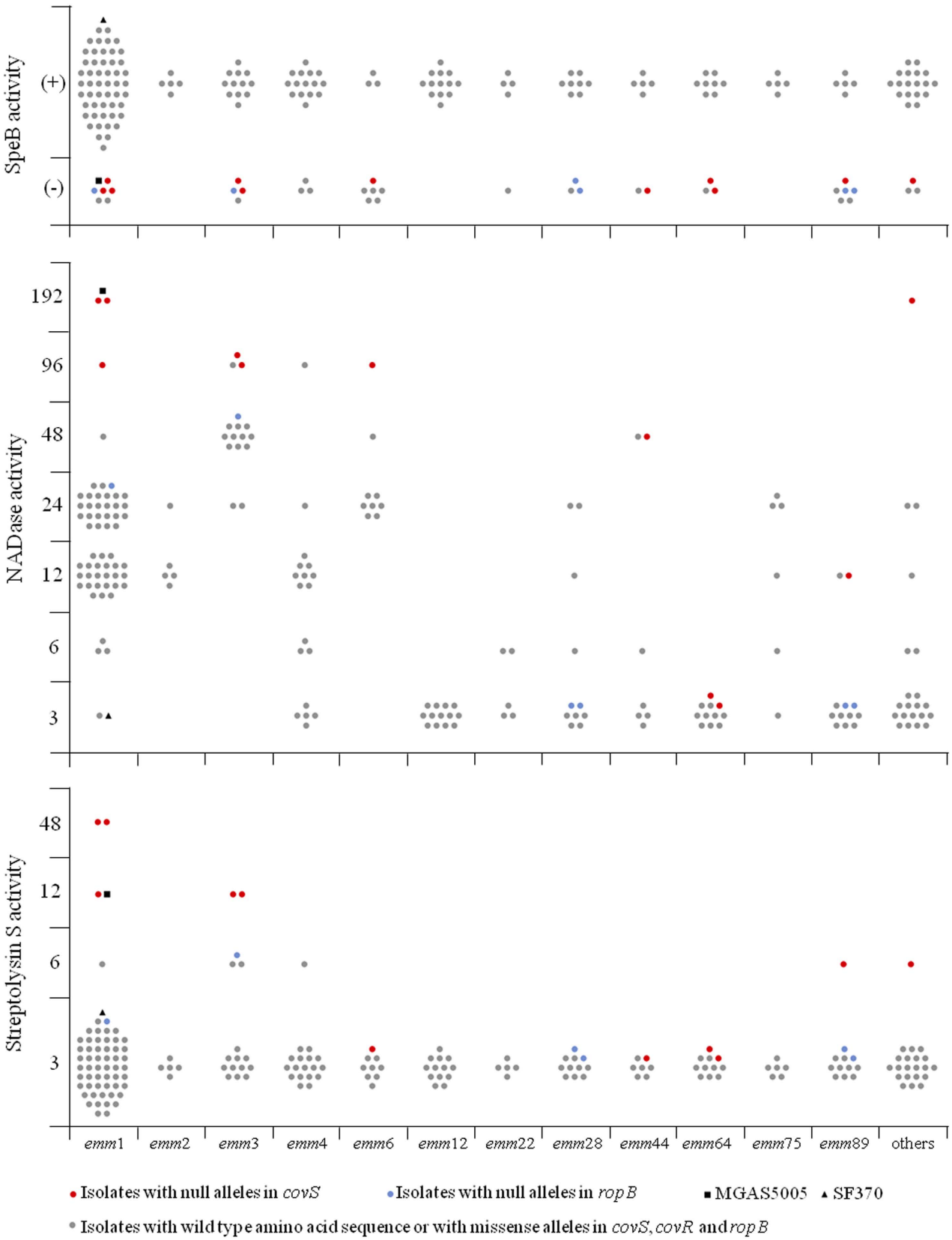


Figure 3. SpeB production and NADase and SLS activity determined for the 191 isolates analysed in this study, according to *emm* type. The *emm* types represented by < 5 isolates are grouped into "Others" and include *emm* 9, 11, 18, 29, 43, 53, 58, 74, 77, 78, 87, 94, 102 and 113. The results of reference strains SF370 and MGAS5005 are also presented.

SpeB production ($P = 0.0008$). Only two SpeB-negative isolates presented alleles encoding the reference amino acid sequence in all three genes. The absence of SpeB activity in isolates carrying non-null alleles in the three genes could be due to alterations in the *speB* gene, its promoter, or in other mechanisms involved in the post-transcriptional or post-translational regulation of SpeB[14].

There was no significant association between the SpeB activity of the isolates and infection type. The overall distribution of *emm* types differed significantly between SpeB-producing isolates and those without protease activity ($P = 0.0030$), but after FDR correction, only *emm* types 89 and 6 were significantly associated with the absence of SpeB ($P = 0.0354$ for both).

NADase activity. In this work, the NADase activity of the isolates was measured by an endpoint titre method based on the degradation of β-NAD by culture supernatants, which results in a reduction of the fluorescence emitted by the reduced form of β-NAD. GAS *emm*1 strains SF370 and MGAS5005 were used as controls of the assay. As expected, SF370 presented a low NADase activity (NADase ≤ 3), while MGAS5005 was found to have the highest NADase activity level measured (NADase = 192) (Fig. 3). The high NADase activity of the latter strain can be attributed to the null *covS* allele, as well as to a different NADase locus from the one carried by old *emm*1 strains such as SF370[29].

The gene encoding the GAS NADase, known as *spn* or *nga*, presents multiple variants and has been shown to be diverging into NADase-active and -inactive subtypes, which are correlated with *emm* patterns and tissue tropism[30]. Therefore, variations in the levels of NADase activity exhibited by distinct GAS lineages, which may encode different *nga* alleles, are expected and were observed among the isolates analysed in this study (Fig. 3 and Supplementary Table S1). Despite this lineage-specific variation, it was possible to identify a significant association between the highest activity values, namely NADase = 96 and NADase = 192, and the presence of null alleles in *covS* ($P = 0.0005$), while *covR* and *ropB* changes did not significantly contribute to an increased NADase activity. In agreement, all the strains carrying *covS* null alleles expressed higher levels of NADase activity than the majority of the isolates of the same *emm* type, except for two *emm*64 isolates (Fig. 3). The absence of NADase activity could be an intrinsic characteristic of this lineage, regardless of the CovRS alleles, due to alterations in the *nga* gene, its promoter, or other regulatory pathways. It is also possible that *nga* is not part of the CovRS regulon in the *emm*64 clone. Both inter- and intra-serotype differences in the regulatory activity of CovRS have been proposed[9].

Despite the association between high levels of NADase activity and *covS* null mutations, it was not possible to detect any significant association between the level of NADase activity and invasive disease.

Streptolysin S activity. The SLS activity of the isolates was evaluated by an endpoint titre method based on the amount of released haemoglobin from sheep erythrocytes during incubation with culture supernatants. Although the majority of the isolates ($n = 179$, 94%) presented an SLS activity ≤ 3, it was possible to identify multiple isolates with increased SLS activity and that also carried alterations in CovRS (Fig. 3). Accordingly, the SLS activity determined for the reference strains SF370 (SLS ≤ 3) and MGAS5005 (SLS = 12) were consistent with a de-repression of the *sag* operon in the latter due to the null allele in *covS*, and a significant association between activity values of SLS = 12 and SLS = 48 and *covS* null alleles was observed ($P = 0.0003$ and 0.0040, respectively). In fact, all isolates with an SLS ≥ 12 presented null alleles in *covS* (Fig. 3 and Supplementary Table S1). Four isolates carrying null alleles and belonging to *emm* types 6, 44, and 64 did not exhibit an increased SLS activity. The coherent low NADase and SLS activity among *emm*64 isolates independent of the CovRS alleles is consistent with a potentially different CovRS regulon in these isolates. Among *emm*6 and *emm*44 isolates, it is possible that the *sag* operon, which encodes SLS activity, is not regulated by CovRS. Alternatively, SLS activity may not be significantly increased in these isolates due to mutations in the *sag* operon, which have been infrequently reported among *S. pyogenes* isolates[31].

It was not possible to detect any significant associations of SLS activity levels with the presence or absence of changes in *covR* and *ropB*, nor with infection type.

Discussion

Variations at the level of the regulatory networks governing GAS gene expression may constitute key elements differentiating strains with a high invasive ability from those that carry the same virulence genes, but cause mild infections or asymptomatically colonise the host. Mutations impairing the function of the two-component system CovRS and the stand-alone regulator RopB could play this role and have been shown to be important in animal models of infection, mostly using representatives of the M1T1 clone[2,3,11,17]. However, the actual impact of these mechanisms in invasive human GAS infections remains unclear.

In the characterised collection, all three genes (*covR*, *covS*, and *ropB*) presented a high genetic diversity and the allele distribution was closely associated with *emm* type and ST. Some studies, mostly based on murine infection models, suggest that isolates of the highly invasive M1T1 clone have a higher ability to acquire *covRS* mutations than strains of other *emm* types[17,32]. However, in our study mutations in the *covR*, *covS*, and *ropB* genes were not more prevalent in *emm*1 or any other *emm* type. We conclude that

among the GAS population causing human infections, the acquisition of mutations in any of the three genes can occur in isolates of diverse lineages, in agreement with data from Japan and Taiwan[11,12]. We also did not find a higher prevalence of *covRS* mutations in isolates carrying the DNase gene *sda1* (data not shown), although it has been suggested that the presence of the phage encoding this gene exerts a selective pressure that favours the acquisition of *covRS* mutations, at least in the M1T1 genetic background[33].

Null *covS* alleles were found to have a significant association with invasive infections, while the missense alleles in *covS* and all null and missense alleles in *covR* and *ropB* were not differently distributed among isolates from pharyngitis and from invasive disease. According to these results, only the null alleles in *covS* would contribute to the transition to invasive infection. In addition, the fact that null alleles are significantly more common in *covS* than in *covR* supports the hypothesis that during the infection process, the acquisition of mutations that impair signalling through CovS, while keeping CovR functional and possibly responsive to phosphorylation by other kinases, such as SP-STK, may favour the progression to invasive infection[19,34]. However, this mechanism of transition to an invasive phenotype seems to be uncommon, since it was found in a minority of invasive isolates (10%).

The absence of expansion of isolates carrying null alleles in any of the three analysed genes and the fact that they are all under stabilising selection suggests that although mutations compromising the activity of these regulators may favour the progression to invasive disease, they are not beneficial to the overall fitness of *S. pyogenes*. Strains carrying *covRS* null alleles presented decreased ability to survive in human saliva and to persist in the murine nasopharynx, leading to the proposal that these strains would have an impaired colonisation and transmission capacity[19,35]. In a murine skin and soft tissue infection model, infection with a mixed population containing both the wild-type strain and a derivative *covRS* mutant resulted in higher virulence than infection with either strain alone, highlighting the importance of the presence of isolates carrying a functional CovRS[2]. Therefore, in spite of promoting a switch to a phenotype that favours the survival of GAS in deeper tissues, the loss of CovS may affect the success of the initial stages of infection[8].

In order to evaluate the phenotypic impact of mutations in the *covRS* and *ropB* genes, we determined the activity of three extracellular proteins that are known to be involved in GAS pathogenesis and to be under the direct or indirect influence of these regulators, namely SpeB, NADase, and SLS. A functional RopB is considered an absolute requirement for *speB* transcription[13], while the influence of *covS* mutations in *speB* expression is more controversial[2,5,7,8,17–19,32]. Our results support a role of both CovRS and RopB in the expression of SpeB, since null alleles in both *covS* and *ropB* were associated with the absence of SpeB production. Overall, the absence of SpeB activity was significantly more common among *emm* types 6 and 89, both frequently reported among invasive infections, especially *emm*89, which is one of the leading *emm* types among invasive isolates in several European countries[23]. However, the *emm* types that have been found to be significantly overrepresented among invasive infections in Portugal, namely *emm*1 and *emm*64[21], were not associated with an increased proportion of SpeB-negative isolates. In agreement, the absence of SpeB activity was not associated with invasive infection, indicating that, in spite of preserving several virulence factors that are regarded as important for the invasive process, the abrogation of SpeB activity cannot explain by itself the higher ability of certain clones to cause invasive disease.

Several studies report a significant influence of RopB on the expression of virulence factors other than SpeB, including NADase and SLS, either directly or due to the regulation of SpeB or the interaction with other transcriptional regulators[10,11,24]. However, it has been demonstrated that the RopB regulon is highly variable, even among strains of the same *emm* type and that the core regulon is limited to the *speB* and *spi* genes[36]. Our results are in agreement with this observation, since changes in *ropB* were not significantly associated with altered NADase or SLS activities. This suggests that, despite the influence that the regulator may have in the expression of these proteins in particular lineages, in the overall GAS population and under the growth conditions used in this study, RopB does not contribute significantly to the expression of the *nga* and *sag* operons. In contrast, *covS* null alleles were significantly associated with an increased activity of both NADase and SLS, which is consistent with a repression of the corresponding operons by phosphorylated CovR, as demonstrated for the *sag* promoter[18]. However, exceptions were noted in isolates of specific *emm* types, including *emm*64, supporting the existence of differences in the CovRS regulon of distinct GAS lineages[9]. Although both NADase and SLS have been shown to contribute to GAS cytotoxicity and virulence[37,38], their individual activities were not correlated with the type of infection caused by the isolates.

The association of missense alleles with invasiveness and with the activities of the studied virulence factors is harder to evaluate. Most of the alleles corresponding to amino acid changes in CovS and RopB occurred in multiple (often all) isolates of the same *emm* type, indicating that these are clonal mutations and probably do not represent the mechanism described for GAS in which mutations leading to a hypervirulent phenotype are acquired during infection[2,3]. Missense changes can have distinct effects on protein function depending on the nature of the amino acid replacement and the region where they occur. It is not possible to know from the current data if an altered phenotype in isolates carrying missense alleles is due to amino acid changes in these regulators or to some other characteristic of that specific lineage. Possibly reflecting this, the prevalence of missense alleles in *ropB*, *covR*, and *covS* could not be associated

with the phenotypes tested. When considering only the non-clonal missense alleles, an association with SpeB-negative isolates, but not with invasive disease, was observed in the case of *ropB*. Non-clonal missense alleles in *covRS* were not associated with any particular phenotype regarding the tested virulence factors, nor with disease presentation.

Null *covS* alleles are associated with invasive isolates and present phenotypes consistent with the ablation of this regulatory system. However, none of the phenotypes is itself associated with invasiveness, indicating that other factors under *covRS* regulation may be responsible for this higher virulence. In agreement, impairment of RopB function, known to result in SpeB downregulation, is not associated with invasiveness. The data presented supports the abrogation of CovS function as a mechanism contributing to the pathogenesis of invasive GAS infections, although this is not specifically associated with lineages identified as having enhanced invasiveness, but occurred in isolates of diverse *emm* types and STs. This mechanism is uncommon and the acquisition of such changes, despite being beneficial in the specific context of invasive disease, may incur an overall fitness cost for GAS, preventing their fixation in the population.

Materials and Methods

Bacterial strains and culture conditions. GAS isolates used in this study (104 recovered from pharyngeal exudates of patients with pharyngitis and 87 associated with invasive disease, recovered from normally sterile sites) are listed in Supplementary Table S1 online. The strains were randomly selected among a collection of 480 non-duplicate isolates recovered from human infections between 2000 and 2005 that have been previously characterised[21], so as to include one third of the isolates of each *emm* type present in the collection with $n \geq 3$, as well as all isolates of *emm* types 1 and 64. Strains SF370 (CECT 5109, obtained from Colección Española de Cultivos Tipo) and MGAS5005 (BAA-947, obtained from American Type Culture Collection) were used as controls. Strains were grown at 37 °C in Todd Hewitt broth (THB) (BD, Sparks, MD, USA) or in Tryptone Soya Agar (Oxoid, Basingstoke, UK) supplemented with 5% defibrinated sheep blood.

Molecular typing. In addition to the molecular characterisation of the isolates that had been previously performed[21], multilocus sequence typing analysis was completed for all isolates.

Gene sequencing and analysis. Genomic DNA of the isolates was extracted using cetyl trimethylammonium bromide (CTAB)[39]. All PCR and sequencing primers used in this study are listed in Supplementary Table S3. The sequences obtained for each isolate were assembled and compared with the corresponding regions of the genome of strain SF370 (GenBank AE004092), considered to be the reference wild-type alleles.

Geneious R7 (Biomatters, Auckland, New Zealand) was used to search the GenBank database (accessed on April 7th 2014) using BLAST for all previously deposited sequences of *covR*, *covS*, and *ropB* of *S. pyogenes*, and align them with the trimmed sequences of the isolates analysed in this study, using the MUSCLE algorithm with default settings.

The subset of non-duplicate alleles without indels was used to calculate the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site (*Ka*/*Ks*), and the respective Z test for stabilising selection, using MEGA5, with the Mega-Kumar method (Kimura 2-parameter). Values of $P < 0.05$ were considered statistically significant. Nucleotide diversity was calculated using DnaSP v.5.10.1.

Determination of protease activity. *Azocasein assay.* SpeB protease activity in stationary phase GAS culture supernatants was determined using an adapted azocaseinolytic assay already modified for the 96-well plate format[40,41]. Briefly, 24 h-cultures of each strain were diluted 1:10 in fresh THB and grown for 18 h, in 96-well plates. Bacteria-free supernatants were obtained by centrifugation at 3220 × g for 10 min, and transferred to a new microtitre plate. An equal volume of activation buffer (0.1 M sodium acetate [pH 5], 1 mM EDTA, 20 mM dithiothreitol) was added and the plate was incubated for 1 h at 40 °C. After activation, 2% azocasein (Sigma-Aldrich, St. Louis, MO, USA) (w/v in activation buffer) was added and incubated for 6 h at 40 °C. Samples were then precipitated with 2.5 volumes of 6% trichloroacetic acid and centrifuged at 15000 × g for 5 min. The optical density at 450 nm of the resulting supernatants was determined. The corresponding proteolytic activities were calculated using a calibration curve performed for each plate with known concentrations of proteinase K (Roche Diagnostics). Three independent assays were performed for each strain. The presence of proteolytic activity was considered positive when at least two of the three assays showed a value ≥0.0025 U.

Casein-plate assay. GAS expression of extracellular cysteine protease was determined by a plate assay as previously described[16]. Briefly, single GAS colonies were stab-inoculated into plates of medium containing 0.5-strength Columbia broth (BD), 3% w/v skim milk (BD), and 1% w/v agar (Oxoid). Protease-expressing strains produced a translucent zone surrounding the site of inoculation after 24-h incubation at 37 °C. Three independent assays were performed for each strain. The presence of proteolytic activity was considered positive when at least two of the three assays presented a translucent zone of size similar to the one of strain SF370.

SpeB Western blot analysis. For Western blot analysis, bacterial cultures were grown to late stationary phase (18 h) and centrifuged at 3200 × g for 15 min. Sterile-filtered supernatants were separated by electrophoresis on a 12% SDS-PAGE gel and transferred to a nitrocellulose membrane (Whatman, Dassel, Germany) using a Mini Trans-Blot (Bio-Rad, Hercules, CA, USA). Immunodetection was performed by chemiluminescence using monoclonal anti-SpeB antibody (Toxin Technology, Serasota, FL, USA) and goat anti-rabbit IgG conjugated with horseradish peroxidase (Santa Cruz Biotechnology, Dallas, TX, USA). Detection was performed with ECL detection reagents (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

Determination of NADase activity. The NADase activity of each strain was determined by an endpoint titre method based on previously published assays[37]. Briefly, bacteria-free supernatants of stationary-phase cultures of GAS strains were obtained as described for the azocasein assay and serially diluted (three-fold for the first dilution and two-fold for the second dilution onwards, up to 1/384) in β-NAD (Sigma-Aldrich) dissolved in PBS for a final concentration of 0.67 mM, in black 96-well plates (transparent bottom) (Greiner Bio-One, Frickenhausen, Germany). Wells containing sterile THB serially diluted in PBS and in β-NAD/PBS were used as controls, for each plate. After incubation for 1 h at 37 °C in the dark, an equal volume of NaOH was added, for a final concentration of 2 N, and the plates were incubated for 1 h at room temperature, in the dark. Fluorescence was measured with excitation at 340 nm and emission detection at 460 nm. For each strain, the NADase activity was expressed as the inverse of the highest dilution prior to a greater or equal to two-fold increase in the fluorescence value. This method has a limit of detection of 3, and strains for which a two-fold increase was not observed and with values similar to the THB + β-NAD control were considered to have a NADase activity of ≤3. A minimum of three independent assays were performed for each strain, and the majority rule was used to determine the final NADase activity.

Determination of streptolysin (SLS) activity. The streptolysin activity of each strain was determined by an endpoint titre method adapted from a previously published assay[20]. Briefly, bacteria-free supernatants of stationary-phase cultures of GAS strains were obtained as described for the azocasein assay. The supernatants (or sterile THB for control of spontaneous haemolysis – blank) were serially diluted in PBS in microplate wells (three-fold for the first dilution and two-fold for the second dilution onwards, up to 1/192), and an equal volume of a 2.5% (v/v) suspension of defibrinated sheep erythrocytes was added. Two complete haemolysis controls were included in each plate, by incubating the erythrocytes suspension with 1% Triton X-100 (positive controls). The blank for these wells consisted only of PBS and the erythrocyte suspension. After incubation at 37 °C for 1 h, erythrocytes were pelleted and the absorbance at 570 nm of the supernatants was measured. For each well, the percentage of haemolysis relative to the positive control was calculated as follows:

$$\frac{Abs_{570}\,sample - Abs_{570}\,blank}{Abs_{570}\,positive\ control - Abs_{570}\,PBS\ blank} \times 100 \quad (1)$$

The streptolysin activity was expressed as the inverse of the highest dilution prior to a ≥ two-fold decrease in the percentage of haemolysis. This method has a limit of detection of 3, and strains for which a two-fold decrease was not observed were considered to have a streptolysin activity of ≤ 3. A minimum of three independent assays were performed for each strain, and the majority rule was used to determine the final streptolysin activity. In four strains, two with functional *covRS* alleles (SF370 and SH0959A in Supplementary Table S1) and another two, of the same *emm* types, with null *covS* alleles and increased NADase activity (SH1025A and SH0421A in Supplementary Table S1), streptolysin assays were also performed in the presence of 33.3 μg/ml of trypan blue (Sigma-Aldrich) or of 16.7 μg/ml cholesterol (Sigma-Aldrich). In all four strains, the results were not significantly changed by the presence of cholesterol, while trypan blue completely inhibited haemolysis, indicating that streptolysin O does not contribute to the haemolytic activity determined under these conditions, in agreement with previous reports[20].

Statistical analysis. The allelic diversity of the *covR*, *covS*, and *ropB* genes in the studied collection of GAS isolates was evaluated using the Simpson's index of diversity (SID) and corresponding 95% confidence intervals ($CI_{95\%}$)[42]. The overall association between *covR*, *covS*, or *ropB* alleles and the *emm* type and ST was evaluated with the Adjusted Wallace coefficient with corresponding $CI_{95\%}$[43].

Unless otherwise specified, the statistical significance of pairwise associations was evaluated by calculating the respective odds ratios (when applicable) and the two-tailed Fisher's exact test, correcting the *P* values for multiple testing through the FDR linear procedure[44].

References

1. Cunningham, M. W. Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* **13,** 470–511 (2000).
2. Engleberg, N. C., Heath, A., Miller, A., Rivera, C. & DiRita, V. J. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. *J. Infect. Dis.* **183,** 1043–1054 (2001).
3. Sumby, P., Whitney, A. R., Graviss, E. A., DeLeo, F. R. & Musser, J. M. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* **2,** e5 (2006).
4. Churchward, G. The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. *Mol. Microbiol.* **64,** 34–41 (2007).
5. Graham, M. R. *et al.* Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. *Proc. Natl. Acad. Sci. USA* **99,** 13855–13860 (2002).
6. Horstmann, N. *et al.* Dual-site phosphorylation of the control of virulence regulator impacts group A streptococcal global gene expression and pathogenesis. *PLoS Pathog.* **10,** e1004088 (2014).
7. Hasegawa, T. *et al.* Detection of invasive protein profile of *Streptococcus pyogenes* M1 isolates from pharyngitis patients. APMIS 118, 167–178 (2010).
8. Aziz, R. K. *et al.* Microevolution of group A streptococci *in vivo*: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. *PloS One* **5,** e9798 (2010).
9. Sugareva, V. *et al.* Serotype- and strain- dependent contribution of the sensor kinase CovS of the CovRS two-component system to *Streptococcus pyogenes* pathogenesis. *BMC Microbiol.* **10,** 34 (2010).
10. Carroll, R. K. *et al.* Naturally occurring single amino acid replacements in a regulatory protein alter streptococcal gene expression and virulence in mice. *J. Clin. Invest.* **121,** 1956–1968 (2011).
11. Ikebe, T. *et al.* Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. *PLoS Pathog* **6,** e1000832 (2010).
12. Lin, J.-N., Chang, L.-L., Lai, C.-H., Lin, H.-H. & Chen, Y.-H. Association between polymorphisms in the *csrRS* two-component regulatory system and invasive group A streptococcal infection. *Eur. J. Clin. Microbiol. Infect. Dis.* **33,** 735–743 (2014).
13. Neely, M. N., Lyon, W. R., Runft, D. L. & Caparon, M. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. *J. Bacteriol.* **185,** 5166–5174 (2003).
14. Carroll, R. K. & Musser, J. M. From transcription to activation: how group A *Streptococcus*, the flesh-eating pathogen, regulates SpeB cysteine protease production. *Mol. Microbiol.* **81,** 588–601 (2011).
15. Kansal, R. G., McGeer, A., Low, D. E., Norrby-Teglund, A. & Kotb, M. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. *Infect. Immun.* **68,** 6362–6369 (2000).
16. Ashbaugh, C. D., Warren, H. B., Carey, V. J. & Wessels, M. R. Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. *J. Clin. Invest.* **102,** 550–560 (1998).
17. Mayfield, J. A. *et al.* Mutations in the control of virulence sensor gene from *Streptococcus pyogenes* after infection in mice lead to clonal bacterial variants with altered gene regulatory activity and virulence. *PLoS ONE* **9,** e100698 (2014).
18. Miller, A. A., Engleberg, N. C. & DiRita, V. J. Repression of virulence genes by phosphorylation-dependent oligomerization of CsrR at target promoters in *S. pyogenes*. *Mol. Microbiol.* **40,** 976–990 (2001).
19. Treviño, J. *et al.* CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A *Streptococcus* virulence factor-encoding genes. *Infect. Immun.* **77,** 3141–3149 (2009).
20. Ravins, M. *et al.* Characterization of a mouse-passaged, highly encapsulated variant of group A *Streptococcus* in *in vitro* and *in vivo* studies. *J. Infect. Dis.* **182,** 1702–1711 (2000).
21. Friães, A., Pinto, F. R., Silva-Costa, C., Ramirez, M. & Melo-Cristino, J. Group A streptococci clones associated with invasive infections and pharyngitis in Portugal present differences in *emm* types, superantigen gene content and antimicrobial resistance. *BMC Microbiol.* **12,** 280 (2012).
22. Shulman, S. T. *et al.* Group A streptococcal pharyngitis serotype surveillance in North America, 2000–2002. *Clin. Infect. Dis.* **39,** 325–332 (2004).
23. Luca-Harari, B. *et al.* Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. *J. Clin. Microbiol.* **47,** 1155–1165 (2009).
24. Hollands, A. *et al.* A naturally occurring mutation in *ropB* suppresses SpeB expression and reduces M1T1 group A streptococcal systemic virulence. *PloS One* **3,** e4102 (2008).
25. Friães, A., Pinto, F. R., Silva-Costa, C., Ramirez, M. & Melo-Cristino, J. Superantigen gene complement of *Streptococcus pyogenes*-relationship with other typing methods and short-term stability. *Eur. J. Clin. Microbiol. Infect. Dis.* **32,** 115–125 (2013).
26. Hurst, L. D. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet. TIG* **18,** 486 (2002).
27. Gusa, A. A. & Scott, J. R. The CovR response regulator of group A *Streptococcus* (GAS) acts directly to repress its own promoter. *Mol. Microbiol.* **56,** 1195–1207 (2005).
28. Chaussee, M. S., Gerlach, D., Yu, C. E. & Ferretti, J. J. Inactivation of the streptococcal erythrogenic toxin B gene (*speB*) in *Streptococcus pyogenes*. *Infect. Immun.* **61,** 3719–3723 (1993).
29. Sumby, P. *et al.* Evolutionary origin and emergence of a highly successful clone of serotype M1 group A *Streptococcus* involved multiple horizontal gene transfer events. *J. Infect. Dis.* **192,** 771–782 (2005).
30. Riddle, D. J., Bessen, D. E. & Caparon, M. G. Variation in *Streptococcus pyogenes* NAD + glycohydrolase is associated with tissue tropism. *J. Bacteriol.* **192,** 3735–3746 (2010).
31. Yoshino, M. *et al.* Nonhemolytic *Streptococcus pyogenes* isolates that lack large regions of the *sag* operon mediating streptolysin S production. *J. Clin. Microbiol.* **48,** 635–638 (2010).
32. Maamary, P. G. *et al.* Parameters governing invasive disease propensity of non-M1 serotype group A streptococci. *J. Innate Immun.* **2,** 596–606 (2010).
33. Walker, M. J. *et al.* DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. *Nat. Med.* **13,** 981–985 (2007).
34. Agarwal, S., Agarwal, S., Pancholi, P. & Pancholi, V. Role of serine/threonine phosphatase (SP-STP) in *Streptococcus pyogenes* physiology and virulence. *J. Biol. Chem.* **286,** 41368–41380 (2011).
35. Alam, F. M., Turner, C. E., Smith, K., Wiles, S. & Sriskandan, S. Inactivation of the CovR/S virulence regulator impairs infection in an improved murine model of *Streptococcus pyogenes* naso-pharyngeal infection. *PloS One* **8,** e61655 (2013).
36. Dmitriev, A. V., McDowell, E. J. & Chaussee, M. S. Inter- and intraserotypic variation in the *Streptococcus pyogenes* Rgg regulon. *FEMS Microbiol. Lett.* **284,** 43–51 (2008).
37. Bricker, A. L., Cywes, C., Ashbaugh, C. D. & Wessels, M. R. NAD^+-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group A streptococci. *Mol. Microbiol.* **44,** 257–269 (2002).
38. Molloy, E. M., Cotter, P. D., Hill, C., Mitchell, D. A. & Ross, R. P. Streptolysin S-like virulence factors: the continuing *sagA*. *Nat. Rev. Microbiol.* **9,** 670–681 (2011).

39. Ausubel, F. *et al.* in Current Protocols (ed. Chanda, V.) (Greene Publishing Associates and Wiley-Interscience: J. Wiley, 1999).
40. Collin, M. & Olsén, A. Generation of a mature streptococcal cysteine proteinase is dependent on cell wall-anchored M1 protein. *Mol. Microbiol.* **36,** 1306–1318 (2000).
41. Cole, J. N. *et al.* M Protein and hyaluronic acid capsule are essential for *in vivo* selection of *covRS* mutations characteristic of invasive serotype M1T1 group A *Streptococcus*. *mBio* **1,** e00191–10 (2010).
42. Carriço, J. A. *et al.* Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. *J. Clin. Microbiol.* **44,** 2524–2532 (2006).
43. Pinto, F. R., Melo-Cristino, J. & Ramirez, M. A confidence interval for the Wallace coefficient of concordance and its application to microbial typing methods. *PloS One* **3,** e3696 (2008).
44. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57,** 289–300 (1995).

Author Contributions

A.F. and C.P. performed the experiments. A.F., C.P and M.R. analysed and interpreted the data. A.F., C.P, J.M-C. and M.R. were involved in the conception and design of the study, as well as in drafting the manuscript and revising it critically for important intellectual content.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Accession codes: All new *covRS* and *ropB* sequences identified in this study were deposited in GenBank (accession numbers KM985476 to KM985497 and KP101294 to KP101323).

Competing financial interests: J. M-C. received research grants administered through his university and received honoraria for serving on the speakers bureaus of Pfizer, Novartis, and Gilead. M.R. received honoraria for serving on speakers bureau of Pfizer and for consulting from GlaxoSmithKline. A.F. and C.P. declare no conflict of interest. No company or financing body had any interference in the decision to publish.

How to cite this article: Friães, A. *et al.* Consequences of the variability of the CovRS and RopB regulators among Streptococcus pyogenes causing human infections. *Sci. Rep.* **5**, 12057; doi: 10.1038/srep12057 (2015).