Molecular epidemiology and virulence factors in *Streptococcus agalactiae*

Dissertação para obtenção do Grau de Doutor em Biologia

Orientadora: Doutora Maria José Borrego, Investigadora Auxiliar,
Departamento de Doenças Infecciosas,
Instituto Nacional de Saúde Dr. Ricardo Jorge

Co-orientadoras: Prof. Doutora Barbara Spellerberg, Professora Associada, Institute of Medical Microbiology and Hygiene, University of Ulm
Prof. Doutora Ilda Santos-Sanches, Professor Associada,
Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa

Jún:

Presidente: Prof. Doutora Ana Isabel Nobre Martins Aguiar Oliveira Ricardo
Arguentes: Prof. Doutora Maria Aida Costa Silva Conceição Duarte
Prof. Doutora Maria de Fátima Carvalho Serrano

Vogais: Prof. Doutora Barbara Spellerberg
Prof. Ana Madalena Peres Drummond Ludovice Mendes Gomes
Doutora Maria José Gonçalves Gaspar Borrego
Doutor João Paulo dos Santos Gomes

Setembro de 2014
Molecular epidemiology and virulence factors in *Streptococcus agalactiae*
Acknowledgements

I am very grateful to Dr. Maria José Borrego, PhD, Head of the National Reference Laboratory for Sexually Transmitted Infections, Department of Infectious Diseases at Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA), my supervisor, for accepting me in her laboratory, for her constant availability, interest, encouragement and unconditional support in the last 12 years, in particular for her supervision of experiments and her help in the elaboration of the PhD dissertation, papers and projects.

I am particularly grateful to Professor Barbara Spellerberg, PhD, Head of the Streptococci Laboratory, Institute of Medical Microbiology and Hygiene, Faculty of Medicine, University of Ulm, my supervisor, for accepting being my co-supervisor, for believing in my personal and work capacities and for giving me the opportunity to perform the PhD. I am grateful for the supervision and support, for the enthusiastic scientific talks, critical reviews, and for the continuous motivation and friendship. It has been an honor to be her PhD student.

I am very grateful to Professor Ilda Santos-Sanches, PhD, Associate Professor of the Department of Life Sciences at Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa (FCT/UNL) and Head of the Laboratory “Molecular Characterization of Pathogenic Bacteria: beta-hemolytic streptococci in carriage and disease” – Centro de Recursos Microbiológicos (CREM) – FCT/UNL, my co-supervisor, for accepting me in her laboratory, for her availability, for the scientific supervision and discussions, and critical review of the PhD dissertation, papers and projects.

I am grateful to Professor Isabel de Sá Nogueira, PhD, Coordinator of the PhD program in Biology of FCT/UNL for the availability and leadership, for monitoring my PhD annual reports and for the scientific questions during PhD seminars.

To my PhD Committee members for their continued support and encouragement: João Paulo Gomes, PhD, from INSA, my colleague and friend with whom I began this journey through the scientific research, for his inspiration, leadership, brilliant ideas, scientific explanations, technical support and critical review of the PhD dissertation, papers and projects; Professor Ana Madalena Ludovice, Assistant Professor of the Department of Life Sciences at Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa (FCT/UNL) for the availability, scientific discussions and helpful suggestions.

To the Institutions that contributed for the PhD work by providing excellent research facilities and a great scientific community: INSA, CREM-FCT/UNL and University of Ulm.

To all obstetricians and members of the Group for the Prevention of Neonatal GBS Infection involved in the PhD studies for providing clinical \textit{S. agalactiae} strains or swabs, especially to Dr. Jorge Lima from Hospital CUF Descobertas for his important contributions, clinical discussions, and persuasive power to not let me give up.

To Vitor Borges from INSA, my colleague and friend, for his important contributions regarding real-time PCR and bioinformatic analyses, for his professionalism, contagious enthusiasm, brilliant ideas, encouragement in the most stressful moments, and for scientific and personal discussions.
Acknowledgements

To Dr. Márcia Rato from whom I learned a bit more about the streptococcal world.

To my lab colleagues at INSA, in particular Silvia Viegas, Albertina Paulino, Arminda Louro, and Alexandra Nunes for all their technical support and friendship. To Inês Silvestre and Vera Damião for the important contribution to this work during her Master’s Theses.

To Dr. Ger van Zandbergen, from Paul-Erlich Institut, Germany for believing in my work since the first moments in 2008, for his collaboration in immune assays, scientific talks, and for fantastic hospitality during my stays in Ulm.

To Stefanie Mauerer, Julia Dick, Ulrike Sinnacher, Corinna Foddis, Anubha Sagar, Daniela Asam, and Moad Hay from the University of Ulm for all technical support and good moments during my stays in Ulm.

To Professor Rogério Tenreiro, Dr. Tânia Tenreiro, and Dr. Lélia Chambel from Instituto de Ciência Aplicada e Tecnologia (ICAT), Faculdade de Ciências, Universidade de Lisboa, for allowing the use of the fluorometer applied on DNase quantitative assays and for the support with the Bionumerics software.

To Professor Cristina Lobo Vilela and Dr. Maria Manuela Oliveira from the Laboratory of Microbiology at Centro de Investigação Interdisciplinar em Sanidade Animal (CIISA), Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa (FMV/UAL), for providing clinical strains of S. agalactiae isolated from bovines and for allowing the preliminary experiments regarding the biofilm production.

To Professor Filomena Martins-Pereira, Dr. Rita Castro and Catarina Farinha for their collaboration in providing clinical strains and epidemiological data.

To Fundação para a Ciência e a Tecnologia and to German Embassy in Lisbon for the financial support of this PhD Project (SFRH/BD/48231/2008; PTDC/SAU-MII/105114/2008 and DAAD scholarship/2010) and to the participation in scientific meetings where I had the opportunity to present our work and to share scientific information with other researchers.

Lastly, I would like to thank my wife Inês and my daughters Maria Clara and Catarina (born during the PhD work) for all their love, smiles, encouragement and faithful support in all my pursuits. Without each of you, I’d be nowhere near the person I am. To them I dedicate this Thesis.

The impossible seduces…
**Resumo**

*Streptococcus agalactiae* (estreptococos do grupo B), bactéria comensal dos tratos genitourinário e gastrointestinal animal e humano, é um agente patogênico oportunista que constitui uma das principais causas de pneumonia, septicemia e meningite em recém-nascidos humanos e uma doença infecciosa emergente em adultos. A administração de profilaxia antimicrobiana intraparto a mulheres grávidas colonizadas por *S. agalactiae* resultou num declínio acentuado da transmissão vertical e, por conseguinte, das internações de início precoce nos recém-nascidos, mas sem efeito na frequência de internações por *S. agalactiae* de início tardio, cuja origem das estirpes ainda não está esclarecida. A infecção por *S. agalactiae* é um processo complexo que envolve a adaptação deste agente patogênico a diferentes nichos do hospedeiro e inúmeros fatores de virulência.

Com o objetivo de estudar a estrutura populacional e a diversidade genética de grupos distintos de estirpes de *S. agalactiae* (total, N = 1141): estirpes clínicas de colonização (N = 953) e de infecção (N = 188), de origem humana (N = 1081) e bovina (N = 60), isoladas em Portugal, Alemanha e Angola entre os anos 2001 e 2012, foram utilizadas várias técnicas de tipificação, nomeadamente a serotipagem e a genotipagem capsulares, testes de susceptibilidade a antibióticos, elektroforese em gel de campo pulsado (PFGE), *multilocus sequencing typing* (MLST) e *multilocus variable number tandem repeat analysis* (MLVA). Uma análise molecular mais aprofundada incluiu a determinação de perfiles de proteínas de superfície (Alp, BibA, FbsB e Sip) de *S. agalactiae* implicadas na virulência e a deteção dos elementos genéticos móveis GBS1 e IS1548 na região intergénica *scpB-lmb*. Esta caracterização molecular e fenotípica permitiu selecionar estirpes de *S. agalactiae* pertencentes às principais linhagens genéticas para estudar a atividade das DNases, que constituiu o segundo objetivo deste trabalho. Dada a importância das nuclease de origem estreptocócica na evasão ao sistema inmunitário do hospedeiro, foram realizados ensaios qualitativos e quantitativos para avaliar a produção das DNases e respetivas implicações sobre as neutrophil extracellular traps. Foram também realizados estudos para a identificação de genes que codificam para DNases extracelulares e para a compreensão do seu papel biológico através da interação entre *S. agalactiae* e granulócitos de origem humana.

Os serótipos capsulares Ia, II, III e V foram os mais frequentes entre as estirpes de *S. agalactiae* de colonização isoladas em mulheres em idade fértil na região de Lisboa e Vale do Tejo (2005-2012), a análise de tendências em séries temporais revelou alterações na distribuição capsular, as quais poderão ter influência no desenvolvimento de vacinas baseadas no antigénio polissacarídico. De facto, a frequência de estirpes do serótipo IV, raras a nível mundial, aumentou 20 vezes entre 2006 (1%) e 2012 (20%). A nova associação do serótipo IV com o complexo clonal 17 (associado à hipervirulência) e a sua proporção entre as estirpes resistentes aos macrófagos poderá constituir uma ameaça para a saúde pública.

A pesquisa de estirpes mutantes em *S. agalactiae* com DNases de reduzida atividade permitiu identificar dois genes que codificam DNases, *sak_0220* e *sak_0814*. A construção de
Resumo

estírpes mutantes para os genes das DNases identificados e a sua análise transcriptómica durante o ciclo de crescimento de *S. agalactiae* revelaram que Sak_0814 é, até à data, a principal nuclease de *S. agalactiae*. Ensaios de infeção *in vitro* revelaram que a produção de nuclease é um fator determinante na sobrevivência de *S. agalactiae* no sangue humano, sugerindo o seu envolvimento na fuga ao sistema imunitário do hospedeiro. A determinação qualitativa e quantitativa da atividade das DNases em estírpes clínicas de *S. agalactiae* revelaram que a totalidade das estírpes de origem bovina e a maioria das estírpes de origem humana apresentaram produção de DNases extracelulares; com base em dados genómicos e de expressão génica, foi possível estabelecer uma associação estatisticamente significativa entre a não produção de DNases e a linhagem genética CC19.

Globalmente, os resultados apresentados nesta Tese contribuem para um maior conhecimento sobre a diversidade fenotípica e genética das estírpes de *S. agalactiae*, clarificam (parcialmente) a base genética da produção de DNases em *S. agalactiae*, realçando a sua função biológica na patogénese e permitindo uma correlação entre a atividade das DNases e duas variáveis epidemiológicas: linhagem MLST e origem clínica. Os mecanismos responsáveis pela resistência bacteriana à imunidade inata poderão constituir uma nova linha de investigação, fundamental para a intervenção terapêutica baseada na desativação de fatores de virulência específicos, como as DNases, tornando o agente patogénico vulnerável à ação do sistema imunitário do hospedeiro.

**Palavras-chave:** *Streptococcus agalactiae*, epidemiologia molecular, linhagens genéticas, resistência aos antibióticos, fatores de virulência, DNases extracelulares, imunidade inata
Abstract

*Streptococcus agalactiae* (Group B *Streptococcus*), a commensal bacterium of the animal and human genitourinary and gastrointestinal tracts, is an opportunistic pathogen and one of the leading causes of pneumonia, sepsis and meningitis in human newborns, and an emerging infectious disease among adults. The implementation of intrapartum antimicrobial prophylaxis to pregnant women colonized with *S. agalactiae* led to a sharp decline of the vertical transmission, and consequently in the early-onset *S. agalactiae* disease, but did not affect the frequency of the late-onset *S. agalactiae* disease for which the source of infection remains unclear. *S. agalactiae* infection is a complex process that involves numerous virulence factors, resulting in the adaptation of this pathogen to different host environments.

To study the population structure and the genetic diversity of distinct groups of *S. agalactiae* strains (total, N = 1141): colonizing (N = 953) and invasive (N = 188) clinical strains of human (N = 1081) and bovine (N = 60) origin isolated in Portugal, Germany and Angola between 2001 and 2012, several typing techniques were used, namely capsular serotyping and genotyping, antibiotic susceptibility testing, pulsed-field gel electrophoresis (PFGE), multilocus sequence typing (MLST) and multilocus variable number tandem repeat analysis (MLVA). Further molecular analysis included the surface protein gene profiling of known *S. agalactiae* surface proteins (Alp, BibA, FbsB, Sip) with a recognized role in virulence, and the detection of the mobile genetic elements GBS1 and IS1548 within the *scpB-lmb* intergenic region. This molecular and phenotypic characterization allowed the selection of *S. agalactiae* strains belonging to the major genetic lineages to study the DNase activity, the second major goal of this work. Considering the importance of streptococcal nucleases on the evasion from the host immune system, in particular from neutrophil extracellular traps, qualitative and quantitative DNase assays were performed. In addition, the identification of genes encoding extracellular DNases was done to understand their biological role in the interaction between *S. agalactiae* and human granulocytes.

Capsular serotypes Ia, II, III, and V were the most common among the colonizing *S. agalactiae* strains recovered from women of child bearing age in the Lisbon and Tagus Valley region (2005-2012); temporal trend analysis evidenced changes in the capsular distribution, which may impact in the development of a polysaccharide vaccine. In fact, the frequency of serotype IV strains, rare worldwide, increased 20-fold between 2006 (1%) and 2012 (20%). The novel association of serotype IV with the clonal complex (CC) 17 (associated to hypervirulence) and its considerable proportion among macrolide resistant strains may represent a threat to public health in a near future.

The screening of *S. agalactiae* mutants displaying a diminished DNases activity allowed the identification of two DNase encoding genes, *sak_0220* and *sak_0814*. The construction of mutant strains for both DNases genes and their transcriptomic analysis during the growth cycle of *S. agalactiae* evidenced Sak_0814 as, so far, the major nuclease of *S. agalactiae*. *In vitro* infection experiments revealed that the nuclease production represents a major determinant for the survival
of *S. agalactiae* in human blood, suggesting their involvement in the escape mechanisms from the host immune system. Qualitative and quantitative DNase assays among clinical strains of *S. agalactiae* revealed that 100% of the bovine and the majority of human *S. agalactiae* strains produced extracellular DNases; a statistically significant association could be established between DNase non-production and the CC19 genetic lineage, which was supported by genomic and transcriptomic data.

Globally, this Thesis contributes to a better knowledge on the phenotypic and genetic diversity of the circulating *S. agalactiae* strains, clarifies (in part) the genetic basis of nuclease production in *S. agalactiae*, highlighting their biological role in pathogenesis and correlating the DNase activity with two epidemiological variables: MLST lineage and clinical origin. The understanding of the mechanisms underlying the bacterial innate immune resistance open research frontiers regarding therapeutic interventions geared for disabling specific virulence factors such as DNases, that should make the pathogen prone to host innate immune clearance.

**Keywords:** *Streptococcus agalactiae*, molecular epidemiology, genetic lineages, antibiotic resistance, virulence factors, extracellular DNases, innate immunity
Table of Contents

Acknowledgements iii
Resumo v
Abstract vii
Table of Contents ix
Figure Index xii
Table Index xiii
Abbreviations xiv
Thesis Outline xvi

1. Chapter I: General Introduction 1
   1.1 The genus Streptococcus 2
   1.2 Streptococcus agalactiae 4
      1.2.1 The organism: historical and morphologic aspects 4
      1.2.2 Clinical relevance and treatment 6
      1.2.3 Molecular epidemiology 10
      1.2.4 Genomics 15
      1.2.5 Pathogenesis 18
      1.2.6 Extracellular DNases 21
   1.3 Objectives 23

2. Chapter II: Accuracy of prenatal culture in predicting intrapartum 24
   group B Streptococcus colonization status
      2.1 Introduction 26
      2.2 Methods 26
      2.3 Results 28
      2.4 Discussion 28

3. Chapter III: Molecular characterization and antimicrobial susceptibility 31
   profiles in S. agalactiae colonizing strains; association of erythromycin resistance with subtype III-1 genetic clone family
Table of Contents

3.1 Introduction 33
3.2 Materials and Methods 33
3.3 Results 36
3.4 Discussion 38
3.5 Acknowledgements and transparency declaration 40

4.1 Introduction 43
4.2 Materials and Methods 44
4.3 Results 45
4.4 Discussion 48
4.5 Members of the Group for the prevention of Neonatal GBS infection 50
4.6 Acknowledgements and transparency declaration 50

5. Chapter V: Molecular epidemiology of group B streptococcal meningitis in children beyond the neonatal period from Angola 51
5.1 Introduction 53
5.2 Methods 54
5.3 Results and Discussion 55
5.4 Acknowledgements and transparency declaration 59

6. Chapter VI: Selection of reference genes for real time expression studies in S. agalactiae 60
6.1 Introduction 62
6.2 Methods 63
6.3 Results 67
6.4 Discussion 72
6.5 Acknowledgements 73
7. Chapter VII: Nuclease production represents a major determinant for survival of *S. agalactiae* in human blood

7.1 Introduction ................................................. 76
7.2 Materials and Methods ................................. 77
7.3 Results .................................................. 83
7.4 Discussion ............................................... 91
7.5 Acknowledgements and transparency declaration 93

8. Chapter VIII: Evaluation of the DNase activity in clinical strains of *S. agalactiae* of human and bovine origin

8.1 Introduction ................................................. 96
8.2 Material and Methods ................................. 97
8.3 Results .................................................. 102
8.4 Discussion ............................................... 107

9. Chapter IX: Concluding Remarks

9.1 Insights into the carrier state ............................ 111
9.2 Molecular epidemiology and antibiotic resistance 112
9.3 Study of extracellular DNases ........................ 114
9.4 Future Perspectives ....................................... 116

References ................................................................ 118

Annex ..................................................................... 143
<table>
<thead>
<tr>
<th>Figure Index</th>
</tr>
</thead>
</table>

1.1 Phylogenetic tree of 25 streptococcal species based on 16S rRNA gene sequences  
1.2 Scheme of the cell wall structure of S. agalactiae  
1.3 MLST/eBurst diagram showing genetic relationships between human S. agalactiae clinical strains  
1.4 Neutrophil extracellular traps associated to bacteria; electron microscopy  
3.1 eBurst diagram of genetic lineages among erythromycin-resistant S. agalactiae strains belonging to subtype III-1  
4.1 Serotype distribution among S. agalactiae colonizing strains (n=953) from women of reproductive age, Lisbon and Tagus valley (2005-2012)  
4.2 Serotype distribution among S. agalactiae strains resistant to erythromycin and clindamycin, Lisbon and Tagus valley (2005-2012)  
5.1 Genetic characteristics of 21 invasive S. agalactiae strains from Angola  
6.1 Growth curves of three S. agalactiae prototype strains at 37°C and 40°C  
6.2 Expression stability of 10 candidate reference genes of O90R, NEM316 and 2603V/R during 5 time-points of growth  
6.3 Expression stability of 10 candidate reference genes of O90R, NEM3  
2603V/R during the logarithmic growth phase  
7.1 Quantiative DNase assay by using gel electrophoresis  
7.2 DNase activity of recombinant DNase proteins (Sak_0220 and Sak_0814)  
7.3 S. agalactiae DNase mutants on DNA methyl green agar plates  
7.4 S. agalactiae double DNase mutant on DNA methyl green agar plates  
7.5 Phylogenetic analysis of sak_0220 and sak_0814 of 10 S. agalactiae strains  
7.6 Western blot of recombinant Sak_0220 of S. agalactiae  
7.7 Expression analysis of sak_0220 and sak_0814 by FACS  
7.8 Growth curves of S. agalactiae strains (wild type and DNase mutants)  
7.9 Infection experiments between S. agalactiae and human granulocytes  
8.1 Algorithm for the evaluation of DNase activity among S. agalactiae strains  
8.2 Semiquantitative DNase assay in gel electrophoresis (NEM316 vs CC19 strain)  
8.3 DNase activity displayed by S. agalactiae strains over time (1h, 2h, 4h, 17h)  
8.4 Expression of sak_0814 of NEM316 and 2603V/R strains during Log phase  
8.5 Expression of sak_0220 of NEM316 and 2603V/R strains during Log phase
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Primers for PCR and sequencing used in the study presented in chapter III</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Antibiotic resistance and molecular typing of colonizing <em>S. agalactiae</em> strains</td>
<td>37</td>
</tr>
<tr>
<td>6.1 Top five stable genes at 37°C for kPCR normalization by using geNORM and NormFinder bioinformatic tools</td>
<td>71</td>
</tr>
<tr>
<td>6.2 Top five stable genes at 40°C for kPCR normalization by using geNORM and NormFinder bioinformatic tools</td>
<td>71</td>
</tr>
<tr>
<td>7.1 Bacterial strains and plasmids used in the study of chapter VII</td>
<td>78</td>
</tr>
<tr>
<td>7.2 Oligonucleotide primers used in the study of chapter VII</td>
<td>79</td>
</tr>
<tr>
<td>8.1 Collection of <em>S. agalactiae</em> strains used in the study of chapter VIII</td>
<td>98</td>
</tr>
<tr>
<td>8.2 DNase activity among <em>S. agalactiae</em> strains of human origin, from colonization or infection</td>
<td>102</td>
</tr>
<tr>
<td>A.1 Primers for qRT-PCR analysis of 10 candidate reference genes (chapter VI)</td>
<td>144</td>
</tr>
</tbody>
</table>
Abbreviations

Alp  Alpha-like protein
ATCC  American Type Culture Collection
BibA  Group B Streptococcus immunogenic bacterial adhesin
BLAST  Basic local alignment search tool
bp  Base pair
°C  Celsius degrees
CAMP  Christie Atkins Munch-Petersen
CC  Clonal complex
CDC  Centers for Disease Control and Prevention
cDNA  Complementary deoxyribonucleic acid
CFU  Colony forming units
CLSI  Clinical and laboratory standards institute
cMLS_b  Constitutive resistance to macrolides, lincosamides and streptogramins B
CPS  Capsular polysaccharide
CREM  Centro de Recursos Microbiológicos
Ct  Threshold cycle
Da  Dalton
DNA  Deoxyribonucleic acid
DNase  Deoxyribonuclease
dNTPs  Deoxynucleotide triphosphate
DLV  Double-locus variant
EDTA  Ethylenediamine tetra acetic acid
EGFP  Enhanced green fluorescent protein
EOD  Early-onset disease
erm  Erythromycin resistance methylase genes
fbsB  Fibrinogen-binding protein B gene
FACS  Fluorescence-activated cell sorting
FCT/MEC  Fundação para a Ciência e a Tecnologia, Ministério da Educação e Ciência
FCT/UNL  Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
GAS  Group A Streptococcus (S. pyogenes)
GBS  Group B Streptococcus (S. agalactiae)
h  Hours
IAP  Intrapartum antibiotic prophylaxis
ICE  Integrative conjugative elements
ICAT  Instituto de Ciência Aplicada e Tecnologia, Universidade de Lisboa
IL  Interleukin
iMLS_b  Inducible resistance to macrolides, lincosamides and streptogramins B
INSA  Instituto Nacional de Saúde Dr. Ricardo Jorge
IPTG  Isopropyl β-D-1-thiogalactopyranoside
IR  Inverted repeat
IS  Insertion sequence
LGT  Lateral gene transfer
lmb  Laminin-binding protein gene
LOD  Late-onset disease
LPS  Lipopolysaccharide
M  mefA-encoded efflux pump phenotype
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb</td>
<td>Mega base-pair</td>
</tr>
<tr>
<td>mef</td>
<td>Efflux-mediated macrolide resistant genes</td>
</tr>
<tr>
<td>MEGA</td>
<td>Molecular evolutionary genetics analysis, software</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MGE</td>
<td>Mobile genetic element</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimum inhibitory concentration</td>
</tr>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>MLST</td>
<td>Multilocus sequence typing</td>
</tr>
<tr>
<td>MLVA</td>
<td>Multilocus variable-number tandem repeat analysis</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>N</td>
<td>Number</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NET</td>
<td>Neutrophil extracellular trap</td>
</tr>
<tr>
<td>NT</td>
<td>Nontypeable</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PAI</td>
<td>Pathogenicity island</td>
</tr>
<tr>
<td>PBP</td>
<td>Penicillin-binding protein</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PFGE</td>
<td>Pulsed-field gel electrophoresis</td>
</tr>
<tr>
<td>PI</td>
<td>Pilus island</td>
</tr>
<tr>
<td>PMA</td>
<td>Phorbol 13-myristate 12-acetate</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative real-time polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNAse</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>s</td>
<td>Second</td>
</tr>
<tr>
<td>scpB</td>
<td>C5 peptidase gene</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate - polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>sip</td>
<td>Surface immunogenic protein gene</td>
</tr>
<tr>
<td>SLV</td>
<td>Single locus variant</td>
</tr>
<tr>
<td>ST</td>
<td>Sequence type</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate EDTA buffer</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA buffer</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>TSA</td>
<td>Tryptone soya agar</td>
</tr>
<tr>
<td>UPGMA</td>
<td>Unweighted pair group method with arithmetic mean</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
This doctoral dissertation is based on seven manuscripts (mentioned below). Five manuscripts have been published in international peer reviewed journals and the other two are in final preparation for submission.

The chapter presentation order of this PhD dissertation does not necessarily reflect a chronological order, as some of the tasks/studies were done simultaneously, where the results obtained during one task/study influenced the progress of the other and vice-versa. Also, the time between the submission of an article and its publication date largely depends on the journal, and on the necessary revisions. According to this, we have presented the chapters in agreement with the objectives that were defined for this doctoral project.

**Chapter I** – consists of a general introduction, presenting the state of the art of some biological aspects of *S. agalactiae*, contextualizing the problems addressed in the Thesis, by briefly reviewing essential aspects of *S. agalactiae* colonization and infection, as well as the preventive strategies. The contribution of virulence factors and antimicrobial resistance for *S. agalactiae* epidemiology, the importance of surveillance and characterization of clinical strains, the major findings during recent years on the structure of the *S. agalactiae* population, and the putative role of extracellular DNases in pathogenesis are also reviewed in this chapter. Here were introduced and described the main objectives of this PhD work.


Chapter IX - presents an integrated discussion of the major findings of the Thesis, highlighting unsolved questions that could be addressed in the future follow-up of these investigations.
Chapter I

General Introduction
1. General Introduction

1.1 The genus *Streptococcus*

The genus *Streptococcus* belongs to the class Bacilli and the order Lactobacillales. The group is large and comprises numerous clinically significant species which are responsible for a wide variety of infections in human and animals (Nobbs *et al.*, 2009). Structurally, these microorganisms are spherical Gram-positive, catalase negative cocci, non-motile, non-spore forming, which often form chains of two or more cells with 0.5 to 2 μm in diameter and may exhibit capsule. Most streptococci are facultative anaerobes, capable of propagation in the presence of oxygen, being however devoided of respiratory metabolism. Streptococci are nutritionally demanding, and optimal *in vitro* growth rates are observed at temperatures between 35°C and 37°C in media enriched with blood, which is often used to distinguish phenotypic characteristics among streptococcal species (Facklam, 2002; Schuchat, 1999).

The species belonging to the genus *Streptococcus* are found among several hosts and associated with pathological conditions, whereby different protocols have been used for the identification of these bacteria. Still, precise identification of streptococci is laborious. Colony morphology, type of hemolysis on blood agar plates, biochemical reactions, and serologic specificity based on antigenic differences in cell surface composition, constitute phenotypic features used for streptococci classification (Facklam, 2002; Wyder *et al.*, 2011). Streptococci are divided into three groups by the type of hemolysis reaction on sheep blood agar plates: a) beta-hemolytic, which consists in clear, complete lysis of red blood cells; b) alpha-hemolytic, which consists in incomplete lysis of red blood cells forming a green halo surrounding the bacterial colonies; and c) non-hemolytic, when the lysis of red blood cells is absent (gamma-hemolysis) (Facklam, 2002; Spellerberg *et al.*, 1999a). However, hemolytic activity differs within species and depends on incubation conditions. The identification of non-hemolytic strains among *Streptococcus pyogenes*, *Streptococcus agalactiae* and *Streptococcus dysgalactiae* subsp. *equisimilis*, which are considered as typically beta-hemolytic bacteria, constitute a limitation of this methodology.

The pioneer serological classification proposed by Rebecca Lancefield in 1933 (Lancefield, 1933) based on the antigenicity of carbohydrate composition of the cell wall, enhanced the discrimination within streptococcal species. Lancefield classification of streptococci is still used nowadays, and allows the identification of 20 serogroups, named Lancefield groups A to V (excluding I and J). However, subsequent studies demonstrated that merely beta-hemolytic streptococci may be characterized via Lancefield serotyping (Farrow & Collins, 1984; Lawrence *et al.*, 1985). Indeed, *Streptococcus pneumoniae* and *Streptococcus uberis*, which exhibit alpha- and gamma-hemolysis, respectively, do not display specific group antigens and thus cannot be classified by the Lancefield system (Facklam, 2002). Also, *Streptococcus dysgalactiae* subsp. *equisimilis* belong to Lancefield serogroups A, C or G (Vandamme *et al.*, 1996).
Based on the increased availability of genetic information, the *Streptococcus* genus has undergone considerable taxonomic revisions and has been divided into two major groups based on 16S rRNA gene sequence similarity, designated Pyogenic and Viridans, where the latter comprises the groups Anginosus, Mitis, Mutans, Salivarius and Bovis (Kawamura et al., 1995) (Figure 1.1).

![Phylogenetic tree](image)

**Figure 1.1** Phylogenetic tree comprehending 25 streptococcal species based on 16S rRNA gene. (Adapted from Nobbs et al., 2009).

The analysis of 16S rRNA gene sequences represents one of the most powerful tools for the classification of microorganisms (Stackebrandt et al., 1987; Woese, 1987) and has been used for the identification and classification of clinically relevant microbes (Claridge, 2004; Janda & Abbott, 2007). However, various alternatives to this single gene analysis have been employed for the identification of streptococcal isolates such as the multilocus sequence analysis (MLSA) schemes developed for the Mitis and Anginosus groups (Bishop et al., 2009; Jensen & Kilian, 2012). MLSA could also identify novel species within the genus *Streptococcus*, in particular within complex groups, such as Anginosus whose members reveal a large antigenic heterogeneity, as they react with Lancefield groups A, C, F and G antisera and can be β-hemolytic, α-hemolytic or non-hemolytic (Jensen & Kilian, 2012; Jensen et al., 2013).

Currently, hundreds of complete and draft streptococcal genomes have been deposited in the GenBank database and it should improve streptococci taxonomy and contribute to a better understanding of the evolutionary diversification of streptococci. In fact, recent comparative genomic analyses of the genus *Streptococcus* (138 genomes analyzed) indicated that all strains
branched into two distinct populations, with Pyogenic, Bovis, Mutans and Salivarius species groups forming one population, and Mitis, Anginosus and Unknown groups clustering into another population, suggesting that there are two major evolutionary lineages within this genus (Gao et al., 2014).

1.2 *Streptococcus agalactiae*

1.2.1 The organism: historical and morphologic aspects

*Streptococcus agalactiae* (*S. agalactiae*) are a Gram-positive diplococcus belonging to the Lancefield group B. Before Lancefield’s classification of hemolytic streptococci in 1933 (Lancefield, 1933), this microorganism was known to microbiologists by its characteristic colonial morphology (gray-white flat-nucoid with 3 to 4 mm in diameter), surrounded by a narrow zone of beta-hemolysis in sheep blood agar plates. Occasionally, *S. agalactiae* strains (approximately 1-5%) may also be nonhemolytic (Facklam, 2002; Spellerberg et al., 1999a).

Some tests allow the presumptive identification of *S. agalactiae*, namely bacitracin and sulfamethoxazole-trimethoprim disk susceptibility testing (92% to 98% of strains are resistant), pigment production during anaerobic growth on Granada media (96% to 98% of strains produce an orange-red pigment), and CAMP testing (Christie-Atkins-Munch-Petersen: 98% to 100% of strains are CAMP-positive) (Facklam et al., 1979; Christie et al., 1944; Rosa-Fraile et al., 1999; Tapsall et al., 1986). The CAMP test detects a diffusible, heat-stable, extracellular protein produced by *S. agalactiae* that enhances the hemolysis of sheep erythrocytes by *Staphylococcus aureus*. The CAMP factor acts synergistically with the beta hemolysin produced by *S. aureus* (Podbielski et al., 1994).

Besides the Group B specific antigen, the cell wall of *S. agalactiae* is composed by a polysaccharide located in the outer cell wall that constitutes the capsule, defining the capsular types of this bacterium and linked together with the capsular polysaccharide (CPS) to the peptidoglycan layer, involving the N-acetylmuramic acid and N-acetylglucosamine residues, (Baron & Kasper, 2005; Caliot et al., 2012; Deng et al., 2000; Lancefield, 1933, 1934; Lancefield & Hare, 1935) (Fig. 1,2).
Figure 1.2. Scheme of the cell wall structure of *S. agalactiae*. Electron micrograph image of a type III *S. agalactiae* organism; the inset shows the *S. agalactiae* inner cell membrane, the peptidoglycan layer that anchors the negatively-charged capsular polysaccharide (CPS) and the group B antigen polysaccharide. Lipoproteins and glycolipids extend from the cell membrane, including an anchor for lipoteichoic acid (LTA). Most surface proteins attach to the peptidoglycan. CPS, LTA, and cell-surface proteins contribute to the organism's ability to adhere, to invade host cells and to evade from host immune defenses in the course of infection contributing for *S. agalactiae* pathogenesis. (Adapted from Baron & Kasper, 2005).

Until 1930s, and before it was considered pathogenic for humans, *S. agalactiae* was known to cause infections in animals, in particular bovine mastitis (Stableforth, 1950). In humans, *S. agalactiae* was firstly observed on vaginal secretions from asymptomatic women and was considered as a colonizing agent (Lancefield & Hare, 1935). However, in 1938, Fry (Fry, 1938) reported three patients with fatal puerperal sepsis caused by *S. agalactiae*, constituting the first report of this etiological agent in human hosts. Sporadic cases were reported during the next three decades, but *S. agalactiae* remained unfamiliar to most clinicians until the 1970s, when a dramatic increase in the incidence of septicemia and meningitis in neonates caused by group B streptococci was documented from geographically diverse regions (Baker et al., 1973; Barton et al., 1973; Franciosi et al., 1973; Howard & McCracken, 1974). Emergence of group B streptococcal infections in neonates was accompanied by an increasing number of infections among pregnant and nonpregnant adults, whose incidence remained stable through the early 1990s. In response, since 1996, the Centers for Disease Control and Prevention (CDC) have been publishing guidelines for the prevention of *S. agalactiae* infections that significantly contributed for the reduction of neonatal mortality from greater than 50% to 4-6% (CDC, 1996; Dermer et al., 2004; van Dyke et al., 2009; Verani et al., 2010).
1.2.2 Clinical relevance and treatment

*S. agalactiae* is a commensal bacterium that asymptptomatically colonizes the gastrointestinal and genitourinary tracts of healthy humans and animal hosts (Nobbs *et al.*, 2009). In European countries, the frequency of *S. agalactiae* colonization among pregnant women ranges from 6% to 36% (Barcaite *et al.*, 2008, Florindo *et al.*, 2010 – Chapter III), with most studies reporting rates higher than 20%, with important variations from country to country and even from hospital to hospital in the same country (Barcaite *et al.*, 2008; Stupak *et al.*, 2010). The variations in the frequency of *S. agalactiae* colonization could be related to population sample sizes, distinct diagnostic methods, demographic differences, among others. Nevertheless, pregnancy status has no influence over anogenital colonization, as similar rates were observed among nonpregnant and pregnant populations (Brumil *et al.*, 2006; Florindo *et al.*, 2010 – Chapter III).

In neonates, *S. agalactiae* may be responsible for fatal invasive infections, constituting the leading cause of neonatal pneumonia, septicemia, and meningitis in industrialized countries (Gascghinard *et al.*, 2011; Le Doare & Heath, 2013, Levent *et al.*, 2010; Libster *et al.*, 2012).

In early reports from 1980s, mortality due to *S. agalactiae* meningitis ranged from 20% to 30%, and survivors were at risk for long-term sequelae (Edwards *et al.*, 1985; Wald *et al.*, 1986). Despite prompt antimicrobial treatment and improvement in neonatal care, contemporary surveys showed that up to 14% of neonates with *S. agalactiae* meningitis die acutely and approximately 40% of meningitis survivors develop long-term/permanent severe neurologic or functional sequelae such as cerebral palsy, hearing loss, blindness, cognitive delay, speech/language disorders or motor deficits (Gascghinard *et al.*, 2011; Georget-Bouquinet *et al.*, 2008; Libster *et al.*, 2012, Thigpen *et al.*, 2011).

*S. agalactiae* is also a significant cause of morbidity and mortality in nonpregnant adults, particularly for those with underlying immunocompromised diseases, such as malignancies, diabetes mellitus, cirrhosis, and HIV infection and for the elderly (Phares *et al.*, 2008). Although *S. agalactiae* infections among adults are most frequently community acquired, nosocomial acquisition is also of concern (Farley, 2001). The spectrum of *S. agalactiae* disease in adults is broad, most frequently including bacteremia, pneumonia, osteoarticular, skin or soft tissue and urinary tract infections (Farley, 2001; Skoff *et al.*, 2009). Less frequent clinical presentations include meningitis and endocarditis, which are, however, associated with significantly higher morbidity and mortality in comparison to neonates (Domingo *et al.*, 1997; Edwards & Baker, 2005; Skoff *et al.*, 2009).

In neonates, two distinct *S. agalactiae*-associated clinical syndromes, referred to as early-onset disease (EOD; age 0–6 days) and late-onset disease (LOD; age 7–89 days) have been recognized (Edwards & Baker, 2005; Le Doare & Heath, 2013). LOD is characterized by
bloodstream infection of the infant with a high incidence of meningeal involvement. Community or nosocomial acquisition of *S. agalactiae* may be responsible for LOD, although this infectious process remains poorly understood, as vertical transmission and prematurity may also be implicated (Gagneur et al., 2009; Lin et al., 2003; Mullany, 2001; Schrag & Verani, 2013).

EOD represents the majority of cases. It usually manifests within the first 24 hours of life, and it is characterized by a fulminating course that often leads to death. EOD is associated with maternal colonization and vertical transmission, through aspiration of infected amniotic fluid or direct contact with the colonized birth canal during delivery, regularly manifesting as pneumonia and bacteraemia (Gibbs et al., 2004; Liu & Nizet, 2004; Schrag & Verani, 2013; Trager et al., 1996). Vertical transmission occurs at delivery in about 20% to 50% of neonates whose mothers are colonized with *S. agalactiae*, but only 1-2% of these newborns will develop EOD (Baker & Barrett, 1973; Berardi et al., 2013; Blumberg et al., 1996; Edwards & Nizet, 2011; Kunze et al., 2011). Thus, maternal carriage is considered the most important risk factor for the establishment of invasive *S. agalactiae* disease in newborns. Other maternal risk factors and obstetrics complications that favor the development of EOD include chorioamnionitis, preterm delivery, prolonged rupture of fetal membranes (>18 h), intrapartum fever (≥38°C), *S. agalactiae* bacteriuria during the current pregnancy and a previous delivery of a *S. agalactiae*-infected infant (Schrag et al., 2002, Verani et al., 2010).

Prevention of EOD through the use of intrapartum antibiotics given to pregnant women with obstetric risk factors or known carriage of *S. agalactiae*, was first proposed by the Centers for Disease Control and Prevention in 1996 (CDC, 1996). The subsequent updates of CDC guidelines (Schrag et al., 2002; Verani et al., 2010) excluded the risk-based approach, as the majority of EOD cases occur in newborns whose mothers did not present any obstetric risk factors (Neto, 2008; Schrag et al., 2002). Thus, since 2002, the prevention of *S. agalactiae* infections has been focused on the rectovaginal screening of pregnant women at 35–37 weeks of gestation and on the antibiotic prophylaxis of all colonized women (Schrag et al., 2002; Verani et al., 2010).

Considering the improvements in therapy and clinical management, the mortality associated with EOD decreased from higher than 50% in the 1970s, to 4-6% in recent years (Edwards & Nizet, 2011; Schrag et al., 2000; Schrag & Verani, 2013; Verani et al, 2010). Accordingly, in developed countries, a significant decrease in EOD occurred after the establishment of *S. agalactiae* prevention policies; e.g., in the United States, the incidence declined from 1.8 cases per 1000 live births in 1990 to 0.32 cases/1000 live births in 2003 (Puopolo et al., 2005; Zangwill et al., 1992); in Australia, the incidence fell from 1.43 cases per 1000 live births in 1993 to 0.25/1000 live births in 2001 (Daley et al., 2004); and in Spain, the incidence declined from 2.4 cases/1000 live births in 1996 to 0.33 cases/1000 live births in 2008 (Lopez-Sastre et al.,
2009). In Portugal, a national epidemiological survey performed between 2001 and 2005 (Neto, 2008) revealed an incidence of EOD of 0.44/1000 live births. This study further evidenced similar mortality rates for EOD and LOD (6.7% vs. 6.3%, respectively), which varied with gestation: 4.6% for term infants, 15.2% for preterm and 18% for very low birth weight infants (<1500 g).

Considering a worldwide meta-analysis comprehending 56 studies in infants younger than 90 days, the mean incidence of EOD (0.43 cases/1000 live births) and associated case fatality (12.1%) were two-fold higher than the average values for LOD. African countries evidenced the most concerning data, where the mean incidence of EOD was 1.21 cases per 1000 live births and the mortality was three-fold higher in comparison to developed countries (Edmond et al., 2012). This might reflect difficulties to implement the CDC guidelines in many low- and middle-income countries (Edmond et al., 2012), namely the rationale administration of intrapartum antibiotic prophylaxis (IAP).

Penicillin is the first-line group of antibiotics for intrapartum antimicrobial prophylaxis (and treatment) of S. agalactiae infections, displaying high pharmacokinetic and bioavailability by easily crossing the placental blood barrier; moreover penicillin is cheap and no cases of resistant S. agalactiae strains have been described yet. However, during the last decade there were descriptions in the United States, Japan and Sweden of S. agalactiae strains displaying reduced susceptibility to penicillin and other beta-lactams which may have clinical implications (minimum inhibitory concentrations, MICs, for penicillins slightly higher than the susceptibility breakpoint criteria defined by the Clinical and Laboratory Standards Institute, that is, MICs ≤ 0.12 mg/L (Chu et al., 2007; Dahesh et al., 2008; Nagano et al., 2008; Kasahara et al., 2010; Persson et al., 2008). This phenomenon has been attributed to the occurrence of mutations in the genes encoding penicillin binding proteins (PBPs), the target of beta-lactam antibiotics. Nevertheless, no outbreak of an epidemic clone with increased penicillin MIC has occurred, as phylogenetic comparative analysis between PBP encoding genes from those strains and the PBP encoding genes from penicillin-susceptible S. agalactiae strains revealed that genetic lineages of penicillin non-susceptible strains have been independently emerging through the accumulation of mutations inducing amino acid substitutions, namely Q557E in PBP 2X (Nagano et al., 2008). This substitution corresponds to the Q552E substitution in PBP 2X of S. pneumoniae that has been reported to be the major responsible for the reduction of susceptibility to beta-lactams (Pernot et al., 2004).

In cases of reduced penicillin susceptibility or for people allergic to penicillin, second-line antibiotics must be chosen as an alternative (Schrag et al., 2002; Verani et al., 2010). For penicillin-allergic pregnant women without a history of anaphylaxis, angioedema, respiratory distress or urticaria, cefazolin, a cephalosporin, is the preferred drug, whereas erythromycin, clindamycin and vancomycin are recommended for penicillin-allergic women at high risk for
anaphylaxis (Schrag et al., 2002; Verani et al., 2010). However, survey studies involving invasive and noninvasive *S. agalactiae* strains evidenced a high rate of resistance to tetracycline, macrolides, lincosamides, as well as the emergence of resistance to fluoroquinolones (Florindo et al., 2010, 2014b - Chapters III and IV; Gherardi et al., 2007; Kawamura et al., 2003). Furthermore, a high frequency of fluoroquinolone- and macrolide-resistance (100% and 47.4%, respectively) among *S. agalactiae* clinical strains with reduced penicillin susceptibility has been recently reported (Kimura et al., 2013), representing a serious public health threat.

Resistance to macrolides and to lincosamides, such as erythromycin and clindamycin, respectively, is commonly mediated by two major mechanisms, namely the modification of the antibiotic binding site at 23S rRNA by methylation (most common) and the active efflux of the antibiotic (Leclercq, 2002). The methylation at 23S rRNA region is mediated by a ribosomal methylase encoded by *erm*-class genes (“erythromycin ribosome methylation”), which adds two methyl groups to the 23S rRNA adenine residue, causing configurational changes in the ribosome, and consequently, a decrease binding of macrolides, lincosamides and streptogramin B; this resistance phenotype is designated MLS\(_B\). The production of the methylase may be constitutive, when it is naturally produced by the cell itself (cMLS\(_{B}\) phenotype), or inductive if the presence of an antibiotic is required to induce its expression (iMLS\(_{B}\) phenotype). The second mechanism is based on the active efflux of macrolides resulting from an energy-dependent efflux pump encoded by *mef*-class genes (“macrolide efflux”), which confers resistance to macrolides only (M phenotype) (Leclercq, 2002).

The resistance to erythromycin and clindamycin has been estimated between 10 and 30% in Europe and in the United States (Barcaite et al., 2008; de Azavedo et al. 2001; Farley, 2001; Fitoussi et al. 2001; Florindo et al. 2010, 2014b - Chapters III and IV; Gherardi et al., 2007; Gygax et al., 2006; Phares et al., 2008); nonetheless, in a study held in Taiwan (Hsueh et al., 2001), the resistance rate to erythromycin and clindamycin reached 46% and 43%, respectively. Several clinical studies have evidenced a lower level of macrolide resistance among invasive strains in comparison to colonizing *S. agalactiae* (Borchardt et al., 2006; Domelie et al., 2008; Fluegge et al., 2005).

Vancomycin is among the antimicrobial agents recommended for perinatal *S. agalactiae* infection prevention within a very small subset of women and is rarely used, namely due to its toxicity (Verani et al., 2010); however, vancomycin is commonly used on empirical therapy of severe mixed infections among adults, which sometimes include *S. agalactiae*. Although the resistance to vancomycin among enterococci and other Gram positive cocci continues to rise in the hospital setting, the first two vancomycin-resistant *S. agalactiae* strains were only identified in early of 2013 (Park et al., 2014; Srinivasan et al., 2014). The collective observations made from
these strains isolated from adults in the United States revealed a putative hot spot for insertion of vanG elements and suggested Enterococcus faecalis as the donor (Mckessar et al., 2000; Park et al., 2014; Srinivasan et al., 2014). These data confirms the threat that resistance to vancomycin may be emerging in S. agalactiae, in which this antibiotic is considered the last resort (Verani et al., 2010).

1.2.3 Molecular Epidemiology

Ten S. agalactiae capsular serotypes have been identified (Ia, Ib, II–IX) based on the different arrangements of the cell surface capsular polysaccharides. Despite the limited diversity of monosaccharide composition and structural motifs, serotypes are antigenically distinct, and can be identified by serotyping which is considered a reference tool for investigating S. agalactiae epidemiology (Haguenoer et al., 2011). The most widely used serotyping method is based on latex agglutination (Afshar et al., 2011; Zuerlein et al., 1991). Moreover, antibiotic resistance among S. agalactiae strains can be related to particular serotypes. In fact, erythromycin resistance has been notable among serotypes Ia, III, and V isolated from neonates and up to 35% of serotype V clinical strains displayed resistance to erythromycin (Fernandez et al., 1998; Figueira et al., 2004; Florindo et al., 2010 - Chapter III; Gherardhi et al., 2007; Lin et al., 2000; Manning et al., 2008).

However, capsular serotyping has several limitations: serotype-specific antibodies are expensive, the technique has a low discriminatory power and the results may be erroneous due to agglutination misinterpretation or due to lack/low expression of CPS (Afshar et al., 2011; Radtke et al. 2010; Sorensen et al. 2010). The capsular switching phenomena (Bellaïs et al., 2012; Davies et al., 2004; Luan et al., 2005), along with the high percentage of nontypeable strains of human origin (up to 32%), either from colonization or infection, constitute other limitations that reinforce the non-applicability of serotyping to determine the clonal relatedness of S. agalactiae strains (Kalliola et al., 1999; Gherardhi et al., 2007; Ramaswamy et al., 2006). For S. agalactiae strains of animal origin, in particular from bovines, this limitation is more evident as the proportion of nontypeable strains is higher (up to 77%) (Bisharat et al., 2004, Ekin & Guturk, 2006; Rato et al., 2013; Zhao et al., 2006).

As the structure of S. agalactiae capsule is determined by genes encoding enzymes responsible for the biosynthesis of the polysaccharides (eps - capsular polysaccharide synthesis), molecular capsular typing focusing on the polymorphism of the eps gene cluster has been proposed to overcome the limitations of serotyping. In fact, eps typing is characterized by higher sensitivity, specificity, reproducibility, and correlation to conventional serotyping, thereby reducing the percentage of nontypeable (NT) strains. The lack of standardization and consensus regarding capsular genotyping methods in S. agalactiae culminated in the appearance of different approaches, all involving the conjugation of two different techniques, such as the PCR amplification and the sequencing of partial regions of eps (Amundson et al., 2005; Borchardt et al., 2004; Florindo et al., 2010).
2010 - chapter III; Kong et al., 2002; Poyart et al., 2007; Sellin et al., 2000; Wen et al., 2006). Except the multiplex PCR assay described by Imperi and co-authors (Imperi et al., 2010), a drawback of the capsular gene typing methods is the impossibility to detect the newest serotype, IX, proposed in 2007 (Slotved et al., 2007).

*S. agalactiae* serotype/genotype distribution among reproductive age women (15 to 49 years of age) that varies according to geographical, temporal and ethnic variables is important due to its correlation with pathogenicity and antibiotic susceptibility (Campbell et al., 2000; Florindo et al., 2010, 2014b – chapters III, IV; Fluegge et al., 2005; Gherardhi et al., 2007; Johri et al., 2006; Manning et al., 2008; Martins et al., 2007, 2012; Slotved et al., 2007; Tazi et al., 2010). For this reason, the development of multivalent capsular polysaccharide-based vaccines for reducing maternal colonization and for preventing vertical transmission to neonates depends on accurate epidemiological data regarding serotype distribution (Johri et al., 2006; Rodriguez-Granger et al., 2012).

Serotypes Ia, Ib, II, III and V have been the most frequently described in European, African, and American studies involving colonizing or invasive *S. agalactiae* strains isolated from neonates and adults, accounting for ≥ 80% of all serotypes (Brunil et al., 2006; Dore et al., 2003; Dutra et al., 2014; Edmond et al., 2012; Florindo et al., 2010, 2014b – chapters III, IV; Gherardhi et al., 2007; Hakansson et al., 2008; Huber et al., 2011; Ippolito et al., 2010; Jones et al., 2006; Kunze et al., 2011; Le Doare & Heath, 2013; Madzivhandila et al., 2011; Manning et al., 2009; Martins et al., 2012; Mottola et al., 2004; Oviedo et al., 2013; Phares et al., 2008; Shabayek et al., 2014; Tsolia et al., 2003; van der Mee-Marquet et al., 2009). Although serotypes VI to IX are rarely found in the aforementioned continents, studies performed in Japan between 1980s and early 2000s showed that serotypes VI and VIII accounted for approximately 60% of the colonizing strains recovered from reproductive age women (Lachenauer et al., 1999; Matsubara et al., 2001; Terakubo et al., 2003). The frequency of the different serotypes change along years, and between 2007 to 2010, serotype distribution studies held in the Japanese cities of Saitama and Kobe (Kimura et al., 2013; Ueno et al., 2012) revealed a clear decrease in the frequency of serotypes VI and VIII.

Until 2010, only one study carried out in Abu Dhabi, United Arab Emirates (Amin et al., 2002), reported a high frequency of serotype IV among colonized pregnant women (26% of all *S. agalactiae* strains isolated between 1998 and 1999). Since 2010, serotype IV became more frequent in studies from Brazil, Ireland, Portugal and United States (Diedrick et al., 2010; Ferrieri et al., 2013; Florindo et al., 2014b – chapter IV; Kiely et al., 2011; Palmeiro et al., 2010) either among colonization or infection cases, suggesting the possibility that this serotype could be emerging as an
important pathogen, as happened with serotype V during the 1990s (Blumberg *et al.*, 1996; Elliott *et al.*, 1998; Skoff *et al.*, 2009).

Epidemiological data collected worldwide revealed that capsular types can be associated with disease condition, where *cps* types Ia, Ib, II, III and V account for 96% and 88% of cases of neonatal and adult invasive *S. agalactiae* infections, respectively (Edmond *et al.*, 2012; Edwards *et al.*, 2005; Martins *et al.*, 2012; Phares *et al.*, 2008). A substantial proportion of EOD and the majority of LOD cases have been associated to *cps* type III (Gherardhi *et al.*, 2007; Jones *et al.*, 2003; Lamy *et al.*, 2006; Manning *et al.*, 2009; Martins *et al.*, 2007; Tazi *et al.*, 2010).

Other molecular biology methods have been developed to improve diagnosis, prognosis of *S. agalactiae* infections and to evaluate the genetic structure of *S. agalactiae* strains. These include pulsed-field gel electrophoresis (PFGE) (Elliott *et al.*, 1998), multilocus sequence typing (MLST) (Jones *et al.*, 2003), and more recently the multilocus variant repeat assay (MLVA) (Haguenoer *et al.*, 2011; Raditke *et al.*, 2010). Pulsed-field gel electrophoresis based on macrorestriction fragment analysis of genomic DNA discriminates *S. agalactiae* strains isolated from different hosts and clinical origins (Benson & Ferrieri, 2001; Gherardhi *et al.*, 2007; Martins *et al.*, 2007 Rato *et al.*, 2013; Tenover *et al.*, 1995). However, PFGE is a labor-intensive, requires experienced personnel, takes several days of work and the limited number of restriction fragments may be suboptimal with respect to its discriminating capacity; moreover, PFGE is an image based method, which makes interlaboratory comparison difficult.

*S. agalactiae* MLST, is based on the sequencing and analysis of internal portion of seven selected housekeeping genes (*adhp, pheS, atr, glnA, sdhA, gicK and tkt*) from the core genome, considered to be selectively neutral and that are located in different parts of the genome. MLST therefore provides a better tool for determination evolutionary relationships contributing to the better resolution of *S. agalactiae* strains and the identification of several genogroups (Jones *et al.*, 2003; Maiden *et al.*, 1998, 2013). The results are unambiguous and easily allow interlaboratory comparison of *S. agalactiae* genetic profiles from distinct geographic areas, clinical origins and hosts and are also suitable for outbreak investigations (Zhao *et al.*, 2008). In addition, MLST has the advantage of reproducibility, being a standard tool for delineating the population genetics structure of *S. agalactiae* (Jones *et al.*, 2003; Lin *et al.*, 2006; Maiden *et al.*, 2013; Sun *et al.*, 2005). The development of curated online MLST reference databases (such as those found in http://pubmlst.org/sagalactiae/, where a total of 710 sequence types, STs, are deposited; accessed on the 18th September, 2014) provided both portable nomenclature schemes and the possibility to infer evolutionary relationships in the population structure of *S. agalactiae* by using eBURST program (Feil *et al.*, 2004), which identifies clonal complexes based on variations in the allelic MLST profiles of analyzed strains and allows a graphic representation of the genetic relatedness
(Fig. 1.3). Although this huge number of STs, MLST studies evidenced that the majority are grouped and constitute a limited number of clonal complexes (CCs) by their similarity to a central allelic profile, the ancestral ST, which is the founder and representative of each CC, namely CC1, CC7, CC12, CC17, CC19, CC23 and CC67/61 (Bisharat et al., 2004; Bohnsack et al., 2008; Fluegge et al., 2011; Jones et al., 2003; Luan et al., 2005, Martins et al., 2007, 2012). \textit{S. agalactiae} strains belonging to these seven CCs have been associated with particular hosts and clinical manifestations. As an example, CC67/61 strains are adapted to the bovine host causing dairy cow’s mastitis, but were never isolated from humans. In addition, \textit{S. agalactiae} isolated from bovines belong to CC67/61 in about two-thirds of the cases (Bisharat et al., 2004; Rato et al., 2013), but may vary by geographic region (Yang et al., 2013). Occasional strains isolated from bovines belong to other STs, such as ST1 and ST23. On the other hand, CC17 strains, which includes the highly virulent serotype III clone identified by Musser and colleagues in 1989, is exclusively adapted to human neonates displaying a rapid and global dissemination (Musser et al. 1989; Sorensen et al., 2010).

![MLST/dBURST diagram](image)

**Figure 1.3.** Example of a MLST/dBURST diagram showing the genetic relationships between human \textit{S. agalactiae} clinical strains isolated from patients with cystic fibrosis (19 strains) and from women with anogenital colonization (72 strains). The \textit{S. agalactiae} strains were clustered within the clonal complexes (CC) indicated in the picture. Sequence types that vary by one allele in their MLST profiles (single locus variants) are arranged in circles around the primary founder sequence type. STs found among \textit{S. agalactiae} colonizing strains are depicted as closed circles; STs found in the \textit{S. agalactiae} strains from patients with cystic fibrosis are shown as open squares. (Adapted from Eickel et al., 2009).

Despite human and bovine \textit{S. agalactiae} strains may represent separate populations, occasional interspecies \textit{S. agalactiae} transmission, namely from bovine hosts to humans was recently observed, but not involving CC67/61 strains (Bisharat et al., 2004; Manning et al., 2010; Martinez et al., 2000). On the basis of MLST data, it has been proposed that ST17 and ST67 strains has arisen from a common bovine ancestor but phylogenetic analysis based on concatenated
sequences of 15 housekeeping genes, including the MLST gene panel of *S. agalactiae*, contradicted this hypothesis (Bisharat *et al.*, 2004; Héry-Arnaud, *et al.*, 2005; Sorensen *et al.*, 2010).

*S. agalactiae* strains of human origin present higher levels of genetic diversity, where the most prevalent CCs are CC1, 12, 17, 19 and 23. Strains belonging to CC1, CC12 are often associated with infections in adults; strains from CC23, CC19 are associated with vaginal carriage and EOD in newborns; and CC17 strains which are mainly responsible for LOD cases, including more than 80% of cases of *S. agalactiae* meningitis (Lin *et al.*, 2006; Manning *et al.*, 2009; Martins *et al.*, 2007; Salloum *et al.*, 2010; Straková *et al.*, 2010; Tazi *et al.*, 2010). Because ST17 strains more frequently cause meningitis than sepsis, and LOD than EOD, it supports the idea that ST17 strains have a particular ability to invade the central nervous system of the neonates (Bisharat *et al.*, 2004, 2005, Jones *et al.*, 2003; 2006; Lartigue *et al.*, 2011; Manning *et al.*, 2009; Tazi *et al.*, 2010). Strains belonging to CC1, 17, 19 and 23 exhibit complex phylogenetic relationships, exhibit diverse capsular types, and include both colonizing and invasive strains, which could suggest an opportunistic pathogenic behaviour. As each MLST lineage contains several different capsular types, the capsular type alone cannot be used as a marker to identify a phylogenetic lineage and it evidences that invasiveness is independent of capsular type (Jones *et al.*, 2006). This phenomenon also suggests that serotype switching occurs within lineages, presumably by horizontal transfer of *cps* genes (Jones *et al.*, 2006, Luan *et al.*, 2005). In fact, a capsular switching from type III to type IV within the highly homogeneous ST17 was recently demonstrated (Bellais *et al.*, 2012). Furthermore, genome sequencing analysis revealed that this capsular switch was due to the exchange of a 35.5 kb DNA fragment containing the entire *cps* operon (Bellais *et al.*, 2012). The changes at the capsular locus were proposed to be driven by the equilibrium between the selective pressure imposed by host immunity and conservation of a particular capsular polysaccharide, as an adaptive advantage of virulent clones (Brochet *et al.*, 2006; Cieslewicz *et al.*, 2005; Luan *et al.*, 2005).

Molecular typing methods based on short clustered regularly short palindromic repeats (CRISPR) and on variable number of tandem repeats at multiple loci (MLVA) of *S. agalactiae* genomes have been shown to be useful for genotyping and showed a perfect match with MLST clonal groups (Haguenoer *et al.*, 2011; Lopez-Sanchez *et al.*, 2012; Radtke *et al.*, 2010). Interestingly, both methods provide higher degrees of diversity by defining subgroups among some MLST lineages, such as CC17 and CC23. Depending on the method and genetic lineage, correlations between subgroups and geographical origin of the strains, host or capsular type were observed (Haguenoer *et al.*, 2011; Lopez-Sanchez *et al.*, 2012); the outcome is a comprehensive characterization of *S. agalactiae* population structure.
1.2.4 Genomics

In 2002, the genome sequencing of two *S. agalactiae* clinical strains of human origin, 2603V/R and NEM316 [genotypes V/ST110 (Genbank Accession no. NC_004116.1) and III/ST23 (Genbank no. NC_004368.1), respectively] has provided unprecedented genetic information (Glaser *et al.*, 2002; Tettelin *et al.*, 2002). The genome of *S. agalactiae* strains 2603V/R and NEM316 consist of a circular chromosome of 2,160,267 bp and 2,211,485 bp, encoding 2,175 and 2,118 predicted genes, with a G+C content of 35.7% and 35.6%, respectively. In both genomes, seven sets of 23S, 5S and 16S ribosomal RNA (rRNA) were identified, as well as 80 transfer RNAs (tRNAs). When compared with other streptococcal species, namely *Streptococcus pyogenes* strain M1 (Ferretti *et al.*, 2001) and *Streptococcus pneumoniae* strain TIGR4 (Tettelin *et al.*, 2001), the genome size of *S. agalactiae* strains was similar to the pneumococcal strain (2.16 Mb), but larger than the *S. pyogenes* genome (1.85 Mb). The G+C content of both *S. agalactiae* strains was lower than the determined for genomes of *S. pyogenes* (38.5%) and *S. pneumoniae* (39.7%), whereas only four and six rRNA operons and 60 and 58 tRNAs were identified, respectively. Although these three streptococcal species colonize different anatomical sites of the human body, they are all capable of causing severe invasive disease. Hence, it is likely that these species share some virulence factors, as well as other genetic features that will surely determine specific colonization, invasion and disease characteristics. In fact, approximately 50% of the genes have homologs among the three species, despite the maintenance of genome architecture being more pronounced between *S. agalactiae* and *S. pyogenes*. These findings may reflect a closer evolutionary relationship between *S. agalactiae* and *S. pyogenes*, which both belong to the pyogenic group of streptococci, while *S. pneumoniae* belongs to the group of viridans streptococci.

The gene repertoire of *S. agalactiae* strains 2603V/R and NEM316 revealed the genetic basis of two surface polysaccharides, the cell wall-associated group B antigen common to all *S. agalactiae* strains (13 and 16 encoding genes, respectively), and the type-specific capsular polysaccharide encoded by *cps* cluster (*cpsA-L*) (Glaser *et al.*, 2002; Tettelin *et al.*, 2002).

One striking genomic feature is that *S. agalactiae* does not have the biosynthetic machinery to produce many essential amino acids, vitamins and co-factors, which must be acquired from external sources (Glaser *et al.*, 2002; Tettelin *et al.*, 2002). This dependence upon host-derived nutrients is reflected in a large investment in import systems; in fact over 10% of the genes in the genome encode components of transporters, where the most abundant class is constituted by ATP binding cassette (ABC) transporters, one of the largest protein families in bacteria (example: almost 5% of the *E. coli* genome) (Linton & Higgins, 1998). Also, this large diversity of transport systems may enable *S. agalactiae* to adapt to different environments and might be involved in its capacity to cause disease (Glaser *et al.*, 2002).
Another finding is the triple presence of a 47,068 bp sequence flanked by inverted repeat (IR) sequences in the genome of NEM316, which was designated pNEM316-L1, a putative integrative plasmid (Glaser et al., 2002). Notably, NEM316-L1 copies and 50% of S. agalactiae genes without an ortholog in S. pyogenes are located within 14 genetic islands (I to XIV) dispersed around the chromosome of NEM316, constituting a particular feature of S. agalactiae (Glaser et al., 2002). These 14 islands are adjacent to tRNAs and are composed by 11 to 77 genes, mostly related to virulence factors, surface proteins and mobile elements. For example, island XII contains the linb and scpB genes, island XIII the CAMP factor, island IV the olp2 gene and island VI contains the cly operon essential for S. agalactiae hemolytic activity and production of pigment (Pritzlaff et al., 2001; Rosa-Fraile et al., 2014; Spellerberg et al., 1999a). The 14 islands are highly variable in frequency and in composition among the different S. agalactiae strains tested and they constitute part of the accessory S. agalactiae genome (Glaser et al., 2002; Herbert et al., 2005; Richards et al., 2011; Rosinski-Chupin et al., 2013; Tettelin et al., 2005). Although these islands could be considered as pathogenicity islands (PAIs), only four (I, VI, XII, and possibly X) could be a true PAI, especially due to the presence of virulence and mobilization genes. While no correlation between PAI and colonization versus invasive strains is noted (Herbert et al., 2005), many findings suggest a correlation of PAI XII with S. agalactiae origin from humans (Al Safadi et al., 2010; Franken et al., 2001; Rosinski-Chupin et al., 2013).

The genetic knowledge on S. agalactiae was incremented in 2005 with the release of six additional full genome sequences of strains isolated from human hosts (Tettelin et al., 2005). Together with NEM316 and 2603V/R genomes, the eight S. agalactiae strains, evidenced similar genome sizes, similar number of predicted genes, an overall identity ranging from 85% to 95% and a high degree of gene synteny (Tettelin et al., 2005). Moreover, genomic analysis suggests that S. agalactiae can be described by its pan-genome, comprising a core genome containing genes present in all strains (approximately 80% of genes) and an accessory genome consisting of partially shared and strain-specific genes (Tettelin et al., 2005), which may influence the adaptation of particular S. agalactiae strains to specific ecological niches and their pathogenic potential (Tettelin et al., 2005). According to Tettelin and co-authors (Tettelin et al., 2005), the gene reservoir of S. agalactiae appears to be vast, as new genes could be successively identified and added to the pan-genome (average of 33 genes per genome sequenced). Mathematical extrapolations predicted that new genes will still be found even after sequencing more S. agalactiae genomes, suggesting that the pan-genome of S. agalactiae is open, as verified in S. pyogenes (Tettelin et al., 2005).

As strain-specific genes tend to cluster into genomic islands, it suggests that lateral gene transfer (LGT) is an important evolutionary force within S. agalactiae (Tettelin et al., 2005). Subsequent studies supported this proposal, where 35 putative integrative conjugative elements (ICE) were identified within the first eight S. agalactiae available genomes, while a combination of
experimental and in silico approaches have shown that large genomic segments (up to 334 Kb) can be exchanged via conjugation among *S. agalactiae* strains, which contributes to the genome dynamics (Brochet et al., 2008). Furthermore, LGT via ICE or other mobile genetic elements (MGE) such as phages have been implicated in LGT between *S. agalactiae* and other streptococcal species (Beres & Musser, 2007; Davies et al., 2005, 2009; Domelier et al., 2009; Ferretti et al., 2001; Haenni et al., 2010; Salloum et al., 2010).

In 2011, Richards and colleagues (Richards et al., 2011) provided the first genome sequence of a bovine *S. agalactiae* strain (FSL S3-026, genotype III-ST67) responsible for clinical mastitis. Remarkably, the total length of this genome was 2 455 848 bp, being 290 Kb (12%) larger than the mean of the genomes of *S. agalactiae* strains A909, 2603V/R and NEM316, with a G+C content of 36.1% (Richards et al., 2011). A distinctive feature of this bovine *S. agalactiae* genome in comparison to *S. agalactiae* genomes of human origin was the high frequency of insertion sequences (IS), which have been proposed to be largely responsible for gene deletions and genome rearrangements (Richards et al., 2011). In fact, the bovine *S. agalactiae* genome contained 97 IS grouped into 14 clusters, representing an unusual high proportion of IS within the genome (4.3%), whereas the human *S. agalactiae* strains presented an average value of 20.4 IS per genome (Richards et al., 2011). According to Richards et al. (Richards et al., 2011), this high proportion of IS in bovine *S. agalactiae* strains may result from evolutionary effects after the isolation of *S. agalactiae* population in cattle (Mahillon & Chandler, 1998). In fact, as IS are often associated with selective disadvantages they are removed from through competition; however, they seem to accumulate only in very small populations where competition is low or nonexistent (Mahillon & Chandler, 1998; Siguié et al., 2006). This hypothesis is consistent with the analysis of Bisharat and co-authors (Bisharat et al., 2004), who suggested a limited genetic diversity of bovine *S. agalactiae* population in comparison to *S. agalactiae* strains from humans. In addition, the identification of 183 bovine strain-specific genes clustered into eight genomic islands, corroborated previous findings showing that strains of bovine *S. agalactiae* strains are evolutionary distinct from human strains (Sorensen et al., 2010). A major factor responsible for this distinctiveness appears to be the LGT (Richards et al., 2011). Indeed, the comparison with other bovine streptococci occupying the same habitat and responsible for mastitis, provided strong evidence of interspecies LGT involving the lactose operon, namely between *S. agalactiae* strain FSL S3-026 and *S. dysgalactiae* subsp. *dysgalactiae* strain ATCC 27957 (Richards et al., 2011). This exchange of genetic material between streptococcal species causing mastitis may have aided the continued adaptation of *S. agalactiae* to the bovine environment by incorporating potential virulence factors into their genome.

The evolution of particular *S. agalactiae* lineages during the time-course of speciation to fish host was recently revealed (Rosinski-Chupin et al., 2013). While *S. agalactiae* CC7 strains
isolated from fish are closely related to the human CC7 counterparts according to their gene content (although displaying large differences in gene expression that can be involved in fish adaptation). ST260-261 piscine *S. agalactiae* strains, a fish-associated clonal complex that has never been reported in humans, are distantly related to human and cattle *S. agalactiae* strains through a massive gene inactivation or through an ongoing insertion-deletion process (Delannoy et al., 2013; Liu et al., 2013; Rosinski-Chupin et al., 2013). Indeed, 190 to 220 pseudogenes were identified in ST260-261 strains in contrast to 27 to 41 within human *S. agalactiae* strains. Consequently, ST260-261 genomes are 10 to 25% smaller than the genomes of other *S. agalactiae* strains, but the G+C content is similar and the number of rRNA operons is equal (Pereira et al., 2013; Rosinski-Chupin et al., 2013). Moreover, all the genomes of fish *S. agalactiae* strains lack two important virulence loci, *imb* and *scpB*, which encode the laminin-binding protein and the C5a peptidase, respectively (Liu et al., 2013; Rosinski-Chupin et al., 2013).

**1.2.5 Pathogenesis**

The development of a bacterial disease has been linked to a molecular arms race, in which the host tries to eliminate the bacteria, while the bacteria try to survive in the host (Bush, 2001; Woolhouse et al., 2002). Although most bacteria cause no disease, some are etiological agents of serious infections. Between these two extremes are *S. agalactiae* that can coexist in humans and other animal hosts in a carriage state or causing disease. Failures of innate immunity define the clinical field of infectious diseases.

The development of disease during the course of *S. agalactiae* infection reflects its successful competition with the local microbial flora, adherence and colonization of epithelial surfaces, penetration through cellular barriers, resistance to immune clearance (allowing bloodstream survival), multiplication at the site of infection and, in cases of meningitis, the ability to breach the blood-brain barrier (Doran & Nizet, 2004; Maisey et al., 2009; Spellerberg, 2000). To overcome physiological obstacles, *S. agalactiae* expresses a high number of virulence factors that are either located on the bacterial surface or are secreted into the surrounding environment that mediates specific host-pathogen interactions, contributing for its pathogenicity; for this reason they are called virulence factors (Doran & Nizet, 2004; Nobbs et al., 2009; Spellerberg, 2000).

*S. agalactiae* adhere to a variety of cells including the female vaginal epithelium, placental membranes, respiratory tract epithelium and blood–brain barrier endothelium. *S. agalactiae* continuously monitors the local environment and fine-tunes the production of adhesins, communication systems, and metabolic pathways to optimize fitness under the prevailing conditions (Doran & Nizet, 2004; Kline et al., 2009; Nobbs et al., 2009). Environmental conditions such as pH, temperature, and oxygen availability, influence the development of *S. agalactiae* communities (Nobbs et al., 2009). In fact, maximal adherence occurs at the acidic pH of vaginal mucosa (Tamura et al., 1994), allowing *S. agalactiae* to occupy a niche that places at risk infants
(vertical transmission during labour) or women (immunocompromised). A low-affinity interaction of *S. agalactiae* with human epithelial cells is mediated by its amphiphilic cell wall-associated lipoteichoic acid, while higher affinity interactions with host cells are mediated by extracellular matrix components: fibronectin, fibrinogen and laminin (Kline et al., 2009; Nobbs et al., 2009). SspB encoded by *scpB* mediates *S. agalactiae* binding to human immobilized fibronectin (Beckmann et al., 2002) and also plays a role in the *S. agalactiae* invasion of epithelial cells (Cheng et al., 2002). Binding of *S. agalactiae* to human laminin, a major glycoprotein of the basement membrane is mediated by the surface-associated lipoprotein Lmb (Ragunathan et al., 2009, 2013; Spellerberg et al., 1999b; Tenenbaum et al., 2007), which shows homology to members of the Lrl family that includes the Lsp/Lbp of *S. pyogenes* (Elsner et al., 2002; Spellerberg et al., 1999b), and promotes *S. agalactiae* invasion of human brain microvascular endothelial cells (Tenenbaum et al., 2007). The attachment of *S. agalactiae* to fibrinogen is mediated by repetitive motifs within surface anchored protein FbsA (Schubert et al., 2002). In addition, adhesion to lung and cervical epithelial cells involve pilus-like structures encoded by three genomic islands (PI-1, PI-2a; PI-2b), which also contribute to adherence and invasion of endothelial cells of the human blood-brain barrier (Maisey et al., 2007; Pezzicoli et al., 2008; Sharma et al., 2013). The examination of *S. agalactiae* clinical isolates showed that pilus DNA sequences in each island are conserved but a difference in the biological outcome was found depending on pilus type (Konto-Gliorghi et al., 2009; Margarit et al., 2009). When the *S. agalactiae* populations recovered from carriage and neonatal infections were compared, the concomitant presence of PI-1 and PI-2a were associated with maternal colonization (and frequent among CC19 strains) while PI-1 and PI-2b were associated with neonatal disease and ST17 strains (Martins et al., 2013; Rinaudo et al., 2010; Sorensen et al., 2010).

After cell adhesion and colonization, and depending on specific host factors, such as a immunocompromised system, *S. agalactiae* can spread to different anatomical sites through the secretion of toxins or virulence factors that facilitate the entry and survival within host cells (Adderson et al., 2003; Doran & Nizet, 2004; Nobbs et al., 2009). The intracellular infection by *S. agalactiae* may result in the loss of integrity of host tissues and, consequently, contributes to the development of the infectious process. The β hemolysin/cytolysin, the CAMP factor and the hyaluronidase are important for *S. agalactiae* invasive infection, not only by damaging host cells but also by promoting the release of the intracellular nutrients that are necessary for the survival of the bacterium (Herbert et al., 2004). The β-hemolysin/cytolysin is an important *S. agalactiae* virulence factor encoded by the *cyl* operon and confers the hemolytic activity, which lead to the total destruction of erythrocytes (Rosa-Fraile et al., 2014; Spellerberg et al., 1999a, 2000). Moreover, this cytolsin causes membrane damage in host cells and has been linked to the destruction of the lung epithelium and endothelial cells by compromising their barrier function.
(Gibson et al., 1999; Liu & Nizet, 2004). Also of note for cellular invasion are the CAMP factor, an extracellular protein that has the ability to form pores in host cells and triggering the cell lysis (Lang & Palmer, 2003); hyaluronidase, encoded by hyl gene, promotes the hydrolysis of hyaluronic acid in the extracellular matrix of animal tissues promoting the spread of the pathogen through tissues and providing nutrients after cell lysis, and the surface-anchored alpha C protein, which specifically interacts with the host cell glycosaminoglycan on the epithelial cell surface to promote group B streptococcal internalization (Baron & Kasper, 2005).

Adhesion may come at a cost because *S. agalactiae* attachment can also stimulate immune cell infiltration, activation and phagocytosis, which will facilitate bacterial clearing (Nobbs et al., 2009). In fact, after reaching the host bloodstream or deeper tissue structures, *S. agalactiae* triggers an immune response, particularly involving neutrophils and macrophages. Effective uptake and killing by these cells require opsonization of the bacterium by specific antibodies in the presence of the complement (Edwards et al., 1980); however, *S. agalactiae* has many virulence factors that confer resistance to opsonization and phagocytosis, promoting evasion from the host immune system. The majority of *S. agalactiae* strains associated with invasive disease are encapsulated, and the sialic acid on the capsular surface confers resistance to opsonization by avoiding the deposition of the C3b of the complement system (Lartigue et al., 2011). In addition, the capsule has the capacity to mimic epitopes of the host, preventing the recognition and blocking the access of the host recognition molecules (Doran & Nizet, 2004; Rajagopal, 2009; Spellerberg, 2000). C5a peptidase also plays an important role at this stage, since it has the ability to cleave the C5a component of the complement system, therefore decreasing the recruitment of neutrophils to the site of infection (Beckmann et al., 2002; Cheng et al., 2002; Maiscy et al., 2008; Spellerberg, 2000). The BibA surface adhesin encoded by gbs2018 gene also plays an important role in inhibiting the function of other components of the complement system, promoting resistance to phagocytosis and enhancing adhesion to epithelial cells (Ring et al. 2002; Tazi et al., 2010). Even if the phagocytic uptake of *S. agalactiae* occurs, the bacteria survive for prolonged periods within the phagolysosome of macrophages (Cornacchione et al., 1998; Sagar et al., 2013; Teixeira et al., 2001). In fact, although *S. agalactiae* lack catalase, it is 10 times more resistant to killing by hydrogen peroxide than the catalase-positive *S. aureus* (Wilson & Weaver., 1985).

Failures of host innate immunity and the presence of a virulence portfolio allow *S. agalactiae* to evade host complement deposition and activation, impede phagocyte recruitment and activation, resist the microbicidal activities of host antimicrobial peptides and reactive oxygen species, escape neutrophil extracellular traps (NETs) and, consequently, promote tissue damage which can cause death.
1.2.6. Extracellular DNases

As key players in the host innate immune response, neutrophils are recruited to sites of infection and constitute the first line of defense. Neutrophils freely circulate in blood vessels and are recruited to the inflammatory sites when the human organism responds to microbial infections. Besides microbial uptake, activated neutrophils can release neutrophil extracellular traps (NETs) by a mechanism called NETosis, upon contact with chemical compounds, host factors, such as activated platelets and inflammatory stimuli (e.g., LPS, IL8) and by microbes (Brinkmann et al., 2004; Guimarães-Costa et al., 2012). The main components of NETs, chromatin DNA, histones and granular antimicrobial proteins, determine their antimicrobial properties (Brinkmann et al., 2004; Papayannopoulos and Zychlinsky, 2009). The externalized NETs bind to a broad variety of microbial pathogens (Figure 1.4) including protozoa, parasites, fungi, and bacteria such as streptococci, which are immobilized and killed by NETs through a combination of nonoxidative and oxidative mechanisms (Beiter et al., 2006; Brinkmann et al., 2004; Buchanan et al., 2006; Hermosilla et al., 2014; Nathan, 2006; von Köckritz-Blickwede & Nizet, 2009; Wartha et al., 2007). Potent nonoxidative killing mechanisms include antimicrobial peptides (AMPs) such as cathelicidins, defensins, cathepsins and other degradative proteases; on the other hand, the generation of antimicrobial reactive oxygen species (ROS) occurs (Ermert et al., 2009; Zawrotniak & Rapala-Kozik, 2013). Thus, it is not surprising that microbial virulence factors have evolved to neutralize NETs through the degradation of its DNA-backbone, avoiding neutrophil killing (Beiter et al., 2006; Berends et al., 2010; Derré-Bobillot et al., 2013; Guimarães-Costa et al., 2014; Sumby et al., 2005). In contrast to intracellular bacterial DNAases that participate in replication, recombination or DNA repair, the extracellular DNAases appear to have a distinct function, in bacterial adaptation. In fact, the contribution of extracellular DNAases in the evasion of NET-mediated antimicrobial activity has been described for several species, including some member of Streptococcus genus, such as S. pyogenes, S. pneumoniae, S. suis and also S. agalactiae (Beiter et al., 2006; de Buhr et al., 2014; Sumby et al., 2005).
While the production of three distinct extracellular nucleases, Nuc I, II and III has been reported for *S. agalactiae* more than 30 years ago (Ferrieri *et al.*, 1980), the genetic basis of DNase production has not been fully characterized. Indeed, in 2013, the genomic background and functional role of an *S. agalactiae* extracellular DNase (Gbs0661, NucA from NEM316 prototype strain) was firstly published (Derré-Bobillot *et al.*, 2013), confirming preliminary findings of Spellerberg and co-workers (Chapter VII), who also described another DNAse, Sak_0220, with significant similarity to Spd3 of *S. pyogenes* streptodornase (Sumby *et al.*, 2005). Gbs0661 displays a high degree of sequence identity with the *S. pneumoniae* EndA and *S. pyogenes* Sdal (Derré-Bobillot *et al.*, 2013). Genome analysis of the available *S. agalactiae* genomes confirm the presence of several putative nucleases, but their intrinsic DNase activity remains undocumented, as well as the correlation of DNase activity with clonal complex, host and tropism. Therefore, the identification of all *S. agalactiae* genes coding extracellular DNases and their biological role in NET degradation/pathogenesis is of enormous importance.
1.3. Objectives

Determining the molecular epidemiology and the virulence factors of *S. agalactiae* have been major research areas, guided by scientific questions that include: a) How safe and effective are strategies aimed at preventing severe neonatal GBS infection? b) What are the trends in serotype distribution and in the antimicrobial susceptibility of *S. agalactiae* strains? c) Do *S. agalactiae* strains have a different capability to cause infection? d) What molecular mechanisms and virulence factors support the *S. agalactiae* infection? e) Which factors explain the leading role of particular *S. agalactiae* strains, such as ST17, in neonatal infections? f) What is the biological role of extracellular DNases of *S. agalactiae*? Do DNases contribute to pathogenesis?

In an attempt to answer some of the above questions, two general objectives were designed, starting with an extensive phenotypic and genetic characterization of a *S. agalactiae* collection, followed by the study of the DNase activity of particular *S. agalactiae* strains.

In detail, the following objectives were pursued, and constituted the subject of each chapter of this Thesis:

1) Evaluation of the accuracy of prenatal culture in predicting intrapartum *S. agalactiae* colonization status, by determining the positive predictive value of *S. agalactiae* cultures at 35-37 weeks of gestation in relation to *S. agalactiae* colonization status at delivery;

2) Assessment of *S. agalactiae* *cps* types and population structure by studying simultaneously colonizing and invasive strains obtained in Portugal, Germany and Angola from human and bovine hosts; also, the evaluation of the antimicrobial susceptibility of colonizing *S. agalactiae* clinical strains isolated in Portugal between 2005 and 2012;

3) Identification of genes encoding extracellular DNases in *S. agalactiae* and evaluation of gene expression and biological role *in vitro*, using human granulocytes;

4) Correlation of the DNase activity displayed by *S. agalactiae* strains with several epidemiological variables, such as host species, capsular type, genetic lineage and clinical origin (carriage or infection).
Accuracy of prenatal culture in predicting intrapartum group B

*Streptococcus* colonization status


*Author contributions*

CF, SV, JL and MJB conceived and designed the project; CF and VD performed the experiments and statistics; JL, IN, IR, PC and LR collected the anogenital swabs for *S. agalactiae* isolation and provided clinical data; CF wrote the paper; MJB and JPG supervised the study; Revision of manuscript: CF, MJB; reviewed the manuscript. Approval: all
Abstract

Objective: To evaluate the positive predictive value (PPV) of group B *Streptococcus* (GBS) cultures at 35-37 weeks of gestation relative to *S. agalactiae* colonization status at delivery. Methods: Rectovaginal swabs from 221 women at labor in four Lisbon hospitals were collected for *S. agalactiae* screening according to the CDC guidelines. Results: The PPV was 24.4%. IAP was administered to 100% of prenatally *S. agalactiae* positive women. There was no case of early-onset *S. agalactiae* disease (EOD). Conclusions: Poor accuracy of prenatal cultures in identifying true candidates for IAP highlights the need for Portuguese clinical and laboratory guidelines to prevent EOD and antibiotic overtreatment of pregnant women.

**Keywords:** *Streptococcus agalactiae*, intrapartum screening, intrapartum antibiotic prophylaxis, early-onset disease
2.1 Introduction

*Streptococcus agalactiae*, group B *Streptococcus* (GBS) has multiple serotypes and is an opportunistic human pathogen that can lead to life-threatening infections in newborns and immunocompromised adults (Edwards et al., 2011). Maternal GBS carriage has been recognized as the major risk factor of early onset disease in newborns (EOD, <7 days of age) (Edwards et al., 2011; Verani et al., 2010). Up to 30% of pregnant women are anogenital colonized, although the carrier status is considered dynamic during pregnancy (Florindo et al., 2010 – chapter III, Verani et al., 2010). CDC guidelines (Schrag et al., 2002; Verani et al., 2010) recommend *S. agalactiae* screening at 35-37 weeks of gestation in order to identify women at risk that should undergo intrapartum antibiotic prophylaxis (IAP) to avoid transmission to the newborn during labor; IAP became the major major responsible for the reduction of EOD in developed countries. Nevertheless, strategies to prevent late onset disease (LOD), which occurs after the first week of life, have yet to emerge, as IAP is unable to avoid LOD. The screening-based approach is challenging, as its efficacy relies on its capacity to predict GBS colonization status at the time of labor. Published reports (El Helali et al., 2009; Lin et al., 2011) showed that both negative (NPV) and positive (PPV) predictive values of prenatal *S. agalactiae* cultures relatively to the *S. agalactiae* status at delivery are suboptimal, especially the PPV. We aimed to evaluate the PPV of *S. agalactiae* positive culture at 35-37 weeks of gestation considering the *S. agalactiae* colonization status at delivery.

2.2 Methods

2.2.1 Patients and study design

Between March 2008 through June 2009, 221 pregnant women presenting a positive result for *S. agalactiae* at 35-37 weeks of gestation from 4 hospitals (Dona Estefânia Hospital, N = 9; Maternity Alfredo da Costa, N = 42; Fernando Fonseca Hospital, N = 67; and CUF Descobertas Hospital, N = 103) were selected for this study. It was not possible to determine the laboratories (private and/or public) where pregnant women performed their *S. agalactiae* prenatal nor the methodologies were used by those laboratories. The unknown colonization status at delivery was also used as an inclusion criterion in order to verify the intrapartum positivity of *S. agalactiae* in this group of women (N = 88). Considering the main focus of this study, and due to budget constraints, women with negative *S. agalactiae* cultures at 35-37 weeks of gestation were excluded. All pregnant delivering before 35 weeks of gestation as well as pregnant that had received antibiotic treatment up to three weeks before admission were excluded.
This study was approved by the ethics board of the involved institutions, and a written informed consent was obtained from all women prior to their enrolment in the study. Information about age, obstetric risk factors, and type of delivery were collected. Later, information on whether newborns developed EOD during the hospital stay was also acquired.

2.2.2 Collection and culture of specimens

A combined rectovaginal swab was collected from each parturient on admission for delivery. Swabs were then maintained in a non-nutritive Amies medium (Biomérieux) at room temperature until processing at the National Institute of Health in Lisbon within 24 hours, according to the described by the CDC guidelines (Schrag et al., 2002). Briefly, each swab was inoculated in Todd Hewitt selective media broth at 37°C, 5% CO₂ for 18 hours and subcultured on Columbia agar supplemented with 5% sheep blood (Biomérieux) at 37°C in 5% CO₂ for an additional period of 24 - 48 hours.

2.2.3 S. agalactiae identification and antibiotic susceptibility testing

S. agalactiae isolates were identified by standard criteria on the basis of colony morphology, Gram staining, nonhydrolysis of aesculin on bile-aesculin agar and group B latex-agglutination test. Antimicrobial susceptibility testing (penicillin G, erythromycin, clindamycin and vancomycin) was performed by E-test according to the Clinical and Laboratory Standards Institute guidelines (CLSI, 2009).

2.2.4 Capsular typing and screening of ST-17 hypervirulent lineage

Capsular typing was performed by using specific antisera for serotypes Ia to V (Essum AB) and cps genotyping (Florindo et al., 2010 – chapter III). The detection of ST17 lineage was achieved by PCR, as described elsewhere (Lamy et al., 2006).

2.2.5 Statistics

Positive predictive value (PPV) of prenatal S. agalactiae (GBS) cultures was calculated through the following formula: [(Number of women GBS +/+ ) ÷ (number of women GBS +/+ and GBS +/–)] × 100%, where GBS +/+ and GBS +/– correspond to intrapartum positive and negative results, respectively (all samples had been positive positive at prenatal stage) (Lin et al., 2011).
2.3 Results

Vaginal-rectal cultures were obtained from 221 *S. agalactiae* positive women on admission for delivery, 118 (53.4%) and 103 (46.6%) from three public and one private hospitals, respectively. Overall, the average maternal age was 30.4 years (range, 14 to 45 years) and the average gestational age at labor was 39.0 weeks (range, 35.9 to 41.4 weeks). The mode of delivery was vaginal or cesarean in 166 (75.1%) and 55 (24.9%) women, respectively. Among 55 women giving birth by cesarean section, 38 (69.1%) were performed electively and 17 (30.9%) were performed after labor, of whom 38 (69.1%) occurred at the private hospital. Of 221 prenatally *S. agalactiae* positive women, only 54 remained positive at delivery, corresponding to a positive predictive value of 24.4%. All these 54 prenatal *S. agalactiae*-positive women received IAP (ampicillin was the first choice for IAP in the four hospitals). However, on a risk-based screening (e.g. preterm delivery), only 11 (5%) would have justified antibiotic treatment.

None of the 88 parturients without prior *S. agalactiae* screening revealed intrapartum *S. agalactiae* colonization; however, 9/9 attending to the private hospital and 17/79 attending to public hospitals [the ones presenting risk factors: preterm deliveries (n = 14); *S. agalactiae* bacteriuria during the current pregnancy (n = 2); previous child with EOD (n = 1)] received IAP.

The serotype distribution showed the predominance of serotype III [25/54 (46.3%)] followed by serotypes Ia [10/54 (18.5%)], II [9/54 (16.7%)], V [7/54 (12.9%)], Ib [2/54 (3.7%)] and IV [1/54 (1.9%)], which was quite similar to that described in Portugal for *S. agalactiae* colonization in last trimester of pregnancy (Florindo et al., 2010 – chapter III). The lineage ST17 was identified in 56% (14/25) of the isolates belonging to serotype III; however, no newborn developed EOD during hospital stay. All clinical isolates were fully susceptible to penicillin G or vancomycin. We observed a resistance rate of 7.4% to erythromycin and 1.8% to clindamycin, which were lower when compared to our previous data (Florindo et al., 2010 – chapter III).

2.4 Discussion

In the present study, public and private hospitals evidenced differences regarding both the *S. agalactiae* screening during pregnancy and the selection of candidates for IAP. As an example, only nine pregnant presented to the private hospital without *S. agalactiae* screening at 35-37 weeks of gestation, and all received IAP vs. 79 attending to public hospitals, where IAP was provided exclusively to the 17 parturients comprehending risks. The lack of *S. agalactiae* screening during pregnancy suggests unawareness, or indifference regarding the free health care provided under the supervision of low-risk pregnancy in Portugal. Thus, social factors contributing to the exclusion from pregnancy surveillance in Portugal seem to need urgent assessment and adjustment.

Although a great heterogeneity of PPVs has been described (El Helali et al., 2009; Lin et al., 2011; Vankenburg-van den Berg et al., 2010) ranging (43% to 100%), our study revealed a
considerably low PPV of 24.4%. This weak concordance between prenatal and intrapartum culture results could be attributed to several variables, namely 1) timing of prenatal *S. agalactiae* screening; 2) laboratory methodologies; and 3) antibiotic usage. The influence of the timing of prenatal *S. agalactiae* screening in this study would be neglectable; in fact, all enrolled *S. agalactiae* positive women were screened at 35-37 weeks of gestation, which has been considered ideal for correlating with *S. agalactiae* colonization status at delivery, by longitudinal studies (Vankenburg-van den Berg et al., 2010). The observed discrepancy could be explained, in part, by methodological heterogeneity (sampling, swab storage and transport, and culturing procedures) that could not be determined in the present study. In fact, in Portugal, the health system allows pregnant women to freely choose the laboratories where antenatal *S. agalactiae* screening is performed (screening at the same hospital of delivery is rare), implicating that a multitude of laboratories were involved, each one using their particular *S. agalactiae* detection protocols. There are neither Portuguese *S. agalactiae* laboratory screening guidelines nor recommendations for following scientific guidelines internationally accepted, such as the provided by the CDC. In this scenario, a heterogeneity of methodologies applied to *S. agalactiae* detection are to be expected, comprehending the proposed by the CDC guidelines but also less expensive and time consuming procedures, such as direct plating in both Columbia 5% sheep blood agar and chromogenic medium (such as Strepto B ID or Granada). Each procedure has inherent limits and drawbacks that can lead to *S. agalactiae* misidentification. Another technical explanation could hold on the proliferation of non-*S. agalactiae* isolates during storage and transport, such as *Enterococcus* and *Proteus* species, impairing the identification and recovery of *S. agalactiae* on blood agar plates. Indeed, and consistent with published data (Tazi et al., 2008), 12.2% of our intrapartum cultures from prenatally *S. agalactiae* positive women evidenced an overgrowth of Gram negative bacteria in blood agar plates (not supplemented with antibiotics), which might have obscured *S. agalactiae* colonies culminating in false-negative results. This emphasized the need to improve the subculture system by using selective *S. agalactiae* media (Columbia agar with colistin and nalidixic acid, or a commercial chromogenic agar), as is currently recommended by the 2010 CDC guidelines (Verani et al., 2010). Indeed, Van Dyke and colleagues (Van Dyke et al., 2009) revealed that 61.4% of EOD cases occurred in term newborns whose mothers were *S. agalactiae*-negative at 35-37 weeks. Whether those negative cultures were false-negative results or the parturients acquired *S. agalactiae* during the interval between pregnancy screening and delivery is unknown, but it surely evidences major variations in *S. agalactiae* colonization status during pregnancy. As we excluded pregnant subjected to antibiotic treatment within three weeks before delivery, we would expect no influence of this factor for the low PPV; however, we cannot exclude that some pregnant women during their hospital admission questionnaire omitted (by unidentified reasons) taking medication, namely antibiotics. In fact, although in Portugal a medical prescription is required for antibiotic purchase, irregularities to this rule exists, allowing self-medication.
In conclusion, the reasons underlying a low PPV of prenatal culture in predicting *S. agalactiae* colonization during labor in Portugal are hard to determine due to the lack of national clinical and laboratory guidelines for *S. agalactiae* prevention that would contribute to the uniformity and quality of *S. agalactiae* screening. This requisite would surely contribute to a higher PPV that would prevent EOD while avoiding overtreatment of pregnant women.

Also, and until the availability of an effective *S. agalactiae* vaccine, new reliable and fast intrapartum diagnostic tools should be developed to supplement antenatal *S. agalactiae* screening.

### 2.5 Acknowledgements

The authors thank Albertina Louro (National Institute of Health); Conceição Telhado, Mariana Loureiro, Glória Carvalhosa, Mafalda Lucas, Maria João Santos, Maria do Céu Ramalho, Rosa Portela, Luís Mós, Anaelde Araújo, and Cristina Castel Branco (CUF Descobertas Hospital); Ana Lima, and Patrícia Silva (Maternity Alfredo da Costa); Graciete Novais, Elsa Guerreiro, Cecília Arrilha, Pedro Mendes, António Faria, and Lilita Miranda (Fernando Fonseca Hospital) for collecting/processing clinical samples.

Declaration of Interest: The authors report no conflicts of interest. This work was supported by Fundação para a Ciência e Tecnologia grant PTDC/SAU10MII/105114/2008 from Ministério da Ciência, Tecnologia e Ensino Superior, Portugal and, PhD grant SFRH/BD/48231/2008 awarded to CF.
Molecular characterization and antimicrobial susceptibility profiles in *Streptococcus agalactiae* colonizing strains: association of erythromycin resistance with subtype III-1 genetic clone family

Florindo C, Viegas S, Paulino A, Rodrigues E, Gomes JP, and Borrego MJ

Manuscript with minor changes published in 2010 in *Clinical Microbiology and Infection* 16: 1458-1463. *With kind permission from Wiley*

*Author contributions*
CF, SV and MJB conceived and designed the study; SV and AP isolated *S. agalactiae* strains; SV and CF performed the antibiotic susceptibility testing; CF performed all the molecular characterization; ER performed the statistics; CF, SV, JPG and MJB analyzed the data; CF wrote the paper; MJB, JPG and SV supervised the study; CF, JPG, SV and MJB reviewed the manuscript. Approval: all.
Abstract

Knowledge of the epidemiology of *Streptococcus agalactiae* in Portugal is limited: therefore, the present study aimed to investigate the carriage rate of *S. agalactiae* among Portuguese women of reproductive age and the prevalence of antibiotic resistance, as well as to perform a molecular characterization of the clinical isolates. *S. agalactiae* was recovered from 6.2% of 4269 women during the period 2005–2007, with a predominance of capsular genotypes III (35%), V (33%), Ia (16%) and II (10%) in a sample of 100 isolates. To our knowledge, this is the first report of the *S. agalactiae* colonization rate in Portugal determined according to CDC guidelines. All isolates were susceptible to penicillin and vancomycin, whereas resistance to clindamycin and erythromycin was detected in 10% and 19% of isolates, respectively. Among the 19 erythromycin-resistant isolates, ten (53%) displayed the constitutive MLS$_B$ phenotype (conferring high level resistance to macrolides), eight (42%) had the inducible MLS$_B$, and the M phenotype accounted for one isolate (5%). *erm* methylase genes were exclusively associated with MLS$_B$ phenotype isolates, whereas the M phenotype was a result of the presence of *mefA*. Multilocus sequence typing analysis of the genetic relatedness among isolates presenting resistance to erythromycin demonstrated a novel association between erythromycin resistance and the subtype III-1/ST19 genetic clone family.

**Keywords:** *S. agalactiae*, antibiotic resistance, capsular genotyping, MLST, Portugal
3.1 Introduction

*Streptococcus agalactiae* is a commensal bacterium of the human gastrointestinal and genital tracts (Schuchat, 1999) and recent studies have reported asymptomatic colonization rates of up to 36% in healthy women (Brimil *et al.*, 2006; Hansen *et al.*, 2004; Morlova *et al.*, 2004; Yucesoy *et al.*, 2004). Moreover, it represents a major cause of bacterial infections in newborns (Schuchat, 1999). Penicillin is the antibiotic of choice for the prophylaxis and treatment of *S. agalactiae* infections. However, there are many penicillin-allergic individuals who require the use of second-line antibiotics (Schrag *et al.*, 2002) to which a resistance increase has been described in several countries (de Azavedo *et al.*, 2001; Fitoussi *et al.*, 2001; Gygax *et al.*, 2006). Streptococcal resistance to macrolides is commonly mediated by two major mechanisms (Leclercq, 2002): (i) the effects on the macrolide-specific efflux mechanism (*M* phenotype), encoded by the *mefA* gene, and (ii) modification of the antibiotic binding site of 23S rRNA methylases, encoded by the *erm* genes (MLSb phenotype), which can be either inducible (iMLSb) or constitutive (cMLSb). The most common classification of *S. agalactiae* strains is based on the capsular polysaccharide, defining nine recognized capsular serotypes (Ia, Ib, II–VIII) among which the most common are Ia, II, III and V (accounting for 80% or more in the United States and Western Europe) (Hickman *et al.*, 1999). Multilocus sequence typing (MLST) has been used for the evaluation of the *S. agalactiae* population structure, genetic lineages and, most of all, for the investigation of virulence potential and tropism (Jones *et al.*, 2003).

In the present study, we aimed to: (i) evaluate the *S. agalactiae* colonization rate among women of reproductive age living in the Lisbon metropolitan area (because maternal *S. agalactiae* colonization is a major risk factor for early-onset neonatal disease); (ii) define the capsular genotype distribution; (iii) determine the prevalence of antibiotic resistance in *S. agalactiae* and its mechanisms; and (iv) identify the genetic lineages among the antibiotic-resistant *S. agalactiae* clones. Clarification of these issues is important for understanding the pathogenicity of *S. agalactiae* and for the implementation of prophylactic measures.

3.2 Materials and Methods

3.2.1 Strain collection

Between January 2005 and December 2007, 4269 women of reproductive age (15–49 years) of which 1310 were pregnant, and attending general practice, gynaecology and family planning clinics located in the Lisbon area, were screened for *S. agalactiae* colonization at the National Institute of Health in Lisbon. In accordance with the CDC guidelines (Schrag *et al.*, 2002), separate lower vaginal and rectal swabs were collected and both swabs were placed in a single tube containing Todd-Hewitt selective enrichment broth (Oxoid). Subcultures on 5% sheep-
blood agar plates were performed, and *S. agalactiae* strains were identified by standard criteria on the basis of colony morphology, Gram staining, nonhydrolysis of aesculin on bile-aesculin agar and group B latex-agglutination assay (Becton Dickinson). Budget constraints prevented extensive molecular characterization (at least eight loci for each bacterial isolate) and antibiotic susceptibility testing of all *S. agalactiae* isolates. Therefore, a sample of 100 isolates, comprising the first 34, 33 and 33 isolates from 2005, 2006 and 2007, respectively, was adopted (see Statistical analysis).

### 3.2.2 Antimicrobial susceptibility testing and macrolide resistance phenotypes

MICs were determined by E-test (AB Biodisk). Each strain was tested for its susceptibility to four antibiotics (penicillin G, erythromycin, clindamycin and vancomycin), in accordance with the CLSI guidelines (CLSI, 2009). The constitutive, inducible and M resistance phenotypes were identified by the double-disc diffusion method, as described previously (de Azavedo *et al.*, 2001). The presence of the resistance genes *erm*TR, *erm*B and *mef*A was also investigated. In brief, total *S. agalactiae* DNA was isolated by using the QIAamp DNA mini kit (Qiagen) in accordance with the manufacturer’s instructions, and was subsequently amplified through previously reported primers (Table 3.1) and the multiplex PCR technique (Gygax *et al.*, 2006; Sutcliffe *et al.*, 1996). Briefly, the PCR reaction used 1X optibuffer (Bioline), 0.2 mM dNTPs (Bioline), 2.8 mM MgCl₂, 25 pmol of each primer (MGW Biotech), 1.5 U of bio-x-act DNA polymerase (Bioline) and 5 μL of template for a final reaction volume of 25 μL. The thermocycling profile consisted of an initial denaturation step at 95°C for 5 min, followed by 35 cycles at 95°C for 30 s, 56°C for 30 s and 70°C for 50 s. The final extension step consisted of 10 min at 70°C.

### 3.2.3 Capsular genotyping and MLST

The *S. agalactiae* isolates were subjected to capsule genotyping by a method previously described by Kong *et al.* (Kong *et al.*, 2002) with some modifications. New PCR primers (Table 3.1) were designed based on capsular polysaccharide synthesis D, E and F gene cluster from *S. agalactiae* reference strains of genotypes Ia, Ib and II to VII (Kong *et al.*, 2002). For clarification purposes, and because we used DNA-based typing methods (as is most common among recent *S. agalactiae* studies), capsular types are designated as genotypes throughout the text, instead of serotypes. PCR reagents and thermocycling profiles were the same as above, except for the annealing temperature (51°C), and extension step (2 min for 20 s at 68°C). Purified amplicons were sequenced using the ABI Prism 3700 DNA sequencer (Applied Biosystems) (for amplification primers and internal primer, see Table 3.1). Nucleotide sequences of 1625 or 1634 bp, within the approximately 1.9 Kb amplicon (positions 565–2189 bp from start codon of *cps*D, relative to the genome sequence of *S. agalactiae* strain 2603VR (GenBank accession number AE009948) were aligned through Lasergene99 software (DNASTAR) with GenBank available *cps*D-*cps*E-*cps*F sequences of *S. agalactiae* reference strains (GenBank accession numbers AF332893 for Ia/090,
AF332894 for Ia/NCDC SS615, AF332903 for Ib/H36B, AF332905 for II/18RS21, AF332900 for III-1/GB00-009, AF332896 for III-2/M781, AF332897 for III-3/NCDCSS620, AF381030 for III-4/WC3935, AF332908 for IV/3139, AF332910 for V/CJB111, AE009948 for V/2603V-R, AF349539 for V/CNCTC 1-82, AF337958 for VI/NT6 and AF332913 for VII/7271) aiming to identify the capsular type and/or subtype (Ia, Ib, II, III-1, III-2, III-3, III-4 and IV to VII) of each S. agalactiae isolate. Except for serotype VIII and IX, this modified methodology allowed total discrimination between the S. agalactiae genotypes and subtypes of genotype III, as against the 790-bp region described by Kong et al. (Kong et al., 2002).

The MLST analysis of macrolide-resistant strains was performed as described previously (Jones et al., 2003). Alleles for the seven loci were analyzed on the MLST website (http://pubmlst.org/sagalactiae), and the combination of these results provided an allelic profile or sequence type (ST). eBURST V3 software (http://eburst.mlst.net) was used to define relationships between STs.

Table 3.1 Oligonucleotide primers used for PCR and sequencing. (Adapted from Florindo et al., 2010).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ermB</td>
<td>ermB-1</td>
<td>GAAAGGTACTCAACCAATA (upper)</td>
<td>639</td>
</tr>
<tr>
<td></td>
<td>ermB-2</td>
<td>AGTAACCCCTACTTATAATCTGTTG (lower)</td>
<td></td>
</tr>
<tr>
<td>ermTR</td>
<td>ermTR-1</td>
<td>GAAGTGTTACCTTCTCTAA (upper)</td>
<td>395</td>
</tr>
<tr>
<td></td>
<td>ermTR-2</td>
<td>GCTTACGCACTTGTCTAATTTT (lower)</td>
<td></td>
</tr>
<tr>
<td>mefA</td>
<td>mefA-1</td>
<td>AGATCCTTAAATCAGGATG (upper)</td>
<td>346</td>
</tr>
<tr>
<td></td>
<td>mefA-2</td>
<td>TTCTTGCTACTAAAGTGA (lower)</td>
<td></td>
</tr>
<tr>
<td>16S rRNA</td>
<td>16S-1</td>
<td>GGAGGAAAGTGGGAGGATGG (upper)</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>16S-2</td>
<td>ATGGGATGACGCGCGCGTGG (lower)</td>
<td></td>
</tr>
<tr>
<td>cysD²</td>
<td>cysD-1</td>
<td>GTTGTGGATGCCGAGAATACT (upper)</td>
<td>1902 or 1911</td>
</tr>
<tr>
<td>cysF²</td>
<td>cysF-1</td>
<td>CTCAGCGGGCACTGATGATA (lower)</td>
<td></td>
</tr>
<tr>
<td>cysE²</td>
<td>cysE-1</td>
<td>TCTACGCTAAAGTTTACG</td>
<td></td>
</tr>
</tbody>
</table>

²PCR primers also used for automated sequencing,
³Primer only used for automated sequencing.
3.2.4 Statistical analysis

For the capsular genotyping, MLST analysis, and the antimicrobial susceptibility evaluation, the sample size was set at 100 strains, considering an expected macrolide resistance of approximately 18% (based on a previous epidemiological evaluation performed in Lisbon) (Figueira-Coelho et al., 2004) for 95% CI with 6% precision for a finite population. Associations between categorical variables were calculated using the Fisher’s exact test with a significance level of 5%. All presented statistical results were obtained using SPSS, version 16.0 (SPSS Inc).

3.3 Results

Among the 4269 women of reproductive age, 263 (6.2%) were culture positive for *S. agalactiae*. Similar anogenital *S. agalactiae* colonization rates were observed among nonpregnant (181/2959; 6.1%) and pregnant women (82/1310; 6.3%), as described previously (Brimil et al., 2006). The 100 randomly selected isolates were fully typeable and belonged to genotypes I to V, with predominance of genotypes Ia, II, III and V, which together represented 94% of all isolates (Table 3.2). No *S. agalactiae* isolate showed resistance to penicillin G (MIC 0.032–0.125 mg/L) or vancomycin (MIC 0.25–1 mg/L). Erythromycin resistance, however, was identified in 19% of strains (N = 19), whereas ten of these were also resistant to clindamycin (Table 3.2). Regarding the resistance mechanisms of the 19 macrolide-resistant strains, ten displayed the constitutive MLSB phenotype, eight the inducible MLSB phenotype, and one the M phenotype (for which the MICs of erythromycin and clindamycin were 4 and 0.064 mg/L, respectively). With the multiplex PCR assay, it was possible to test for the presence of the genes responsible for the macrolide resistance phenotypes (Table 3.2).
Table 3.2 Antibiotic resistance and molecular typing of colonizing group B streptococcal isolates. (Adapted from Florindo et al., 2010).

<table>
<thead>
<tr>
<th>Capsular type and subtype</th>
<th>Number of strains (n = 100)</th>
<th>Erythromycin resistance only, n (%)</th>
<th>Clindamycin resistance only, n (%)</th>
<th>Resistance to both, n (%)</th>
<th>Total resistance, n (%)</th>
<th>Number of strains carrying antibiotic resistance* genes</th>
<th>MLST genetic lineages (STs)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>1 (6.2)</td>
<td>1 (6.2)</td>
<td>1 0 0</td>
<td>ST-23 (n = 1)</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1 (33.3)</td>
<td>1 (33.3)</td>
<td>1 0 0</td>
<td>ST-28 (n = 1)</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0 0</td>
<td>ST-44 (n = 1); ST-45 (n = 1); ST-369 (n = 1)</td>
</tr>
<tr>
<td>D</td>
<td>15</td>
<td>7 (46)</td>
<td>0</td>
<td>6 (37.1)</td>
<td>12 (73.3)</td>
<td>4 6 1</td>
<td>ST-37 (n = 2); ST-104 (n = 2)</td>
</tr>
<tr>
<td>E</td>
<td>22</td>
<td>7 (31.8)</td>
<td>0</td>
<td>5 (22.7)</td>
<td>12 (54.5)</td>
<td>3 6 1</td>
<td>ST-17 (n = 1)</td>
</tr>
<tr>
<td>F</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0 0</td>
<td>ST-1 (n = 1); ST-2 (n = 2); ST-10 (n = 1)</td>
</tr>
<tr>
<td>G</td>
<td>11</td>
<td>2 (18)</td>
<td>0</td>
<td>2 (18)</td>
<td>1 (9.1)</td>
<td>2 2 0</td>
<td></td>
</tr>
</tbody>
</table>

*None of the strains carried more than one resistance gene.
*Refers to the 19 erythromycin-resistant S. agalactiae strains.
*One strain did not harbour any of the three resistance genes.
*One strain showed the co-resistance MLSB phenotype.
MLST, multi-locus sequence typing.

Each screened strain presented only a single resistance gene. As expected, \textit{S. agalactiae} strains presenting the cMLS\textsubscript{B} phenotype were highly resistant to erythromycin and clindamycin (MICs $\geq 256$ mg/L) because of the presence of the \textit{ermB} gene. One exception occurred, where a highly resistant strain presented the \textit{ermTR} gene. By contrast, almost all of the iMLS\textsubscript{B} resistance phenotypes were conferred by the presence of the \textit{ermTR} gene. An exception occurred for one erythromycin resistant strain (MIC = 2 mg/L), which exhibited the iMLS\textsubscript{B} phenotype but did not yield any PCR product. The occurrence of inhibition was improbable because the amplification of the internal control was observed. As expected, the \textit{mefA} gene was identified in the single \textit{S. agalactiae} strain displaying the drug efflux mechanism. The data obtained in the present study revealed that macrolide resistance was not equally distributed among the \textit{S. agalactiae} genotypes (Table 3.2). Indeed, \textit{S. agalactiae} strains expressing the capsular genotypes III or V accounted for a higher proportion of the erythromycin-resistant strains (17 of 19; 90%), followed by genotypes Ia and Ib (one strain each). Furthermore, we observed that 12 of the 13 resistant strains expressing genotype III belonged to capsular subtype III-1 ($\chi^2$ test, $p < 0.001$), where 50% of these displayed the \textit{ermTR} gene, reflecting macrolide resistance and susceptibility to lincosamides (Table 3.2). The other two resistance genes were also found in the strains with subtype III-1. One III-1 strain did not harbour any of the three resistance genes. MLST analysis of 19 macrolide-resistant strains demonstrated the existence of different genetic lineages, including among strains expressing the same capsular genotype (Table 3.2). In particular, we identified five distinct STs (ST19, ST27, ST44, ST106 and ST369) among the strains with subtype III-1 (N = 12), although the majority presented the ST19 genetic lineage (N = 7). Despite the high number of STs, strains expressing subtype III-1 were clustered together in the same eBURST group (Fig. 3.1), which only corresponds to part of the clonal complex 19 (CC19); this phenomenon may be a result of the
limited number of strains that were studied. Thus, the majority of ST’s were single-locus or double-locus variants of each other, constituting a clonal cluster represented by ST19 and associated ST’s.

Figure 3.1 eBURST diagram of the genetic lineages among erythromycin-resistant *Streptococcus agalactiae* strains belonging to subtype III-1. Two single locus variants (ST44 and ST27) derived from the primary founder, ST19. ST106 and ST369 are descendents of ST27, and double locus variants of the primary founder. The diameter of the circles is proportional to the number of strains. (Adapted from Florindo et al., 2010).

3.4 Discussion

In Portugal, antenatal screening for *S. agalactiae* in routine clinical settings often reveals substantial discrepancies from the CDC guidelines. Because maternal colonization is the most important risk factor for invasive *S. agalactiae* neonatal disease, and because a previous Portuguese study reported 6.6% mortality in *S. agalactiae* infections of newborns (Neto, 2008), we aimed to determine the streptococcal colonization rate among women of reproductive age according to CDC guidelines.

The colonization rate obtained (6.2%) was low compared to a recent Portuguese study (20%) (Martinho et al., 2008); this could be explained by the application of different experimental methodologies. The results obtained in the present study are in accordance with the low prevalence rates described for some Southern European countries, namely Greece (6.6%) and Turkey (6.5%) and in contrast to the high rates described in Scandinavia (25.4–36%) (Hakansson et al., 2008; Hansen et al., 2004; Tsolia et al., 2003; Yuccsoy et al., 2004). The capsular cps-based genotyping identified several *S. agalactiae* types. Genotypes III (35%), V (33%), Ia (16%) and II (10%) were the most common, contrasting with the low prevalence of genotypes Ib and IV (3% each). The *S. agalactiae* capsular distribution was similar to that previously described in Lisbon (1999–2002, vaginal strains) (Figueira-Coelho et al., 2004) and these apparently have been the most successful capsular genotypes in the Lisbon metropolitan area in last decade. However, *S. agalactiae* capsular
typing in other studies demonstrated the predominance of other genotypes (IV in United Arab Emirates and VI–VIII in Japan) (Amin et al., 2002; Lachenauer et al., 1999); this could reflect specificities of immune responses which may vary according to the studied population. The recent emergence of S. agalactiae strains with reduced penicillin susceptibility in Japan and in the United States constitutes a major threat to the use of penicillin in prophylaxis (Dahesh et al., 2008; Kimura et al., 2008). The molecular characterization of those strains revealed a mutagenic pathway similar to that observed a few decades ago, when the first β-lactam resistant S. pneumoniae strains were identified. This suggests a potential risk of failure for intrapartum antibiotic prophylaxis with β-lactams for S. agalactiae in the near future. Moreover, 19% of the S. agalactiae isolates were resistant to erythromycin and 53% of these were resistant to clindamycin, which comprise be second-line choices of antibiotic. In respect of erythromycin resistance among colonizing strains of S. agalactiae, our results were similar to those described in France and Canada (18%) (de Azavedo et al., 2001; Fitoussi et al., 2001), but differed considerably from the 3.8% reported in Czech Republic (Motlova et al., 2004) and 38–41.9% in the United States (Gygax et al., 2006; Borchardt et al., 2006). Together with social determinants and differences of health care structures, the factors most frequently associated with these large discrepancies in antimicrobial resistance are the consumption and inappropriate use of antibiotics (Brounzwor et al., 2002). In Portugal, no significant change in the consumption of macrolide antibiotics was registered during the period 1997–2006 (Coenen et al., 2006). However, the uptake of intermediate and long-acting macrolides (clarithromycin and azithromycin, respectively) increased significantly during this period, and these agents potentially enhance resistance selection compared to shortacting macrolides (e.g. erythromycin) (Coenen et al., 2006). In the present study, macrolide resistance was found to be predominantly the result of ribosomal methylation (18 of the 19 resistant strains). Indeed, only one strain displayed the M phenotype, as confirmed by the presence of the mefA gene. The prevalence of iMLS$_B$ and cMLS$_B$ phenotypes was similar (eight and ten strains, respectively); however, one strain with the iMLS$_B$ mechanism was not associated with either the mef or the erm genes, suggesting that it could be related to point mutations in the ribosomal L4 and L22 proteins preventing antibiotic binding, as suggested by Diner and Hayes (Diner & Hayes, 2009). The mefA gene has been shown to be mobile in a variety of Gram-positive bacteria (Luna et al., 1999) and its low frequency among our isolates of S. agalactiae is surprising. Indeed, considering the remarkable differences in the microflora of diverse anatomic sites (e.g. genital vs. respiratory tract), dissimilar horizontal gene transfer events are expected to occur, leading to diverse antibiotic resistance spread. The study of these phenomena will be important for understanding how macrolide resistance is evolving. In agreement with previous studies (Figueira-Coelho et al., 2004; Fitoussi et al., 2001), the data obtained in the present study demonstrate that erythromycin resistance was more frequent among genotype III strains (13/19; 68%); this could comprise a major public health concern which originated through the diversification of the founding sequence type (ST19). These
results suggest the genetic clustering of a macrolide-resistant clone family within the capsular subtype III-1 population. Indeed, the only available data worldwide, along with those obtained by ourselves, also showed the clonal spread of macrolide-resistant *S. agalactiae* strains, despite the involvement of a different genotype (V/ST1) (Manning et al., 2008). To our knowledge, these two studies comprise the only data available correlating macrolide resistance and MLST.

In conclusion, the present study demonstrates the higher prevalence of *S. agalactiae* genotypes III and V among Portuguese women of reproductive age. We also found an important association between macrolide resistance and the subtype III-1/ST19 clonal complex, where the MLSB phenotype was the most frequent. Knowledge of the local distribution of capsular genotypes and MLST lineages is crucial for the development of an effective vaccine against *S. agalactiae*. This could be an attractive alternative to the intrapartum antibiotic prophylaxis, which may soon be of limited value owing to the emergence of antibiotic-resistant strains.

### 3.5 Acknowledgements and Transparency Declaration

The authors are grateful to Prof. Dr. Barbara Spellerberg (Institute of Medical Microbiology and Hygiene, University of Ulm) for critically reviewing this manuscript.

This work was partly supported by Comissão de Fomento da Investigação em Cuidados de Saúde (Grant 123/2007). The authors have no conflicts of interest to declare.
Epidemiological surveillance of colonizing group B Streptococcus epidemiology in the Lisbon and Tagus Valley regions, Portugal (2005 to 2012): emergence of a new epidemic type IV/clonal complex 17 clone


Manuscript with minor changes published in 2014 in Euro Surveillance 19: pii=20825
With kind permission from European Centre for Disease Prevention and Control (ECDC)

Author contributions
CF, ISS, MJB, FPM, RC involved in the methodological design; CF, VD, IS, CFa, and the members of the Group for the Prevention of Neonatal GBS Infection were involved in strain characterization and data analysis; ISS, FMP and MJB supervised the study; CF wrote the first draft; Draft revision and approval: all

Note
Part of the results were included in the Thesis of Inês Silvestre for Master Degree in Medical Microbiology, entitled Evolução dos genótipos de Streptococcus agalactiae associados à colonização na grávida, FCT/UNL, December 2013.
Chapter IV

Abstract

This study presents the serotype distribution and the antibiotic resistance profile of 953 colonizing group B *Streptococcus* recovered from women of child bearing age (15 to 49 years) between 2005 and 2012 in the Lisbon and Tagus Valley region, Portugal. Overall, serotypes Ia, II, III, and V were the most common, accounting 752 of the 953 isolates (about 80%). However, there were changes in *S. agalactiae* distribution, in particular in the two last years of the study. Of note, the proportion of serotype IV isolates increased from 1% (2/148) in 2006 to 20% (19/97) in 2012. Also, considerable proportions of serotype IV isolates from 2010 to 2012 were respectively resistant to erythromycin (9/43; 21%) or clindamycin (6/43; 14%). The identification of nine serotype IV isolates presenting a novel association with the clonal complex (CC) 17 lineage, involving a putative capsular switch, may accentuate their virulence potential and ecological success. Molecular analysis of this subgroup of isolates revealed the presence of *rib*, IS861 (insertion sequence) and GBS1I within the C5a peptidase gene (*scpB*) – laminin-binding protein gene (*lmb*) region, reflecting high clonality and a putative common origin. A close surveillance of the emergent type IV/CC17 isolates is crucial considering the potential impact over *S. agalactiae* treatment guidelines and capsular vaccine development.

**Keywords:** *Streptococcus agalactiae*, antibiotic resistance, capsular genotyping, serotype IV
4.1 Introduction

*Streptococcus agalactiae*, group B *Streptococcus* is an opportunistic microbial agent of neonatal pneumonia, sepsis and meningitis in human newborns (Edwards & Nizet, 2011). *S. agalactiae* is also a significant cause of morbidity and mortality in non-pregnant adults, particularly those with underlying medical conditions and in the elderly (Edwards & Nizet, 2011). Up to 36% of pregnant women are anogenitally colonised, although the carrier status is considered dynamic during pregnancy (Barcaite et al., 2008; Edwards & Nizet, 2011). In newborns, maternal *S. agalactiae* carriage has been recognised as the major risk factor of early onset disease (EOD, <7 days of age), but bacteria can also be acquired through horizontal nosocomial transmission (Edwards & Nizet, 2011).

Classification of *S. agalactiae* serotype is based on 10 immunologically unique capsular polysaccharides (Ia, Ib, II-IX), whose prevalence varies according to geographical location, time of study and ethnicity (Edwards & Nizet, 2011; Slotved et al., 2007). Thus, the continuous monitoring of circulating *S. agalactiae* isolates is important in assessing changes in *S. agalactiae* serotype distribution, which is essential for the development of polysaccharide-based vaccines suitable for different geographical areas (Johri et al., 2006; Rodriguez-Granger et al., 2012). Serotypes Ia, II, III and V have been the most frequently described in European countries such as the Czech Republic, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Sweden, and the United Kingdom (Florindo et al., 2010 – chapter III; Ippolito et al., 2010), as well as in the United States (US) (Ippolito et al., 2010), whereas serotypes VI and VIII, to date scarcely found in these countries, could frequently be identified in Japan (Lachenauer et al., 1999). With the exception of a study carried out in Abu Dhabi, United Arab Emirates, where serotype IV predominated among colonized pregnant women (15/57, 26% of the *S. agalactiae* isolates) (Amin et al., 2002), there are few reports among other countries worldwide of serotype IV as a predominant serotype both in cases of colonization and infection (Bellais et al., 2012; Figueira-Coelho et al., 2004; Florindo et al., 2010 – chapter III; Fluegge et al., 2011; Ippolito et al., 2010; Lachenauer et al., 1999; Martins et al., 2007).

Previous reports from Portugal, for the period from 2002 to 2007 (Florindo et al., 2010 – chapter III; Martins et al., 2007), have shown a low and stable prevalence of serotype IV (6/269 (2%) and 3/100 (3%) among colonized women of reproductive age (15 to 49 years). The same scenario was observed among neonatal (2/64 cases; 3%) for the years 2000 to 2004 (Martins et al., 2007) and non-pregnant adult infections (2/225 cases; 1%) from 2001 to 2008 (Martins et al., 2012). After 2010, reports from Brazil, Ireland and the US (Diedrick et al., 2010; Ferrieri et al., 2013; Kiely et al., 2011; Palmeiro et al., 2010) revealed an increased prevalence of serotype IV in
colonisation and infection, suggesting the possibility that this serotype could be emerging as an important pathogen, as happened with serotype V during the 1990s (Elliot et al., 1998).

In this report we describe the annual serotype distribution and the antimicrobial susceptibility of *colonizing S. agalactiae* isolated in the Lisbon and Tagus Valley region in Portugal from 2005 to 2012, revealing the increasing frequency of serotype IV and a novel serotype IV clone defined by its clonal complex (CC) 17 hypervirulent lineage, recently identified in Taiwan, France and the US (Belais et al., 2012; Ferrieri et al., 2013; Tien et al., 2011).

### 4.2 Materials and Methods

#### 4.2.1 Group B Streptococcus collection

A total of 953 non-redundant *S. agalactiae* carriage isolates recovered from rectovaginal specimens of healthy women in reproductive age (668 pregnant) were included in this study. *S. agalactiae* were isolated according to the US Centers for Diseases Control and Prevention (CDC) guidelines (Schrag et al., 2002; Verani et al., 2010). The Portuguese National Institute of Health and six tertiary hospitals (Maternidade Alfredo da Costa, Hospital Garcia de Orta, Hospital Dona Estefânia, Hospital CUF Descobertas, Hospital Fernando Fonseca and Hospital Distrital de Santarém) located in the Lisbon and Tagus Valley region, Portugal, participated in this survey between January 2005 and December 2012. *S. agalactiae* isolates were identified to the species level by standard criteria based on colony morphology, Gram staining, catalase test, and commercial group B *Streptococcus* latex-agglutination assays.

#### 4.2.2 Capsular serotyping

All isolates were serotyped by slide agglutination using specific rabbit antisera against *S. agalactiae* polysaccharide antigens Ia, Ib, II to VIII (Essum AB) according to the instructions of the manufacturer. Non-serotypeable isolates were subjected to capsular (cps) genotyping, through the polymorphism analysis of cpsD-cpsE-cpsF region (Florindo et al., 2010 – chapter III). All serotype IV isolates were further confirmed through capsular genotyping. Non-typeable isolates after both serotyping and cps genotyping procedures were designated as NT.

#### 4.2.3 Antimicrobial susceptibility profile

All *S. agalactiae* isolates were tested for penicillin G, erythromycin, clindamycin and vancomycin susceptibility by Epsilometer (E)-test, in accordance to the Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2009), to determine the minimum inhibitory concentration (MIC). The constitutive and inducible macrolide-lincosamide-streptogramin resistance phenotypes (cMLSβ and iMLSβ, respectively) were identified by the double-disc
diffusion method, as well the macrolide-specific efflux resistance phenotype (M) (Florindo et al., 2010 – chapter III; CLSI, 2009). Macrolide resistance genes \textit{ermTR}, \textit{ermB} and \textit{mefA} were also investigated by polymerase chain reaction (PCR) (Florindo et al., 2010 – chapter III).

Considering that tetracycline is nowadays not recommended for the prophylaxis of \textit{S. agalactiae} neonatal infection (Schrag et al., 2002; Verani et al., 2010), this antibiotic was not tested by all laboratories involved in the present study; consequently, only a subset of 372/953 (39\%) \textit{S. agalactiae} isolates was tested for tetracycline by disc-diffusion in accordance to the CLSI guidelines (CLSI, 2009).

\textbf{4.2.4 Molecular analysis of serotype IV isolates}

In order to estimate the frequency of type IV isolates belonging to the sequence type (ST) 17 lineage, the presence of the \textit{lvgA} gene (encoding a surface adhesin characteristic of the hypervirulent \textit{S. agalactiae} CC17) was achieved by PCR, as described elsewhere (Lamy et al., 2006) Serotype IV \textit{lvgA}-positive isolates were further subjected to multilocus sequence typing (MLST) analysis (Jones et al., 2003), including the partial sequencing (about 500 bp) of seven housekeeping loci. Alleles of all loci were examined through the \textit{S. agalactiae} MLST database (http://pubmlst.org/sagalactiae/) providing an allelic profile or ST.

Serotype IV characterization also included the study of the Alp family, a major streptococcal antigen, by using multiplex PCR for direct identification of the \textit{alpha-C}, \textit{rib}, \textit{epsilon} and \textit{alp2-alp4} genes (Gherardhi et al., 2007). The prevalence of mobile genetic elements (MGEs), IS (insertion sequence) \textit{86I}, \textit{IS1381}, \textit{IS1548} and GBSi1 group II intron within the C5a peptidase gene (\textit{scpB}) – laminin-binding protein gene (\textit{hmb}) region within type IV/CC17 isolates were also evaluated by PCR, as previously described (AI Safadi et al., 2010).

\textbf{4.3 Results}

\textbf{4.3.1 Annual distribution and frequency of serotypes}

Among the 953 isolates analyzed, serotypes III, Ia, and V were the most frequent ones during the whole study period (2005–2012) (222 (23\%), 203 (21\%), and 192 (20\%), respectively), followed by serotypes II, IV, Ib and NT (135 (14\%), 89 (9\%), 72 (8\%) and 40 (4\%), respectively) (Figure 4.1). Serotypes VI to VIII were not found.
Variations in the distribution of *S. agalactiae* serotypes were observed, especially in 2011 and 2012, when the proportion of the serotypes III and V decreased whereas the proportion of serotypes IV and Ib increased. Indeed, a remarkable increase in serotype IV frequency has been observed, from 1% (2 of 148 isolates) in 2006 to 20% (19 of 97 isolates) in 2012 (20-fold), ranking this serotype as the second most detected in 2012 (Figure 4.1). In contrast, serotype II remained stable during the eight years study period, as its frequency ranged between 12% (N = 116) and 16% (N = 151).

![Figure 4.1 Serotype distribution among group B Streptococcus colonizing isolates (N = 953) from women of reproductive age, Lisbon and Tagus Valley regions, Portugal, 2005-2012. NT, non-typeable. (Adapted from Florindo et al., 2014b)](image)

### 4.3.2 Susceptibility to antimicrobials

Neither resistance nor reduced susceptibility to vancomycin or to penicillin G, a first-line antibiotic for the prophylaxis and treatment of *S. agalactiae* infections, were detected. For the total isolates in the 2005 to 2012 period, the percentage of *S. agalactiae* isolates that were resistant to erythromycin ranged from 14% (21/148) in 2006 to 23% (22/95) in 2011, whereas the percentage of *S. agalactiae* isolates with resistance to clindamycin ranged from 6% (7/120) in 2009 to 18% (17/97) in 2012 (Figure 4.2). Of note, the higher resistance rates for both antibiotics respectively were observed in the two last years of the study (2011 and 2012) (Figure 4.2).
Figure 4.2 Percentage of the different group B *Streptococcus* serotypes among erythromycin (A) (N = 162) and clindamycin (B) (N = 98) resistant isolates, Lisbon and Tagus Valley regions, Portugal, 2005-2012. NT, non-typeable. (Adapted from Florindo *et al.*, 2014b).

Among the 162/953 (17%) erythromycin-resistant isolates, 99/162 (61%) displayed the cMLS$_B$ phenotype, 56/162 (35%) had the iMLS$_B$, and the M phenotype accounted for 7/162 isolates (4%). All of the cMLS$_B$ and iMLS$_B$ resistance phenotypes were conferred by the presence of the *ermB* and *ermTR* genes, respectively, whereas the M phenotype was related to the presence of the *mefa* gene. Among the 372 *S. agalactiae* isolates tested for tetracycline, 306 (82%) were resistant to this antibiotic. Only 41/162 erythromycin-resistant *S. agalactiae* isolates were tested for tetracycline and all were resistant to the latter, which could be expected considering a putative horizontal gene transfer event involving the same conjugal transposon carrying both genetic resistance determinants (Gherardini *et al.*, 2007). We verified that the erythromycin (n = 162) and clindamycin (N = 98) resistant isolates involved multiple serotypes (Figure 4.2), despite the predominance of serotypes III and V from 2005 to 2008; however, the distribution profile remained very similar during the last three years (2010–2012), which could contradict the association between serotype III and macrolide resistance, previously demonstrated in Portugal and Spain (Florindo *et al.*, 2010 – chapter III; Martins *et al.*, 2011). This situation constitutes a new scenario involving other serotypes, namely Ib and IV (Figure 4.2). In fact, in 2007 none of the four isolated strains serotyped as Ib was resistant to macrolides, but during 2010 to 2012, 19/28 (68%) and 16/28 (57%) serotype Ib isolates were resistant to erythromycin and clindamycin, respectively; however, the number of Ib isolates was relatively low during this triennium. In 2006 and 2007, none of the 11 serotype IV isolates was resistant to erythromycin or clindamycin, whereas during 2010 to 2012,
9/43 (21%) and 6/43 (14%) serotype IV isolates were resistant to erythromycin and clindamycin, respectively.

### 4.3.3 Frequency of clonal complex 17 lineage in serotype IV isolates

Nine of 89 (10%) serotype IV isolates collected over the eight-year period belonged to the hypervirulent CC17 lineage, and all displayed ST291 (a single locus variant of ST17); these nine isolates were recovered in 2008 (N = 3), 2009 (N = 1), 2010 (N = 1) and 2012 (N = 4). Concerning their susceptibility to antimicrobials, with one exception (one isolate from 2012, which was co-resistant to clindamycin and erythromycin (MIC ≥ 256 µg/ml), the remaining eight isolates were fully susceptible to penicillin G, erythromycin, clindamycin, and vancomycin. All displayed the \textit{S. agalactiae} surface protein \textit{rib} gene, the GBSII in the \textit{scpB-lob} intergenic region and the IS861. Excluding one serotype IV isolate from 2010, the insertion sequence IS\textit{1381} was not detected.

### 4.4 Discussion

The \textit{S. agalactiae} capsule has long been recognized as one of the most important virulence factors. Variations of the polysaccharide structure allow the antigenic distinction of 10 different serotypes (Edwards & Niset, 2011; Slotved \textit{et al.}, 2007). It has been reported that predominating serotypes change over time, vary by geographical region and ethnic origin and can be associated with different diseases. The existence of several serotypes together with their differential distribution constitutes a major obstacle for the development of a global and effective \textit{S. agalactiae} vaccine to prevent \textit{S. agalactiae} neonatal infections (Johri \textit{et al.}, 2006).

Due to its low prevalence in European countries and in the US, serotype IV was not selected for the development of capsular polysaccharide-based vaccines (Johri \textit{et al.}, 2006; Rodriguez-Granger \textit{et al.}, 2012). This situation has changed in the last decade, when some countries, including the US, saw the emergence of serotype IV among colonizing and invasive \textit{S. agalactiae} isolates (Diedrick \textit{et al.}, 2010; Ferrieri \textit{et al.}, 2013). This scenario may become risky if the emergence of serotype IV combines with antibiotic resistance, which was the case in our study where co-resistance to second-line macrolide antibiotics was observed in recent years (2010–2012). Corroborating our findings, resistance to macrolides and clindamycin has been described in the US (Ferrieri \textit{et al.}, 2013) among invasive serotype IV isolates, predicting the emergence of serious problems for the intrapartum antibiotic prophylaxis in pregnant women allergic to penicillin. \textit{S. agalactiae} serotype distribution changes and antibiotic resistance trends constitute emerging phenomena that emphasize the need for constant monitoring, in order to develop accurate \textit{S. agalactiae} prevention strategies.
Another major concern is the association of serotype IV with the ST17 lineage identified in our study, supporting that previously described in a few other geographical regions, such as France, Taiwan and US (Bellaïs et al., 2012; Ferrieri et al., 2013; Tien et al., 2011). It is worth noting that ST17 lineage was long considered as a homogeneous epidemic clone, almost exclusively composed by serotype III isolates, and characterized by its rapid global dissemination and successful adaptation to human neonates (Sorensen et al., 2010). The origin of the novel association of CC17 with serotype IV can be due to an exchange of a 35.5 Kb DNA segment containing the entire capsule operon, culminating in a type III to type IV capsular switch, as described by Bellaïs et al. (Bellaïs et al., 2012). This phenomenon predicts an important epidemiological success for this new clone. As both French and Portuguese type IV/CC17 *S. agalactiae* isolates were recently identified (after 2008), and as they share the same ST291, we could speculate on a common ancestor; however, this hypothesis needs further evaluation as this ST was also described among serotype IV invasive isolates from Minnesota, US (Ferrieri et al., 2013).

In our study, the clonal origin hypothesis was evaluated through the screening of specific mobile genetic elements among our type IV/ST291 isolates, as their acquisition via recombination or horizontal transfer events are linked with the evolution and niche adaptation of bacterial species or particular clones. We verified that all type IV/ST291 isolates shared the same MGE profile composed by IS861 and *S. agalactiae* within the scpB-lmb intergenic region in the absence of IS1381. Only one variant carrying this latter IS has been identified in 2010. This MGE profile strongly correlates to the evolutionary scheme proposed by Héry-Arnaud et al. (Héry-Arnaud et al., 2005) for the ST17 lineage; however, the existence of type IV/ST291 variants, containing IS1381 or displaying antibiotic resistance, suggests differential evolutionary status from a common ancestor.

In conclusion, a novel epidemic *S. agalactiae* type IV/CC17 clone seems to be emerging through a putative clonal expansion among neonates and adults, as might have occurred since the 1960s with type III/ST17, an ‘epidemic clone’ with a rapid global dissemination and adaptation to human neonates (Sorensen et al., 2010).

The sudden increase of *S. agalactiae* serotype IV detection in different countries does not rely on the emergence of type IV/CC17 only, as other genetic lineages (such as CC1 and CC23) or different types of pulsed-field gel electrophoresis have been identified, constituting the majority of the serotype IV isolates (Diedrick et al., 2010; Palmeiro et al., 2010; Elliot et al., 1998; Tien et al., 2011). A careful surveillance of *S. agalactiae* type IV/ST291 emergence is recommended, in order to define its host specificity, tropism, virulence potential and antibiotic resistance phenotype.
4.5 Members of the Group for the Prevention of Neonatal GBS Infection

Emilia Prieto (Instituto de Higiene e Medicina Tropical, Lisbon, Portugal), Filomena Nogueira and José Diogo (Hospital Garcia de Orta, Almada, Portugal), Luisa Sancho, Maria João Brito and Rosalina Barroso (Hospital Fernando Fonseca, Amadora, Portugal), Isabel Peres, Rosa Barros and Teresa Neto (Hospital D. Estefânia, Lisbon, Portugal), Eulália Galhano and Isabel Silva (Maternidade Bissia Barreto, Coimbra; Portugal), Alexandra Almeida, Eugénia Fernandes and Paulo Sarmento (Maternidade Júlio Dinis, Porto, Portugal) and Isabel Lourenço (Maternidade Alfredo da Costa, Lisbon, Portugal).

4.6 Acknowledgements and Transparency Declaration

The authors would like to thank to Albertina Paulino (Instituto Nacional de Saúde Dr. Ricardo Jorge), Jorge Lima (Hospital Cuf Descobertas, Lisbon, Portugal), Isabel Nogueira (Maternidade Alfredo da Costa, Lisbon, Portugal), Paula Caetano (Hospital Dona Estefânia, Lisbon Portugal and Hospital Cuf Descobertas, Lisbon Portugal), Isilda Rocha (Hospital Fernando da Fonseca, Amadora, Portugal), and Patrícia Guimarães (Instituto Higiene e Medicina Tropica, Lisbon, Portugal). Funding: This work was supported by projects PTDC/SAU-MII/105114/2008, Pest-OE/BIA/UI0457/2011-CREM, both funded by Fundação para a Ciência e a Tecnologia/Ministério da Educação e Ciência (FCT/MEC), Portugal, PROC 60839, funded by Fundação Calouste Gulbenkian, Portugal, and Grant 123/2007 funded by Comissão de Fomento da Investigação em Cuidados de Saúde, Portugal. CF was supported by PhD grant SFRH/BD/48231/2008 (FCT/MEC). Conflict of interest: None declared.
Molecular epidemiology of group B streptococcal meningitis in children beyond the neonatal period from Angola

Florindo C, Gomes JP, Rato MG, Bernardino L, Spellerberg B, Santos-Sanches I, and Borrego MJ

Manuscript with minor changes published in 2011 in Journal of Medical Microbiology 60:1276-1280. With kind permission from Society for General Microbiology, UK

Author contributions
CF, JPG, MJB conceived the study; LB diagnosed S. agalactiae, provided the strains and clinical data; CF and MGR performed the PFGE analysis; CF performed additional molecular analyses of strains; CF, JPG, MGR, BS, ISS and MJB analyzed the data; BS, ISS and MJB supervised the study; CF wrote the manuscript; Manuscript revision and approval: all.
Abstract

*S. agalactiae* is a major pathogen of neonates and immunocompromised adults. Prior studies have demonstrated that, beyond the neonatal period, *S. agalactiae* rarely causes invasive infections in children. However, during 2004–2005, *S. agalactiae* was the causative agent of 60 meningitis episodes in children aged 3 months to 12 years from Angola. To identify and study the specific causative genetic lineages of *S. agalactiae* childhood meningitis, which lack characterization to date, we conducted an extensive molecular analysis of the recovered isolates (N = 21). This constitutes what we believe to be the first molecular study of the population structure of invasive *S. agalactiae* isolates from Africa. A low genetic diversity was observed among the isolates, where the majority belonged to clonal complex (CC) 17 presenting the capsular subtype III-2 (86% of cases) and marked by the genetic element GBSi1, which has previously been observed to be associated with neonatal hosts. The predominance of single-locus variants of sequence type (ST) 17 suggested the local diversification of this hypervirulent clone, which displayed novel alleles of the *fbsB* and *sip* virulence genes. The absence of the *scpB–imb* region in two *S. agalactiae* isolates with the Ia/ST23 genotype is more typical of cattle than human isolates. Globally, these data provide novel information about the enhanced invasiveness of the CC17 genetic lineage in older children and suggest the local diversification of this clone, which may be related to the future emergence of a novel epidemic clone in Angola.

Keywords: *Streptococcus agalactiae*, meningitis, Angola, CC17 lineage
5.1 Introduction

*Streptococcus agalactiae*, group B *Streptococcus* (GBS), is the leading cause of neonatal invasive infections in industrialized countries, and ten polysaccharide capsule types (serotypes) (Ia, Ib and II–IX) have been identified (Schrag et al., 2000; Slotved et al., 2007). *S. agalactiae* disease in newborns is classified as early-onset disease (EOD) or late-onset disease (LOD), depending on the age of the infant at the time of disease manifestation. EOD (< 7 days of age) represents the majority of cases and is associated with transmission from colonized mothers to the newborn through aspiration of infected amniotic fluid or passage through the birth canal, regularly manifesting as pneumonia and bacteremia (Liu & Nizet, 2004; Schrag et al., 2000; Trager et al., 1996). LOD (7 – 89 days of age) is characterized by bloodstream infection with a high incidence of meningeal involvement. The source of causative *S. agalactiae* for LOD is still not completely understood, but community or nosocomial acquisition as well as vertical transmission and prematurity may be implicated (Gagneur et al., 2009; Lin et al., 2003; Mullaney, 2001; Schrag et al., 2000). The vast majority of LOD episodes are caused by a homogeneous capsular type III genetic clone, defined by multilocus sequence typing (MLST) as sequence type (ST) 17 (Gherardi et al., 2007; Jones et al., 2003; Manning et al., 2009; Tazi et al., 2010). Dissimilarities in the pathogenic potential between carriage and invasive isolates have raised the question of whether the latter possess unique biological features that would favour crossing of the blood–brain barrier to cause meningitis. More recent studies have shown that the highly virulent clone ST17 presents an exclusive protein pattern, such as BibA, FbsB and CspA variants, which seem to be crucial for disease pathogenesis (Brochet et al., 2006; Springman et al., 2009; Tazi et al., 2010). The population structure and virulence traits of invasive *S. agalactiae* have been elucidated in recent studies from Europe (Gherardi et al., 2007; Luan et al., 2005; Jones et al., 2003, 2006) and North America (Bohnscak et al., 2008; Manning et al., 2009). In contrast, the few studies performed with African isolates have been restricted to capsular typing of invasive *S. agalactiae* (Gray et al., 2007; Madhi et al., 2003) or MLST data on *S. agalactiae* from maternal carriage (Brochet et al., 2009). *S. agalactiae* could represent a serious public health problem in Angola, as it constitutes a significant cause of bacterial meningitis and this country has the second highest mortality rate for under fives in the world (220 deaths per 1000 live births; WHO, 2010). Unfortunately, no *S. agalactiae* screening programs during pregnancy along with intrapartum antibiotic prophylaxis are available, which increases the risk of vertical transmission and, consequently, the probability of EOD or LOD. In the present study, we analyzed the phenotypic and genomic characteristics of *S. agalactiae* isolates responsible for meningitis in Angolan children beyond the neonatal period.

To our knowledge, this is the first study that combines several molecular methods for the characterization of invasive African *S. agalactiae* isolated from children belonging to an age group
for which this meningitis aetiological agent is uncommon (Kim, 2010; Sáez-Llorens & McCracken, 2003; Tzanakaki & Mastrantonio, 2007).

5.2 Methods

5.2.1 Study population and bacterial isolates

The Paediatric Hospital of Luanda is a reference hospital in Angola, and contains the only laboratory in the whole country with the skills to diagnose bacterial meningitis. This laboratory was established in 2002 in collaboration with the Portuguese National Institute of Health in response to the Angolan bacterial meningitis endemic situation (Bernardino et al., 2003). Patients attending this hospital belong to a low socioeconomic group and come from all 18 provinces of Angola, either independently or transferred from other hospitals. We analyzed 21 S. agalactiae isolates responsible for meningitis in children aged 91 days to 12 years from a total of 60 cases of S. agalactiae meningitis diagnosed at the Paediatric Hospital of Luanda during the years 2004 (N = 33) and 2005 (N = 27). Due to hospital constraints, namely regarding its ability for long-term storage of biological material, only 21 of the 60 S. agalactiae isolates were kept at -80°C, and only those 21 were sent to the Portuguese National Institute of Health for further characterization.

5.2.2 S. agalactiae identification and antimicrobial susceptibility profile

S. agalactiae isolates were obtained from cerebrospinal fluid cultures and confirmed at the species level, as described previously (Florindo et al., 2010 – chapter III; Pelkonen et al., 2009). Antimicrobial susceptibility testing (penicillin G, erythromycin, clindamycin and vancomycin) was executed by E-test according to Clinical and Laboratory Standards Institute guidelines (CLSI, 2009), and the presence of macrolide resistance-associated genes (ermTR, ermB and mefA) was analysed by PCR amplification, as described elsewhere (Gygax et al., 2006; Sutcliffe et al., 1996).

5.2.3 Capsular genotyping, PFGE and MLST

Capsular genotyping was carried out by PCR and DNA sequencing of the cpsD-cpsE-cpsF region, as documented previously (Florindo et al., 2010 – chapter III). Genomic DNA was digested with Smal and the fragments were resolved by PFGE as described elsewhere (Rato et al., 2008). Cluster analysis was performed using BioNumerics software (Applied Maths) to create UPGMA dendrograms. The Dice similarity coefficient of the Smal restriction PFGE profiles was used with optimization and position tolerance settings of 0 and 1%, respectively. Distinct PFGE types were assigned based on a similarity coefficient of < 80% (Rato et al., 2008). Clones were defined as clusters of isolates (three or more) when they presented a dendrogram profile similarity of ≥ 80%. For the MLST method (Jones et al., 2003), PCR fragments (~ 500 bp) of seven housekeeping loci were amplified and sequenced. Alleles of all loci were examined on an MLST database
(http://pubmlst.org/sagalactiae/) and the combination provided an allelic profile or ST. Clonal complexes (CCs) comprising isolates sharing six or seven identical alleles were defined.

5.2.4 Alpha-like protein (Alp) family

The molecular characterization included the study of a major antigen, the Alp gene family, which was analysed by multiplex PCR for direct identification of the alpha-C, rib, epsilon and alp2–alp4 genes (Gherardi et al., 2007).

5.2.5 Detection of mobile genetic elements (MGEs)

The presence of two MGEs, IS1548 and GBSi1, within the scpB–lmb intergenic region was evaluated by PCR, as described previously (Al Safadi et al., 2010). In the absence of MGEs, the presence of the flanking genes (scpB and lmb) was verified.

5.2.6 Allelic variation in bibA (gbs2018), fbsB and sip

The genetic polymorphisms of three virulence genes, bibA (encoding a surface protein), fbsB (encoding the fibrinogen-binding protein B) and sip (encoding a surface immunogenic protein), were investigated by PCR and DNA sequencing (Brochet et al., 2006; Springman et al., 2009).

5.3 Results and Discussion

5.3.1 Antibiotic susceptibility profiles

Antibiotic susceptibility testing revealed that all isolates were fully susceptible to penicillin G, as revealed by their MIC values (range 0.047–0.064 mg/ml), which indicated that the empiric antibiotic therapy [dose regimen: penicillin G (100 000 U/kg i.v. every 6 h) plus chloramphenicol (25 mg/kg i.v. every 6 h)] that is applied at the Paediatric Hospital of Luanda whenever there is a suspicion of bacterial meningitis was effective against S. agalactiae. Moreover, no resistance was detected for vancomycin (MIC range 0.38 – 1 mg/ml), clindamycin (MIC range 0.19 – 0.25 mg/ml) or erythromycin (MIC = 0.25 mg/ml), with the exception of a single isolate presenting intermediate erythromycin resistance (MIC = 0.5 mg/ml); however, none of the most common antimicrobial resistance genes were detected in this isolate. The low frequency or absence of macrolide resistance in invasive S. agalactiae has also been observed by other authors (de Azavedo et al., 2001; Gherardi et al., 2007; Zhao et al., 2008), which could suggest that invasive isolates are less likely to carry resistance determinants. We speculate that invasive isolates, mostly confined to sterile anatomical sites, have less contact with commensal or pathogenic microbiota, and thus are less prone to horizontal genetic transfer phenomena.
5.3.2 Population structure of the *S. agalactiae* isolates

Capsular genotyping of the 21 *S. agalactiae* isolates revealed two *cps* genotypes, Ia and III-2, where the latter was predominant (86%) and carried the *rib* gene (Figure 5.1). As no data are available on the *S. agalactiae* colonization rate and genotype distribution in Angola, we cannot draw conclusions about the predominance of these two capsular clones among invasive *S. agalactiae* isolates responsible for meningitis, as observed by others (Bohnscak et al., 2008; Gherardi et al., 2007; Gray et al., 2007; Luan et al., 2005; Manning et al., 2009; Tazi et al., 2010; Zhao et al., 2008).

The discriminatory power of the PFGE method was higher than that of the capsular typing and allowed the identification of 13 different DNA band profiles corresponding to six PFGE types (named A–F) distributed into two major clonal clusters (I and II), three singletons (D–F) and a group of two isolates belonging to PFGE type C (Figure 5.1). Isolates from cluster II with the same PFGE profile (sharing 100% similarity) belonged to different STs (ST17 or ST109). Seven STs were identified among the 21 isolates using MLST. Three isolates, all exhibiting capsular genotype Ia, were ST23, and the remaining 18 isolates displayed the capsular genotype III-2 and were ST17 or single-locus variants (SLVs) of this ST. Within these SLVs, ST450 and ST451 corresponded to novel STs, described for the first time in this study to our knowledge. The absence of other clones with the ability to cause meningitis, such as III/CC19 and V/CC1 as reported by others (Gherardi et al., 2007; Jones et al., 2003; Manning et al., 2009), may reflect local genotype distribution characteristics and/or the limited number of isolates available in the current study.

![Figure 5.1](image.png) Genetic characteristics of the 21 invasive GBS isolates from Angola. The dendrogram was constructed through the BioNumerics software using the UPGMA method. The genetic similarity between isolates is shown on the horizontal scale. PFGE types were defined on the basis of a threshold of 80% similarity (Rato et al., 2008). The simultaneous absence of both MGEs is denoted ΔMGE. (Adapted from Florindo et al., 2011).
The predominance of the CC17 lineage supports the previously reported association between CC17 and neonatal infections (Gherardi et al., 2007; Jones et al., 2003; Manning et al., 2009; Tazi et al., 2010), although the children enrolled in those studies belonged to a different age group (up to 3 months of age). In contrast to other studies (Gherardi et al., 2007; Manning et al., 2009; Tazi et al., 2010), analysis of the CC17 lineage showed an atypical distribution of STs within this lineage, where the majority (77.8%) of the CC17 isolates were SLVs of ST17, suggesting a local diversification of this clone. Moreover, the identification of PFGE and MLST genetic variants among the CC17 isolates corroborated previous studies describing the relative homogeneity of this genetic lineage (Brochet et al., 2006; Gherardi et al., 2007; Rolland et al., 1999; Springman et al., 2009). This limited diversity indicates that CC17 has emerged recently from the core population, reflecting a distinct genome architecture with putative implications in host tropism and virulence (Manning et al., 2009; Sorensen et al., 2010; Tazi et al., 2010), where the presence of MGEs may be relevant, as demonstrated by the up-regulation of the \( \text{lm}b \) gene by IS\( \text{IS}1548 \) (Al Safadi et al., 2010).

### 5.3.3 Genomic organization of the \( \text{scpB-lmb} \) region

In line with the results presented above, we screened for the presence of two MGEs, IS\( \text{IS}1548 \) and GBSi1, situated between the \( \text{scpB} \) and \( \text{lm}b \) genes, and studied the genetic polymorphism of three virulence-associated genes. All isolates belonging to genotype III-2/CC17 carried GBSi1 within the \( \text{scpB-lmb} \) intergenic region (Figure 5.1), which is considered a marker of the CC17 genetic lineage (Al Safadi et al., 2010; Luan et al., 2005; Zhao et al., 2008). The absence of GBSi1 or IS\( \text{IS}1548 \) was observed in one of the three In/ST23 isolates, whereas the other two isolates lacked the \( \text{scpB} \) and \( \text{lm}b \) genes, suggesting that they may have originated either directly or indirectly from cattle, as these genes are usually absent in bovine isolates (Al Safadi et al., 2010; Brochet et al., 2006; Franken et al., 2001). The possibility of other sources for \( \text{S. agalactiae} \) acquisition, namely from the community or cattle (Manning et al., 2010), was further supported by the fact that none of the 21 invasive isolates were recovered from newborns, which contradicted the usual \( \text{S. agalactiae} \) pathogenesis. Nevertheless, data from a previous study in Angola reported that a relevant number of children attending the Paediatric Hospital of Luanda with signs of meningitis died without a laboratory diagnosis (123/717 in 2004) (Pelkonen et al., 2009), suggesting that \( \text{S. agalactiae} \) vertical transmission is probably underestimated in Angola. In addition, the lack of clinical data precluded the establishment of any association between \( \text{S. agalactiae} \) meningitis in older children and the presence of predisposing conditions for \( \text{S. agalactiae} \) infection (such as human immunodeficiency virus infection, malaria or severe malnutrition), which was verified in 72.7% of South African children infected with \( \text{S. agalactiae} \) after the neonatal period (Madhi et al., 2003).
5.3.4 Allelic variation in \textit{bib}A, \textit{fbs}B and \textit{sip}

The relationship between the allelic variation of virulence-associated genes and MLST genetic lineages (Figure 5.1) partially contrasted with the literature data (Brochet et al., 2006; Springman et al., 2009). Indeed, our findings regarding the \textit{sip} and \textit{fbs}B genes revealed: (i) a \textit{sip}\textsubscript{3a} allele, described here for what we believe to be the first time for the CC17 lineage; (ii) a novel minor variant of \textit{sip}2 found only in ST287 isolates (\textit{sip}2\textsubscript{1}; GenBank accession no. HQ267706); (iii) the \textit{sip}2 allele in one ST23 (CC23) isolate (previously considered to be exclusive to CC17 isolates); and (iv) a novel allelic variant of \textit{fbs}B\textsubscript{2b} shared by all CC17 isolates (\textit{fbs}B\textsubscript{2b,1}; GenBank accession no. HQ267707). CC17 and CC23 Angolan isolates presented particular genetic signatures involving ST, \textit{cps} genotype, MGEs and surface protein genes, CC17/III-2/GBS\textsubscript{1}/rib/gbs2018-3 and ST23/1a/\textit{AMGE}/epsilon/gbs2018-1, which was in accordance with studies carried out in other countries (Al Safadi et al., 2010; Brochet et al., 2006; Gherardi et al., 2007; Springman et al., 2009). In addition, the presence of the \textit{scp}B and \textit{imb} genes in 19 out of 21 isolates highlights the hypothesis that the \textit{scp}B–\textit{imb} region may be related to colonization or other mechanisms of human GBS infection (Al Safadi et al., 2010). In contrast, the detection of \textit{sip}2 and \textit{sip}\textsubscript{3a} alleles in genotypes 1a/ST23 and III-2/ST174, respectively (Figure 5.1), indicates the occurrence of recombination events among distant lineages. These findings suggest that the putative existence of an exclusive set of surface proteins in CC17 isolates as epidemiological markers of this highly virulent lineage (Brochet et al., 2006; Springman et al., 2009) should be viewed with caution.

In conclusion, the predominance of CC17 causing episodes of meningitis in older children from Angola could suggest an adaptation of this lineage to childhood infection, as it rarely causes bacteremia or meningitis in the adult population (Jones et al., 2003, 2006; Luan et al., 2005; Tazi et al., 2010); however, vertical transmission and some clinical predisposing conditions cannot be excluded. Thus, further epidemiological studies are required to elucidate the course of \textit{S. agalactiae} infection in neonatal and post-neonatal cases of meningitis, as well as the putative cattle origin of \textit{S. agalactiae}, as suggested from our data. Finally, the use of both colonizing and invasive circulating clones in further studies is mandatory, as they may contain specific implications for the design of a universal \textit{S. agalactiae} vaccine.
5.4 Acknowledgements and Transparency Declaration

This work was partly supported by Fundação para a Ciência e Tecnologia/MEC grant PTDC/SAU-MII/105114/2008, Portugal, and PhD grants SFRH/BD/48231/2008 and SFRH/BD/32513/2006 awarded to C.F. and M.G.R., respectively. We thank the MLST database curator, Kate Dingle, for allocating new STs in the *S. agalactiae* MLST database.
Selection of reference genes for real-time expression studies in *Streptococcus agalactiae*

**Florindo C, Ferreira R, Borges V, Spellerberg B, Gomes JP, and Borrego MJ**


*Author contributions*
CF and JPG designed the study; CF and RF performed the experiments; CF and VB performed the bioinformatic analysis; Data analysis: all; JPG and MJB supervised the study; CF wrote the manuscript; CF, JPG and MJB revised the manuscript; Approval: all.
Abstract

*Streptococcus agalactiae*, group B streptococci (GBS) is the leading cause of severe bacterial infections in newborns. *S. agalactiae* expression studies allowed the identification and characterization of virulence factors and a better understanding of the host–pathogen–environment interactions. The measurement of transcript levels by quantitative real-time PCR (qRT-PCR) is a widely used technique in *S. agalactiae*; however, a systematic evaluation and validation of reference gene stability for normalization purposes in *S. agalactiae* expression studies is currently lacking. Therefore, we analyzed the stability of 10 candidate reference genes (*16SrRNA, glcK, glnA, groEL, gyrA, recA, rpoB, rpsL, sdhA* and *trr*) in three *S. agalactiae* prototype strains (O90R, NEM316 and 2603V/R) grown at different temperature conditions (37°C and 40°C). Our approach was based on the calibration of transcript levels from each gene against the number of bacteria from the same sample (ratio messenger RNA/genomic DNA). As a complementary analysis, reference gene stability was also investigated through the bioinformatic applications, geNorm and NormFinder. Considering the whole *S. agalactiae* development cycle, only a minority of genes were stable under both growth conditions, but this number increased when restricting the analysis to the logarithmic time-points. The range of stable genes was higher at 37°C, where *recA* and *sdhA* were stable simultaneously for the three strains, and six out of 10 genes were stable for at least two strains. At 40°C, *recA* showed up again as one of the best options, suggesting its potential use as reference gene in future qRT-PCR studies. The results generated with geNorm and NormFinder were consistent with those obtained experimentally and evidenced minor variations either among strains or temperature conditions. In conclusion, the fluctuation of expression of reference genes observed among different *S. agalactiae* strains and growth conditions highlights the importance of carefully validating, for each experimental scenario, the use of reference genes for qRT-PCR normalization purposes. Nevertheless, *recA* seems to be a good candidate for such optimizations.

Keywords: *Streptococcus agalactiae*, gene expression, normalization, reference genes
6.1 Introduction

*Streptococcus agalactiae*, group B streptococci (GBS), is a leading cause of bacterial sepsis and meningitis in neonates from industrialized countries (Edmond et al., 2012; Schrag et al., 2000), and an emerging pathogen in nonpregnant adults (Phares et al., 2008; Skoff et al., 2009; Tazi et al., 2011). For the understanding of dissimilarities between carriage and infection, the evaluation of the gene expression in *S. agalactiae* is crucial. Most of these studies have been performed mostly by quantitative real-time PCR (qRT-PCR) (Al Safadi et al., 2010; Gleich-Theurer et al., 2009; Lembo et al., 2010; Quach et al., 2009; Rozhdestvenskaya et al., 2010; Tazi et al., 2010) and whole-genome microarray analysis (Bryan et al., 2008; Jolri et al., 2007; Mereghetti et al., 2008; Sitkiewicz et al., 2009), demonstrating extensive transcriptome remodeling at the various stages of growth and in different biological scenarios. Accurate quantification of these transcriptomic changes requires the use of a proper control to normalize gene expression data, in order to remove or minimize the experimental variables, such as differences in the amount of starting material, RNA extraction yield, RNA quality or PCR efficiencies (Bustin, 2002; Huggett et al., 2005; Nolan et al., 2006). Housekeeping genes (HKGs) are frequently used for qRT-PCR normalization (Bustin, 2002; Huggett et al., 2005; Thellin et al., 1999; Vandecastele et al., 2001), where mRNA of the target genes under study is normalized against the co-extracted mRNA encoded by HKGs.

As a prerequisite, the expression of HKGs is often considered constant with low levels of fluctuation among most experimental conditions (Bustin, 2002; Huggett et al., 2005; Thellin et al., 1999). However, many studies showed that expression of HKGs in both eukaryotes (Cicimnati et al., 2008; Dheda et al., 2004; Huggett et al., 2005; Thellin et al., 1999) and prokaryotes (Borges et al., 2010; Metcalf et al., 2010; Vandecastele et al., 2001) can vary with experimental conditions, in part because these genes may not be strictly involved in the basal cell metabolism (Chuang and Ishitani, 1996). Thus, reference genes need to be properly validated for specific species, biological samples, and growth conditions in order to prevent inaccurate data interpretation, and subsequent biased expression profiles (Bustin, 2002; Dheda et al., 2004). To our knowledge, no comprehensive evaluation has been performed so far concerning the validation of reference genes for expression studies in *S. agalactiae*. Therefore, the present work aims to evaluate the stability of ten candidate reference genes in three *S. agalactiae* prototype strains throughout the bacterial development cycle under different growth conditions.
6.2 Methods

6.2.1 Bacterial strains and growth conditions

Three *S. agalactiae* prototype strains belonging to distinct genetic lineages were used in this study: O90R (genotype Ia/ST25), NEM316 (genotype III/ST23) and 2603V/R (genotype V/ST110). Bacteria were grown in Todd Hewitt broth supplemented with 0.5% yeast extract (THY) in 5% CO₂ at 37°C overnight as standing cultures. Dilutions of 1:50 of these cultures were used to inoculate triplicate cultures of fresh THY broth (50 ml) that were allowed to incubate without shaking at 37°C and 40°C with 5% CO₂. Cell growth was monitored by measuring the optical density at 600 nm (OD₆₀₀) and by viable cell counting. These temperatures were chosen in order to reproduce normal in vitro and in vivo growth conditions (37°C) but also severe human *S. agalactiae* infections, during which inner body temperature can reach 40°C ("fever" conditions) (Freitas Lione et al., 2010). Growth curves of O90R, NEM316, and 2603V/R with identifiable lag, logarithmic and stationary phases were obtained both by OD₆₀₀ reading and qRT-PCR to determine the number of bacterial genomes (Figure 6.1). The triplicate assays of growth curves for each strain were highly reproducible, allowing the accurate establishment of the growth phases.
Figure 6.1 Growth curves of the O90R, NEM316 and 2603V/R wild type strains at 37°C (A) and 40°C (B). Bacteria were inoculated into THY media from a fresh overnight culture and growth was monitored by measuring the OD₆₀₀ for at least 330 min (5.5 h) and by qRT-PCR in five time-points [Lag (OD₆₀₀ ≥ 0.1), early-log (EL, OD₆₀₀ = 0.2), middle-log (ML OD₆₀₀ = 0.5), late-log (LL, OD₆₀₀ = 0.8) and early stationary (ES, OD₆₀₀ = 0.9) phases]. Depicted are mean values and standard deviations of three independent experiments. (Adapted from Florindo et al., 2012).

6.2.2 Nucleic acid isolation

The general strategy of this study consisted on determining the ratio between the amount of mRNA from each gene (numerator) and the number of genomes (denominator). Thus, for each strain, under different experimental conditions, bacterial cells were collected for DNA and RNA extraction at five time-points of growth: lag phase (OD₆₀₀ ≥ 0.1), early, middle and late exponential phases (OD₆₀₀ = 0.2, 0.5 and 0.8, respectively) and early stationary phase (OD₆₀₀ = 0.9). At each time-point, 1 ml of each bacterial culture was collected, homogenized and vigorously divided into two identical aliquots, one of which was immediately stored at −20 °C for further DNA extraction whereas the other was immediately subjected to RNA isolation. The latter comprehended a first step of RNA stabilization [500 µl bacterial suspension plus 1 ml RNA protect bacteria reagent (Qiagen)], followed by harvesting of S. agalactiae cells, and pellet digestion [150 µl of Tris-EDTA buffer, pH 8.0, containing 50 U mutanolysin (Sigma-Aldrich) and 15 mg/ml lysozyme (Sigma-Aldrich), 37°C, 30 min]. Subsequently, total RNA was extracted using the RNeasy mini kit.
(Qiagen) according to manufacturer’s instructions. Residual contaminant DNA was removed by using 30 U RNase-free DNase (Qiagen), and RNA elution was done in 40 μl of RNase-free water.

Extracted RNA (ranging from 20 to 560 ng/μl, depending on the growth phase) was finally stored at −80 °C until use. The absence of genomic DNA contamination was further checked by PCR (using the same primer pairs used for qRT-PCR). Genomic DNA was obtained using the QIAamp DNA mini kit (Qiagen), according to a modified protocol for gram-positive bacteria (Cohen-Poradosu et al., 2004). Briefly, S. agalactiae cells were lysed with 30 U of mutanolysin (Sigma-Aldrich) and 20 mg/ml of lysozyme (Sigma-Aldrich) for 2 h at 37°C, before treatment with proteinase K (10 mg/ml) and buffer AL for 30 min at 56 °C. Eluted DNA (ranging from 13.3 to 95 ng/μl) was then stored at −80 °C.

6.2.3 Reproducibility evaluation of nucleic acid isolation

The reproducibility of RNA and DNA extraction procedures was evaluated in order to avoid biased results on qRT-PCR assays. Thus, during the growth of a S. agalactiae strain, we collected 5 ml of culture and divided it into five aliquots of 1 ml, which were subsequently divided into two aliquots of 500 μl for independent DNA and RNA purifications (techniques described in 6.2.2). This procedure was done for both lag and early stationary growth phases in order to evaluate the extraction reproducibility under low and high amounts of nucleic acids. The concentration of DNA and RNA was assessed at OD260. The reproducibility of both methods was statistically evaluated by calculating the coefficient of variation within each group of five samples.

6.2.4 Generation of standard curves

In order to quantify the number of S. agalactiae genomes in each time-point, a plasmid standard curve was generated as previously described (Gomes et al., 2006). Briefly, an amplicon of the single copy sdhA gene of S. agalactiae was cloned into the TOPO vector using the TOPO TA technology for PCR products, which was used to transform competent E. coli DH5α strain (Invitrogen). Plasmid DNA was isolated and purified using the QIAprep Spin Miniprep kit protocol (Qiagen), according to the package protocol. RNA contamination was avoided by adding RNase A (20 mg/ml). Confirmation of cloning success was performed by EcoRI digestion and sequencing of the cloned fragment. The plasmid copy number was determined at OD260, according to the formula: No. Plasmid/μl = [Avogadro No. × Plasmid conc. (g/μl)]/MW of 1 mol of plasmids (g). A standard curve was generated by using eight 10-fold serial dilutions (representing 10 to 1×10^8 plasmid copies/μl). As S. agalactiae chromosome contains a single copy of the sdhA gene, the number of sdhA copies determined for each blind sample (by using the standard curves), correspond to the number of existing bacteria in those samples.
6.2.5 Reverse transcriptase and qRT-PCR

cDNA was generated from 1 μl (from 20 to 560 ng) of each RNA sample collected at each time-point, by using TaqMan RT reagents (Applied Biosystems). The reaction mixture (50 μl) consisted of 2.5 μM of random hexamers, 5.5 mM MgCl₂, 500 μM of each dNTP, 1× RT Buffer, 0.8 U/μl RNase inhibitor and 1.25 U/μl MultiScribe RT and were performed under the following cycling conditions: 10 min at 25°C, 15 min at 42°C and 5 min at 99°C. To minimize the influence of PCR inhibitors in real-time PCR (personal observation), all cDNA samples were diluted by a factor of 30 and stored in DNase-free microtubes at −80°C.

The pool of genes selected for this study included three sets: i) groEL (60KDa chaperonin), gyrA (DNA gyrase A), recA (recombinase A), rpoB (RNA polymerase beta unit) and rpsL (30S ribosomal protein S12), which were previously used to normalize data in qRT-PCR studies in *S. agalactiae* (Al Safadi et al., 2010; Gleich-Theurer et al., 2009; Lombo et al., 2010; Quach et al., 2009; Rozhdestvenskaya et al., 2010; Santi et al., 2009; Tazi et al., 2010); ii) gcleK (glucose kinase), gltA (glutamine synthetase), sdhA (succinate dehydrogenase) and ikf (transketolase), which seemed to present low expression variation in previous *S. agalactiae* microarrays studies (Mereghetti et al., 2008; Santi et al., 2009; Sikiewicz et al., 2009); and iii) 16S rRNA (16S ribosomal RNA) due to its frequent use as internal control in bacterial expression studies (Cope et al., 2011; Gomes et al., 2005; Jorge et al., 2011; Nunes et al., 2007; Shin et al., 2006).

For each gene, primers were designed using Primer Express (Applied Biosystems) (Annex) based on constant regions determined through comparison of sequences available in GenBank. The qRT-PCR was performed by using ABI 7000 SDS, SYBR Green chemistry and optical plates (Applied Biosystems). The qRT-PCR reagents consisted of 1× SYBR Green PCR Master Mix (Applied Biosystems), 400 nM of each primer and 5 μl of sample DNA (from 66.3 to 475 ng in the 5 μl) or cDNA, in a final volume of 25 μl. All samples were run in duplicate and ‘no template controls’ (NTC) and ‘no-RT’ controls were included in all runs to exclude potential DNA contamination. For each *S. agalactiae* strain, plates included a plasmid standard curve and duplicates of DNA extracted at each one of the five time-points (for absolute quantification of bacterial genomes), together with DNA standard curves and duplicates of cDNA obtained at each one of the five time-points (for quantification of transcripts). The use of DNA standard curves for quantification of transcripts allows a cross-comparison between expression data from different genes (which is not possible by using cDNA standard curves) (Gomes et al., 2005). Thermocycling amplification consisted of an initial denaturation at 95°C for 10 min followed by 40 cycles of 95°C/15 s and 60°C/1 min. The gene expression was determined from the respective standard curves by conversion of the mean threshold cycle (Ct) values. The specificity of the PCR amplicons was verified by melting curve analysis.
Finally, raw qRT-PCR data was normalized against the number of \textit{S. agalactiae} genomes determined for the corresponding sample. The final expression results were based on three independent experiments for prototype strains O90R, NEM316 and 2603V/R.

6.2.6 \textit{geNorm} and \textit{NormFinder} analysis

For the analysis of the reference gene expression stability, two well recognized statistical applications, \textit{geNorm} version 3.5 (Vandesompele \textit{et al.}, 2002) and \textit{NormFinder} version 0.953 (Andersen \textit{et al.}, 2004) were applied. \textit{geNorm} calculates the mean pairwise variation (M value) of a particular gene compared to that of all other genes under study. Subsequently, the genes are ranked and the lowest M value stands the highest expression stability (Vandesompele \textit{et al.}, 2002). \textit{NormFinder} selects the genes with the minimum expression variation throughout the sample. Each gene is ranked with a stability value based on the intragroup and/or intergroup variance. Genes with lower values have higher expression stability (Andersen \textit{et al.}, 2004).

6.3 Results

6.3.1 Reliability of nucleic acid isolation and qRT-PCR data

The repeated RNA and DNA extraction of twin samples demonstrated high reproducibility with a mean coefficient of variation (CV) of 5.3\% and 7.5\% for RNA and DNA, respectively. The high qRT-PCR efficiency of each set of primers (> 99\%) and the correlation coefficients obtained for all standard curves (\(R^2 > 0.99\)) ensured the reliability of our expression data. Moreover, we obtained Ct values from 9.9 to 12.2 for 16S\textit{rRNA} and from 19.6 to 25.2 for the remaining nine genes, which indicated expression levels in an appropriate range. Amplification curves with Ct values > 35 were sometimes obtained for no-RT controls, indicating a residual contamination with DNA. This contamination is neglected as the disparity of > 10 Ct values between samples and no-RT controls corresponds to > 1000-fold difference in the amount of nucleic acids. The specificity of this methodology was confirmed through melting curve analysis in all PCR runs, where a single peak for each amplicon was observed. Moreover, no amplicons were detected for NTC controls.

6.3.2 Overview of gene expression at 37\(^\circ\)C and 40\(^\circ\)C

The growth of \textit{S. agalactiae} strains at 40\(^\circ\)C resulted in a down-regulation of nine out of ten genes (from 1.3 to 2.9-fold decrease of the mean expression values) when compared to the growth at 37\(^\circ\)C. The only exception was the heat-shock protein encoding gene (\textit{groEL}), which presented no variation. For both temperatures, 16S\textit{rRNA} was the most expressed gene in all \textit{S. agalactiae} strains, presenting mean expression values up to 300-fold higher than the remaining nine genes
(data not shown). These high expression values may result from the presence of seven copies in the chromosome of \textit{S. agalactiae}.

### 6.3.3 Validation of reference genes with genomic DNA

We calibrated the mRNA of each candidate gene against the number of bacteria. Thus, for each time-point and for each gene, we determined the absolute fold-difference to the mean expression value of all five time-points, at 37°C or 40°C. We considered a cut-off value of $\leq$ 2-fold to define a stable gene. Only a minority of genes showed stability under this criterion (Figure 6.2). At 37°C, it was the case of \textit{tkr} for O90R and NEM316; at 40°C this was observed for \textit{16S rRNA}, \textit{sdhA} and \textit{tkr} for O90R and \textit{16S rRNA} and \textit{glnA} for 2603V/R. This scenario illustrates the difficulty of selecting a normalizing gene for studies enrolling the whole \textit{S. agalactiae} development cycle. This becomes exacerbated when multiple strains are used. However, as most studies focus on the expression levels of target genes during exponential growth, and as our results evidenced that gene instability was often related to the early stationary point (Figure 6.2), we re-analyzed the data by using solely the three logarithmic time-points. Remarkably, the panel of stable genes increased when restricting the analysis to this growth phase, broadening the range of options per strain and per experimental condition (Figure 6.3). Indeed, at 37°C six out of 10 genes showed stability for at least two strains (where \textit{recA} and \textit{sdhA} were stable for the three strains), and at 40°C, at least two genes were stable for each strain (where, again, \textit{recA} shows up as one of the best options).
Figure 6.2 Expression stability of 10 candidate reference genes for 090R, NEM316 and 2603 V/R strains during the five time-points of growth. Each graph represents the mean absolute fold difference of transcript levels of each time-point relative to the mean expression of all five time-points at 37°C (A) and 40°C (B). The transcript levels of each gene were normalized against the number of bacterial genomes (gDNA). Error bars represent SD. Vertical black dotted lines indicate a threshold of 2-fold expression difference. Stable genes (≤ 2-fold difference) are highlighted in gray. (Adapted from Florindo et al., 2012).
Figure 6.3 Expression stability of 10 candidate reference genes for 090R, NEM316 and 2603V/R S. agalactiae strains during the logarithmic time-points (early-, middle- and late-log). Each graph represents the mean absolute fold-difference of transcript levels of each time-point relative to the mean expression of the three logarithmic time-points at 37°C (A) and 40°C (B). The transcript levels of each gene were normalized against the number of bacterial genomes (gDNA). Error bars represent SD. Vertical black dotted lines indicate a threshold of 2-fold expression difference. Stable genes (≤ 2-fold difference) are highlighted in gray. Horizontal dashed gray boxes (A) indicate stable genes simultaneously for the three strains. (Adapted from Florindo et al., 2012).
6.3.4 Bioinformatic validation of reference genes

The stability of gene expression over the three logarithmic time-points was also evaluated with statistical algorithms. Tables 6.1 and 6.2 present an overview of the five top-ranked genes by using these two software applications. Globally, inter- and intra-strain stability rankings generated by geNorm were similar to the determined using NormFinder despite slight variations in ranking position. In fact, at both 37°C and 40°C, recA, gyrA, and glicK showed up among the top-five ranked genes for both software’s for the three strains. When bioinformatic results were compared with the ones obtained experimentally, a higher agreement was achieved within each strain than between strains. For example, all top-five ranked genes for strain NEM316 at 37°C are among the most stable genes detected through our experimental strategy.

Table 6.1 Top-five ranked genes of *S. agalactiae* based on geNorm and NormFinder bioinformatic tools. Genes of *S. agalactiae* prototype strains (O90R, NEM316, and 2603/V/R) were evaluated throughout three logarithmic time-points of growth at 37°C. The “M value” (geNorm) and “Stability value” (NormFinder) are inversely correlated to the stability of the candidate genes. (Adapted from Florindo et al., 2012).

<table>
<thead>
<tr>
<th>O90R</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
<th>NEM316</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
<th>2603/V/R</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sFD</td>
<td>(0.220)</td>
<td>sFD</td>
<td>gyrA</td>
<td>(0.068)</td>
<td>gyrA</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>recA</td>
<td>(0.320)</td>
<td>recA</td>
<td>gyrA</td>
<td>(0.088)</td>
<td>recA</td>
<td>gyrA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>HsRN</td>
<td>(0.200)</td>
<td>gyrA</td>
<td>recA</td>
<td>(0.088)</td>
<td>gyrA</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>glkK</td>
<td>(0.253)</td>
<td>glkK</td>
<td>recA</td>
<td>(0.175)</td>
<td>glicK</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>gyrA</td>
<td>(0.497)</td>
<td>glicK</td>
<td>recA</td>
<td>(0.200)</td>
<td>glicK</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
</tbody>
</table>

Table 6.2 Top-five ranked genes of *S. agalactiae* based on geNorm and NormFinder bioinformatic tools. Genes of *S. agalactiae* prototype strains (O90R, NEM316, and 2603/V/R) were evaluated throughout three logarithmic time-points of growth at 40°C. The “M value” (geNorm) and “Stability value” (NormFinder) are inversely correlated to the stability of the candidate genes. (Adapted from Florindo et al., 2012).

<table>
<thead>
<tr>
<th>O90R</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
<th>NEM316</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
<th>2603/V/R</th>
<th>geNorm (M value)</th>
<th>NormFinder (stability value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>glicK</td>
<td>(0.168)</td>
<td>glicK</td>
<td>gyrA</td>
<td>(0.062)</td>
<td>gyrA</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>recA</td>
<td>(0.168)</td>
<td>recA</td>
<td>gyrA</td>
<td>(0.062)</td>
<td>gyrA</td>
<td>recA</td>
<td>(0.138)</td>
<td>glicK</td>
</tr>
<tr>
<td>glkK</td>
<td>(0.264)</td>
<td>glkK</td>
<td>recA</td>
<td>(0.108)</td>
<td>recA</td>
<td>glicK</td>
<td>(0.267)</td>
<td>glicK</td>
</tr>
<tr>
<td>gldA</td>
<td>(0.276)</td>
<td>gldA</td>
<td>glicK</td>
<td>(0.393)</td>
<td>glkK</td>
<td>gldA</td>
<td>(0.267)</td>
<td>glicK</td>
</tr>
<tr>
<td>hsp40</td>
<td>(0.417)</td>
<td>hsp40</td>
<td>glkK</td>
<td>(0.384)</td>
<td>glkK</td>
<td>hsp40</td>
<td>(0.267)</td>
<td>glicK</td>
</tr>
</tbody>
</table>

71
6.4 Discussion

The choice of the right reference genes to be used for normalization in expression studies is critical, especially when it is known that expression of HKGs can fluctuate under experimental conditions (Bustin, 2002; Bustin et al., 2005; Dheda et al., 2004; Huggett et al., 2005; Vandecastelee et al., 2001). Therefore, we have assayed an amplification-based strategy, the qRT-PCR, in order to verify the expression stability of 10 candidate reference genes in three GBS prototype strains grown at 37°C or 40°C. The transcripts were normalized against the number of bacteria (through determination of gDNA copies) for each time-point of growth. Still, as mRNA and gDNA quantifications are experimentally independent, the lack of reproducibility on both extraction protocols could constitute a critical point of this normalization strategy. However, considering that we have obtained high reproducibility for both RNA and DNA extraction procedures, these steps had no significant influence in the final output of the gene expression stability. Another critical point of this strategy could be the presence of gDNA from dead cells, but this was minimized by selecting the early stationary phase as the last time-point of the study, considering that a preliminary cell viability counting test (data not shown) showed that this phenomenon was only critical above OD<sub>600</sub> ≈ 0.9, i.e., in the beginning of the stationary phase. Also, we opted for not using the colony forming units (CFU) method to normalize gene expression (ratio cDNA/CFU), because S. agalactiae are Gram-positive cocci occurring in short chains in which more than two dividing bacteria can be counted as just one CFU (Koch & Doyle, 1999; Vandecastelee et al., 2002), which could lead to an underestimation of the number of bacteria. A global comparison between the expression values obtained at different temperatures evidence a down-regulation of all genes (except groEL) at 40°C. For the exposed, the results of the present study should not be extrapolated for assays where there is a temperature shift during the same experiment.

Considering our initial strategy, which included the gene validation during five time-points of growth, we verified that gene expression stability was also dependent on the strain under study. We also observed that the lack of stability for the majority of genes was mainly related to significant variations in early stationary phase, contrasting to a seemingly regular expression during the logarithmic phase. Based on this observation, and because most published studies regarding GBS expression were held during the exponential growth, we re-analyzed the data focusing on the three time-points of the logarithmic phase. This reanalysis evidenced an increase in the number of stable genes (Figure 6.3). In fact, at 37°C two genes, recA and sdiA, displayed stability simultaneously for the three S. agalactiae strains, and six genes were stable for at least two strains. At 40°C the results were not so promising; still, three genes (including recA) were stable simultaneously for two strains. Dissimilar expression stability of the same gene for different S. agalactiae strains may reflect divergent regulation in expression of metabolic and virulence
pathways, despite their highly conserved nucleotide sequences. Indeed, the repertoires of genes under influence of regulatory systems, in response to environmental stimuli, appear not to be identical among *S. agalactiae* strains, which may contribute to a strain-specific adaptation of *S. agalactiae* to a preferred host niche, as a commensal or as an invasive pathogen (Jiang et al., 2008; Lamy et al., 2004). Also, it has been shown for different bacterial strains of the same species that genes with identical nucleotide sequence both in the open reading frame and in the promoter region may yield significant different expression levels (Nunes et al., 2007). As a complementary analysis, we further evaluated the gene stability during the logarithmic phase through geNorm and NormFinder. Both bioinformatic applications produced similar gene rankings which overlapped most of the experimental results, where recA showed up again among the best ranked genes (Tables 6.1 and 6.2). Emphasis should also be given to gyrA, the most used reference gene in *S. agalactiae* expression studies (Al Safadi et al., 2010; Brochet et al., 2008; Gleich-Theurer et al., 2009; Santi et al., 2008), which was always among the top-five ranked genes for all strains and temperatures by using both softwares. Although its “experimental” performance was more heterogeneous, the absolute fold-difference values were always close to the cut-off value (Figure 6.3). In conclusion, by using experimental and bioinformatic approaches, we identified and validated a list of stable genes for the three *S. agalactiae* prototype strains grown at 37°C and 40°C, where recA seems to be the best choice. Nevertheless, due to *S. agalactiae* genome diversity, the identification of appropriate reference genes for expression studies may be challenging. We believe that the genes presented here as “stable” are excellent candidates to be tested in future studies involving other *S. agalactiae* strains. Ultimately, the normalization with gDNA proved to be a strong alternative to the use of reference genes.

**6.5 Acknowledgements**

The present work was supported by a Fundação para a Ciência e Tecnologia grant PTDC/SAU-MII/105114/2008 from the Ministério da Educação e Ciência, Portugal, and PhD fellowships SFRH/BD/48231/2008, SFRH/BD/68527/2010, and SFRH/BD/68532/2010 awarded to CF, VB, and RF, respectively.
Nuclease production represents a major determinant for survival of *Streptococcus agalactiae* in human blood


Manuscript in final preparation

Author contributions
BS and GZ designed the study; CF optimized and participated in the semi-quantitative DNase assays, the cloning assays, the production of recombinant proteins, and phylogenetic analyses of DNase genes. JD, NR, SM also performed the other experiments, including western blots and infection assays. FF kindly provided DNase antisera. Analysis of the data: All; BS supervised the study; BS and JD wrote the first draft

Note
Part of the results were present as a poster:
Dick J. Mauerer S, Florindo C, Spellerberg B. The molecular basis of *Streptococcus agalactiae* DNase activity. Autumn Meeting, Society for General Microbiology, University of Nottingham, Nottingham, UK, September 6-9, 2010.


Abstract

*Streptococcus agalactiae* (Group B streptococci, GBS) is a major cause of severe neonatal infections and is increasingly observed in invasive disease in adult patients. To establish an infection *S. agalactiae* have to escape powerful innate immunity mechanisms. Neutrophil extracellular traps represent an important mechanism of innate immunity that different streptococcal species evade by the production of extracellular DNases. While the production of three distinct extracellular nucleases has been reported for *S. agalactiae* 30 years ago the genetic basis of DNase production has not been fully characterized. Using an insertion mutant library, we screened for *S. agalactiae* mutants showing a diminished DNase production on DNA-methyl green agar. The screen led to the isolation of 22 mutants with diminished nuclease activity. Genetic analysis of the insertion sites resulted in the identification of two putative DNase-encoding genes corresponding to the molecular sizes of the previously published 33 kDa and 26.5 kDa nucleases II and III of *S. agalactiae*. To prove DNase activity of the encoded proteins, both genes were expressed as recombinant His-Tag proteins in *E. coli* and analyzed for nuclease activity. To investigate role of the two *S. agalactiae* nucleases for invasive infections we generated a double mutant by inactivation of both nuclease genes. In comparison to the wild-type strain this nuclease deficient mutant displayed a severely diminished ability for survival in human blood and in infection experiments with human granulocytes.

**Keywords:** *Streptococcus agalactiae*, extracellular DNases, neutrophil extracellular traps, immune evasion
7.1 Introduction

*Streptococcus agalactiae* (Group B streptococci, GBS) represents a major bacterial pathogen causing life threatening diseases like pneumonia, sepsis and meningitis in term and premature neonates. In the perinatal period innate immunity mechanisms play a predominant role in the prevention of invasive bacterial diseases. Consequently the successful evasion of innate immunity is an important virulence strategy for neonatal microbial pathogens. For *S. agalactiae* it is known that a number of well-known and characterized virulence factors like the β-hemolysin and the C5a-peptidase contribute to the escape of innate immunity. Neutrophils or PMNs (polymorphonuclear leukocytes) are the most abundant of innate immune cells and represent a first line of defense (Nathan, 2006). One of their most important functions is the phagocytosis of invading extracellular pathogens. However, several years ago a novel mechanism to eliminate bacteria was described for neutrophils; the generation of neutrophil extracellular traps (NETs) that capture bacterial pathogens via the release of chromosomal DNA and kill the enclosed bacteria by the action of antimicrobial peptides, neutrophil enzymes and histones (Brinkmann et al., 2004; Brinkmann & Zychlinsky, 2012). The generation of NETs is activated through contact with pathogenic bacteria, including streptococci, staphylococci and enterococci. Many different species of pathogenic streptococci and staphylococci have been shown to encode nucleases that are crucial for the escape of pathogenic bacteria from these NETs (Beiter et al., 2006; Berends et al., 2010; Sumby et al., 2005). Consequently bacterial nucleases play an important role for microbial pathogens to survive the encounter with granulocytes. These findings are supported by in vivo experiments with nuclease deficient microbial pathogens (Berends et al., 2010). In the context of neonatal group B streptococcal infections the generation of NETs may be of special interest, since the generation of NETs is impaired in neonatal neutrophils (Yost et al., 2009).

More than three decades ago *S. agalactiae* were shown to harbor three different nucleases designated Nuc I, Nuc II and Nuc III in an elegant biochemical study (Ferrieri et al., 1980). These nucleases were shown to have a size of 18.6, 33, and 26.5 kDa and nuclease activity could be detected in 99% of the investigated *S. agalactiae* strains. Surprisingly despite the publication of multiple *S. agalactiae* genome projects with some of them identifying close to 20 putative nuclease genes based on sequence homologies, the genes encoding these biochemically characterized *S. agalactiae* nucleases have not been clearly identified and functionally studied. To identify the corresponding genes of the *S. agalactiae* nucleases that were characterized so many years ago, we screened an insertion mutant library for isolates displaying a diminished zone of nuclease activity on DNA-methyl green agar plates. This screen led to the isolation of several mutants displaying diminished nuclease activity and the identification of the genes coding for two of the previously described *S. agalactiae* nucleases. The activity of both putative nucleases was characterized by the
expression as recombinant fusion proteins. Following the generation of an *S. agalactiae* double mutant of these genes, the role of *S. agalactiae* nucleases for survival of *S. agalactiae* in human blood and in granulocyte assays was further characterized.

### 7.2 Materials and Methods

#### 7.2.1 Streptococcal strains and growth conditions

Streptococcal and *E. coli* strains used in this study are listed in Table 7.1. *E. coli* DH5α served as a host for recombinant pAT28 plasmids, *E. coli* EC101 as host for recombinant pGh9:ISSI plasmids. Streptococci were grown on Tryptone Soya Agar (TSA) Plates with sheep blood (Oxoid) or in THY-broth (Todd-Hewitt broth, Oxoid, supplemented with 0.5% yeast extract, Difco) at 37°C and 5% CO₂. Mutants with mobilized plasmids were grown at 30°C. Mutant strains harboring chromosomally integrated pGh9:ISSI vectors were cultured in medium containing erythromycin (250 μg/ml for LB medium and 1 μg/ml for THY-broth) at a temperature of >37°C, to ensure chromosomal plasmid stability. Mutant strains harbouring cytoplasmic pAT28 plasmids were grown in liquid medium or on agar plates containing spectinomycin (100 μg/ml for *E. coli* and 120 μg/ml for streptococci).
Table 7.1 Bacterial strains and plasmids used in the study.

<table>
<thead>
<tr>
<th>Strain or plasmid Strains</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S. agalactiae</strong> serotype Ia clinical isolate carrying an integration of pGhost5 in the <em>lmh</em> gene</td>
<td>BSU 4</td>
<td>Spellerberg et al. 1999a</td>
</tr>
<tr>
<td><strong>S. agalactiae</strong> serotype Ia clinical isolate</td>
<td>BSU 6</td>
<td></td>
</tr>
<tr>
<td>BSU 575</td>
<td>BSU 6 derivative <em>sak_0814</em>:pGhost9::ISS1</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 576</td>
<td>BSU 6 derivative <em>sak_0814</em>:pGhost9::ISS1</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 617</td>
<td>BSU 6 derivative <em>sak_0814</em>:ISS1</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 620</td>
<td>BSU 6 derivative <em>sak_0814</em>:ISS1</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 623</td>
<td>BSU 6 derivative <em>sak_0220</em>:ISS1</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 738</td>
<td>BSU 6 harboring pBSU409::<em>sak_0814</em>prom</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 741</td>
<td>BSU 6 harboring pBSU409::<em>sak_0220</em>prom</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 746</td>
<td>BSU 620 harboring pAT28::<em>sak_0220</em></td>
<td>This study</td>
</tr>
<tr>
<td>BSU 766</td>
<td>BSU 623 harboring pAT28::<em>sak_0814</em></td>
<td>This study</td>
</tr>
<tr>
<td>BSU 775</td>
<td>BSU 620 derivative <em>sak_0220</em>:pGhost5</td>
<td>This study</td>
</tr>
<tr>
<td>BSU 795</td>
<td>BSU 6 harboring pAT28</td>
<td>This study</td>
</tr>
<tr>
<td><strong>E. coli</strong> DH5α</td>
<td>endA1 hsdR17 supE44 lacU169(f80lacZD(M15) recA1 gyr96 thi-1 relA1</td>
<td>Boehminger</td>
</tr>
<tr>
<td>EC101</td>
<td><em>E. coli</em> JM101 derivative with repA from pWV01 integrated into the chromosome</td>
<td>Law et al., 1995</td>
</tr>
<tr>
<td>BSU 731</td>
<td><em>E. coli</em> BL21 DE3 carrying pET21* Sak_0814* amplified from strain 6</td>
<td>This study</td>
</tr>
<tr>
<td><strong>plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGhost9::ISS1</td>
<td>Eryr ori Ts SpeI ori pUC ori pAmb1</td>
<td>Maguin et al., 1996</td>
</tr>
<tr>
<td>pAT28</td>
<td>pAT28 derivative carrying the 1032 bp open reading frame of <em>sak_0220</em> from strain BSU 6</td>
<td>This study</td>
</tr>
<tr>
<td>pAT28::<em>sak_0814</em></td>
<td>pAT28 derivative carrying the 786 bp open reading frame of <em>sak_0814</em> from strain BSU 6</td>
<td>This study</td>
</tr>
<tr>
<td>pBSU409</td>
<td>pAT28 derivative, carrying a promoterless effA gene</td>
<td>Gleich-Theuer et al., 2009</td>
</tr>
<tr>
<td>pBSU409::<em>sak_0814</em>prom</td>
<td>pBSU409 derivative carrying the promoter region of <em>sak_0814</em></td>
<td>This study</td>
</tr>
<tr>
<td>pBSU409::<em>sak_0220</em>prom</td>
<td>pBSU409 derivative carrying the promoter region of <em>sak_0220</em></td>
<td>This study</td>
</tr>
</tbody>
</table>

7.2.2 General DNA techniques

For DNA preparation and analysis standard molecular biology techniques were used. PCR was performed with Taq polymerase according to the manufacturer’s protocol (Roche), with 30 cycles of amplification steps of 1 min at 94°C, 1 min at 50°C to 56°C, and 1 to 4 min at 72°C depending on primers (Table 7.2) and product size. Genomic streptococcal DNA was isolated as described elsewhere (Pospiech and Neumann, 1995). Plasmid DNA was isolated and purified using the Qiaprep Spin Miniprep Kit (Qiagen) according to the manufacturer’s instructions. Plasmids and PCR products were sequenced on an ABI 373 automated DNA sequencer using the ABI Prism Dye
terminator cycle sequencing kit (PE Applied Biosystems). Streptococcal strains were transformed according to the protocol of Ricci et al. (Ricci et al., 1994).

<table>
<thead>
<tr>
<th>prime</th>
<th>target gene</th>
<th>sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAT28-2</td>
<td>pAT plasmid</td>
<td>CTC TTC GCT ATT ACG CCA GCT</td>
</tr>
<tr>
<td>pAT28-3</td>
<td>pAT plasmid</td>
<td>GTT GTG TGG AAT TGT GAG CCG</td>
</tr>
<tr>
<td>IspGhostSR</td>
<td>IspGhostSR plasmid</td>
<td>ATC TAC TGA GAT TAA GGT CTT AAT GG</td>
</tr>
<tr>
<td>pGhost KS</td>
<td>pGhost KS plasmid</td>
<td>CGA GGT CCA CGG TAT CG</td>
</tr>
<tr>
<td>Spd3-F1</td>
<td>sak_0220</td>
<td>GCC GGA GGT GCT CTA TTG G</td>
</tr>
<tr>
<td>Spd3-R1</td>
<td>sak_0220</td>
<td>TGA TGC CTT TCG CTT CTT T</td>
</tr>
<tr>
<td>Spd3-F2</td>
<td>sak_0220</td>
<td>CTG CCG CAG GTG CTC TAT</td>
</tr>
<tr>
<td>Spd3-R2</td>
<td>sak_0220</td>
<td>TTT CCC AAT AAA ATG ATA AAA T</td>
</tr>
<tr>
<td>SAK-0814-fwd</td>
<td>sak_0814</td>
<td>GAC ATG CTT AAA TAA TAG GC</td>
</tr>
<tr>
<td>SAK-0814-rev</td>
<td>sak_0814</td>
<td>CGG AAT ACT AAG TAA TGT ATC C</td>
</tr>
<tr>
<td>PAT-for</td>
<td>sak_1847</td>
<td>CCT TGA CAT TGA TGC TGC</td>
</tr>
<tr>
<td>PAT-rev</td>
<td>sak_1847</td>
<td>GGT ACG ACT GAA AGT GCT TG</td>
</tr>
<tr>
<td>SAK-0814-fwd</td>
<td>sak_0814</td>
<td>GAG GTA AAG CTT ATG AAA AGA TTA CAT</td>
</tr>
<tr>
<td>SAK-0814-rev</td>
<td>sak_0814</td>
<td>GCC GCC CTC GAG ATT CAT TAT TCT TGC TCT TG</td>
</tr>
<tr>
<td>SAK-0814-fwd</td>
<td>sak_0814</td>
<td>GAG GTA CAT TGT AAA AGA TTA CAT</td>
</tr>
<tr>
<td>SAK-0814-rev</td>
<td>sak_0814</td>
<td>GCC GCC GAA TTC ATG AAA TTA TCT AAA CAA TTA</td>
</tr>
<tr>
<td>Spd3-His-fdr</td>
<td>sak_0220</td>
<td>GAG GTC CTC GAG TTG TGC TTC AGT AGT GCT GT</td>
</tr>
<tr>
<td>Spd3-His-rdr</td>
<td>sak_0220</td>
<td>GAG GTA AAG CTT AGC CCA CGG AAT TTT GCC</td>
</tr>
<tr>
<td>Komp.SAK.for</td>
<td>sak_0814</td>
<td>ATC GGA TCC GCG GGG AAT TAG AAA GAG GT</td>
</tr>
<tr>
<td>Komp.SAK.rev</td>
<td>sak_0814</td>
<td>TAT GAA TTC CGG CTT GAG AAA TTA TTT GCT C</td>
</tr>
<tr>
<td>Komp.SAK.F2</td>
<td>sak_0814</td>
<td>ATC GGA TCC GCG GGT AGA TAT GGT AGA CGC</td>
</tr>
<tr>
<td>Komp.SAK.R2</td>
<td>sak_0814</td>
<td>TAT GAA TTC GGG CAC TAA AAA GGA TAA CAA C</td>
</tr>
<tr>
<td>Komp.Spd3.for</td>
<td>sak_0220</td>
<td>CTA GGA TCC GCG CCA TAT CAT TTA GAA GGG</td>
</tr>
<tr>
<td>Komp.Spd3.rev</td>
<td>sak_0220</td>
<td>TTG GAA TTA TTC CGG ACC GGC AAA TTT GGC</td>
</tr>
<tr>
<td>SAK-0814-PROMO-F1</td>
<td>sak_0814</td>
<td>GCG CGG GAA TTC TCA CAG AAC AAG CCT CCT TAC</td>
</tr>
<tr>
<td>SAK-0814-PROMO-R1</td>
<td>sak_0814</td>
<td>GAG GAA GGG CAG CTT ATT ACA TAC TAC CCA</td>
</tr>
<tr>
<td>SAK-0814-PROMO-R3</td>
<td>sak_0814</td>
<td>GAG GTA GGA TCC CTT ATC ATA AGC TTT ACC TC</td>
</tr>
<tr>
<td>spd3FORhomologsProm</td>
<td>sak_0814</td>
<td>GCC GGA ATT CAT AGT TAT TAT ACA TGA CTA CC</td>
</tr>
<tr>
<td>spd3REVhomologsProm</td>
<td>sak_0814</td>
<td>GCC GAG ATG TCT TTT AGA TAA TTT CAT AAA AAG</td>
</tr>
<tr>
<td>spd3kofor</td>
<td>sak_0814</td>
<td>GCC GCC GGA TTC CTT AGC TTA TGG ACC GAG GT</td>
</tr>
<tr>
<td>spd3korev</td>
<td>sak_0814</td>
<td>GCC GCC GAA TCC CCT TAT TGT CAA CGT AGA CG</td>
</tr>
</tbody>
</table>

### 7.2.3 Screening for *S. agalactiae* mutants deficient in nuclease activity and identification of chromosomal integration sites

The screening for *S. agalactiae* mutants displaying a diminished nuclease activity was carried out using a previously constructed mutant library (Spellerberg *et al.*, 1999a) that is based on the undirected chromosomal integration of the vector pGh9:ISS1 (Maguin *et al.*, 1996). This library was screened for mutants showing a diminished zone of clearance on DNA methyl green agar plates (Becton Dickinson). In mutants that were selected for further investigation, chromosomal
pGh9:ISS1 integration sites were determined as described previously (Spellerberg et al., 1999a). Mobilization of the pGh9:ISS1 vector from individual mutants leading to stable mutants harbouring a single copy of the ISS1 Insertion element at the original integration site were generated by induction of rolling circle replication at 30°C in the absence of antibiotic pressure.

### 7.2.4 Semiquantitative measurement of *S. agalactiae* nuclease activity

Semi-quantitative nuclease assays were performed as previously described (Sumby et al., 2005), with some modifications. Briefly, filtered supernatants from the *S. agalactiae* wild-type and mutant strains were harvested from overnight liquid THY cultures (stationary growth phase). One microgram of a purified double-stranded PCR amplicon (~500 bp) was incubated with increasing volumes of *S. agalactiae* culture supernatant (1 to 10 μl) in the presence of 1x buffer M (Roche) at 37°C for 0.5, 1 and 2 hours. Nuclease reaction was stopped with EDTA (0.5 M, pH 8.0) at 4°C. The samples were analyzed visually by 1% agarose gel electrophoresis for DNA digestion. A negative control consisting on a reaction mixture without supernatant was used in all experiments.

### 7.2.5 Complementation of nuclease-negative mutants

To complement nuclease deficient mutants with an insertion of ISS1 in *spd3* or *sak0814* primers (Table 7.2) were designed to amplify the respective genes of *S. agalactiae* strain BSU 6 and the resulting PCR products were cloned into the vector pAT28. The vector pAT28: *spd3* was introduced in *spd3:ISS1* mutant strain BSU 623 and vector pAT28: *sak0814* into the *sak0814:ISS1* mutant strain BSU 620 by electroporation. Resulting clones (BSU 766 and BSU 764) were selected on THY-agar plates supplemented with spectinomycin and checked for the presence of the recombinant pAT28 vector. Subsequently, the phenotype of complementation mutants was evaluated on DNA-methyl green agar plates.

### 7.2.6 Generation of an *S. agalactiae* sak_0220 and sak_0814 double nuclease mutant

A double mutant of the nuclease genes *sak_0220* and *sak_0814* was generated by the integration of the vector pGhost5 into the *sak_0220* gene of *S. agalactiae* strain BSU 617. BSU 617 carries an ISS1 insertion in the *sak_0814* gene that is located 270 nucleotides downstream of the ATG start codon. The strain BSU 617 was generated through mobilization of the ISpGhost9 vector from strain BSU 575. BSU 575 was one of the nuclease mutants selected in the initial screen of the ISpGhost 9 library. To integrate the vector pGhost5 into the gene *sak_0220* of BSU 617 a fragment of the gene was amplified through the primers 5'-GGC GGC GGA TCC CTT AGC TTA TGG ACC GAG GT-3', 5'-GCG GGC GAA TTC CCT TAT TTG CAA CGT AGA CG-3' and subcloned into pGost5 in *E. coli* DH5α. Correct construction of the vector was verified by DNA sequencing. Transformation and integration of pGhost5 into the genome of strains BSU 617 was
carried out as described previously (Spellerberg et al., 1999b). The loss of nuclease activity in the resulting strain BSU 775 was assessed visually on DNA methyl green agar plates.

7.2.7 Phylogenetic analysis of sak_0814 and sak_0220 nucleotide sequences

Phylogenetic trees of the nucleotide sequences of sak_0814 and sak_0220 were constructed using the MEGA 4. For the analysis nucleotide sequences of 10 reference S. agalactiae strains (8 human strains and 2 bovine strains) available on the GenBank database were chosen. Analysis was carried out employing the “p-distance” method by calculating the percentage of nucleotide differences in each nuclease gene.

7.2.8 Gene expression of the sak_0220 and sak_0814 genes of S. agalactiae

To characterize the gene expression of the sak_0220 gene and the sak_0814 gene of S. agalactiae the putative promoter regions of the respective genes were cloned into a streptococcal EGFP plasmid pBSU409 (Gleich-Theurer et al., 2009). For this purpose the promoter region of the sak_0220 gene homologue was amplified with the primers: for 5’- GGC GGA ATT CAT AGT TAT TAT ACA TGA CTA CC-3’ and rev 5’- CGC GGG ATC CTG TTT AGA TAA TTT CAT AAA CC-3’ (restriction sites are underlined). The resulting PCR products and the vector pBSU409 were digested with the enzymes BamHI and EcoRI, ligated, and transformed into DH5a cells. The correct construction of the plasmids was verified by PCR with primers flanking the insertion site (pAT28-3 and pAT28-EGFP4) (Table 7.2) of pBSU409 and sequencing of the resulting PCR products. The recombinant plasmids were transformed into the S. agalactiae strain BSU 6 and selected on THY agar plates supplemented with Spectinomycin (120 µg/ml). To investigate the expression of the nuclease II and III gene in different growth phases, FACS analysis (fluorescence-activated cell sorting) of the S. agalactiae strains carrying the pBSU409::spd3prom (BSU 741) and pBSU409::sak0814prom (BSU 738) vector was carried out as described elsewhere (Aymanns et al., 2011) at the OD_{600} 0.2, 0.4, 0.6 and 0.8. To investigate the influence of DNA on the expression of sak_0220 and sak_0814 the strains BSU 741 and BSU 738 harboring the recombinant reporter plasmids were grown to mid-logarithmic phase (OD_{600} = 0.4), bacterial cells were harvested through centrifugation and resuspended in PBS containing 0, 1 and 2 µg/ml of DNA. DNA was provided as PCR products of the S. agalactiae atr gene, generated as described in Jones et al. (Jones et al., 2003). Following incubation at 37°C for 1 hour, bacterial cells were collected through centrifugation, resuspended in PBS and fluorescence was quantified through FACS analysis. To characterize the influence of glucose in the growth medium, overnight cultures of the strains BSU 741 and BSU 738 were grown in THY-broth supplemented with 0, 0.25%, 0.5%, and 1% of glucose. Bacterial cultures were pelleted by centrifugation, washed once in PBS, resuspended in PBS and measured by FACS analysis. The S. agalactiae strains BSU 98 (carrying plasmid pBSU101) and BSU 99 (carrying plasmid pBSU100) were used as positive control and negative controls for FACS analysis (Aymanns et al., 2011).
7.2.9 Western Immuno Blot analysis

Western Immunoblot of recombinant Sak_0220 protein of *S. agalactiae* (computed molecular weight 38.9 kDa). The coding region of the *S. agalactiae sak_0220* was subcloned into the His-Tag vector pET21a, resulting in a C-terminal translational fusion of the protein with the histidine-Tag. Following expression of the construct in *E. coli* strain BL21(DE3). The recombinant protein was purified over a Nickel column under native conditions. 5 μg of protein in each lane were separated by SDS-PAGE electrophoresis blotted onto nitrocellulose membrane and probed with anti-Histag antibody (lane 1) or a polyclonal anti *S. agalactiae*-nuclease antibody (lanes 2-4). The following protein samples were tested for reactivity: Purified protein native conditions, lanes 1+4. *E. coli* BL21(DE3) total protein lysate prior to (lane 2) and after IPTG induction (lane 3).

7.2.10 Whole blood killing assay

To assess the ability of nuclease deficient mutants for survival in human blood, heparinized blood was collected from healthy human volunteers and 1 ml of blood was inoculated with the $10^4$ CFU (colony forming units) of the double nuclease mutant BSU 775. Strain BSU 4 carrying a pGhost5 vector integrated into the *imb* gene of *S. agalactiae* (Spellerberg et al., 1999b) served as a positive control. Samples were incubated at 37°C under shaking conditions for 4 hours. After 1, 2 and 4 hours the number of viable bacteria was determined by plating an aliquot of the sample on THY agar plates containing 1 μg/ml of erythromycin. To supplement the mutant strain BSU 775 by providing external DNAse activity, the assay was performed as described above with the addition of 10 μl of sterile filtered supernatant (native or heat inactivated) from an overnight culture of strain BSU 4.

7.2.11 Granulocyte survival assay

Isolation of human granulocytes was performed from heparinized blood of healthy human volunteers as described elsewhere (Sagar et al., 2013). To assess the survival of the nuclease deficient strain BSU 775. $10^5$ granulocytes were infected with the *S. agalactiae* strains BSU 4 and BSU 775 at a multiplicity of infection (MOI) of 1 and 10. Viable bacterial counts were quantified by plating on THY agar plates after 2 and 4 hours of incubation at 37°C.
7.3 Results

7.3.1 Identification of the *S. agalactiae* nuclease genes

To identify genes responsible for the nuclease activity of *S. agalactiae* a previously constructed *S. agalactiae* mutant library based on the undirected chromosomal integration of the pGhost:ISS1 vector was used (Spellerberg et al., 1999a). Screening was carried out on DNA-methyl green agar plates, assessing the zone of clearance around single colonies. Due to the existence of multiple nucleases in *S. agalactiae*, a total lack of nuclease activity could not be expected for mutants carrying vector insertions in single genes. Therefore all clones displaying a clearly diminished zone of clearance surrounding the colonies were selected in the initial screen, resulting in about 400 clones. Repeated testing of these clones was performed on DNA methyl green agar plates and in semi-quantitative nuclease assays (Figure 7.1) resulting in a reduction of the selected clones to 22 that displayed a reproducibly diminished nuclease activity.

![Image of gel with lanes labeled as 1, 3, 5, 7, 10, neg, WT, BSU575, BSU576](image)

**Figure 7.1** Depicted are the results of a semi-quantitative nuclease assay. 2 μg of DNA were incubated for 30 min at 37°C using increasing amounts of crude GBS culture supernatant (1, 3, 5, 7, and 10 μL). Incubation of DNA with PBS served as negative control. Shown are the *S. agalactiae* strains BSU6 wild-type (WT), BSU 575 and BSU 576, which are both carrying an insertion in the nuclease III gene *sak_0814*.

Determination of the insertion sites of these nuclease mutants revealed that two independent insertions were found in the *S. agalactiae* gene *sak_0220* that displays significant homology to the *spdB* gene of the *S. pyogenes* streptodornase. The overall identity of the two proteins is 62% at the amino acid level, while the *S. pyogenes* streptodornase is however somewhat smaller (266 versus 343 amino acids) than the gene product of *sak_0220*. The mutations we found were located at nucleotide 774 and 794 of *sak_0220*. Furthermore the two mutants displaying the strongest reduction of nuclease activity in the semi-quantitative nucleases assay (BSU 575, BSU 576) carried independent insertions in the gene *sak_0814* at nucleotide 271 and nucleotide 752.
sak_0814 represents a gene that is identical to gbs0661 and has independently from our work just very recently been described as encoding *S. agalactiae* nuclease activity (Derré-Bobillot et al., 2013). It is 261 amino acids long and harbors a motif of the endonuclease NS_2 superfamily. To confirm the nuclease activity of the proteins encoded by the *S. agalactiae* sak_0220 gene and the newly identified putative nuclease gene sak_0814, both genes were expressed as His-Tag proteins and the nuclease activity of the recombinant proteins was analyzed in semi-quantitative nuclease assays (Figure 7.2). In these assays a strong nuclease activity was clearly demonstrated after 1 and 4 hours of incubation for the recombinant sak_0814 gene product.

![Nuclease activity assay](image)

**Figure 7.2** Enzymatic nuclease activity of the two putative *S. agalactiae* nuclease genes sak_0814 and sak_0220 (spd3) was evaluated by expression as recombinant His-Tag proteins. 2 μg of DNA were incubated with increasing amounts of recombinant proteins as indicated, during 60 and 240 min at 37°C. The recombinant laminin adhesin Lmb of *S. agalactiae* served as a negative control. Lane 1, DNA molecular weight 50 bp ladder.

The histidine fusion protein of sak_0220, demonstrated degradation of the DNA sample in this assay, but was not as active as the recombinant histidine fusion protein of Sak_0814. However Sak_0220 displays a high similarity to Spd3, a well characterized nuclease of *S. pyogenes*. To further substantiate the nuclease activity of the proteins encoded by sak_0814 and sak_0220, we complemented the mutants of both genes by introducing the open reading frames of sak_0814 and sak_0220 into the vector pAT28. Upon transfer of the recombinant pAT28 construct into the strains.
BSU 617 and BSU 623 the deficiency of nuclease activity in these strains was partially restored (Figure 7.3). To assess the amount of nuclease production in *S. agalactiae* that can be attributed to the two genes *sak_0814* and *sak_0220*, a double nuclease mutant was created by insertion of the vector pGhost5 into the *sak_0220* gene of the *sak_0814* mutant strain BSU617.

![Image of petri dishes](image)

**Figure 7.3** *S. agalactiae* mutants of the genes *sak_0814* (BSU 617) and *sak_220* (BSU 623) and the respective complementation strains of these mutants (BSU 764 and BSU 766) were grown on DNA methyl green agar plates to assess DNase production. BSU 795 represents the wild-type strain. Nuclease production can be seen as transparent halos surrounding single colonies.
Loss of nuclease activity for this double mutant was almost complete as visualized on DNA-methyl green agar plates (Figure 7.4). In summary, the findings we obtained for the mutants of the sak_0220 gene of S. agalactiae and the sak_0814 gene strongly support the notion that these two genes encode the major nucleases of S. agalactiae.

![Image of agar plates showing nuclease activity]

**Figure 7.4** Nuclease activity of the nuclease producing strain BSU 4, the sak_0814 mutant strain BSU 576 and the double nuclease mutants strain BSU 775 on DNA methyl green agar plates. BSU 4 represents the positive control strain displaying regular nuclease activity.

### 7.3.2 Phylogenetic analysis of sak_0220 and sak_0814 nucleotide sequences

To investigate the phylogenetic relationship of the nucleotides sequences of sak_0814 and sak_0220, ten S. agalactiae reference genomes were selected including eight human and two bovine strains. Analysis using the program MEGA 4 showed a high conservation of the sequences for sak_0814 with an average nucleotide substitution per gene of 2.4 (Figure 7.5). Among the human strains the sak_0814 gene of ST17 and ST19 strains appear to be distinct and different from the other human strains that cluster together, irrespective of sequence type. In contrast to the sak_0814 gene, the nucleotide sequences for the sak_0220 gene are more heterogeneous (Figure 7.5), on average 9.6 nucleotides exchanges per gene are present. For sak_0220 more or less distinct alleles could be observed for each clonal complex. Interestingly in some of the genomes deposited in the GenBank database bigger mutations of sak_0220 are present. In several ST61 strains an insertion of 17 nucleotides resulting in a premature stop codon is present. S. agalactiae strains 2603V/R and 18RS21 revealed the presence of several premature stop codons in sak_0220 that would lead to a truncated version of Sak_0220 protein.
7.3.3 Western blot experiments

In previous biochemical experiments three nuclease of *S. agalactiae* have been characterized at a functional level (Ferrieri *et al.*, 1980). They were shown to have molecular sizes of 18600 Da (± 2800) for nuclease I, 33000 Da (± 8800) for nuclease II and 26500 Da (± 6700) for nuclease III. The computed size for the protein encoded by *sak_0814* (29.4 kDa) corresponds well to the estimated size of nuclease III and the gene product of *sak_0220* with a predicted weight of 38.9 kDa is compatible with the biochemically determined size of nuclease II. Polyclonal rabbit antisera generated against nuclease II and III of *S. agalactiae* were fortunately still available and tested for their reactivity with the recombinant histidine fusion proteins of *sak_0814* and *sak_0220* in Western blot experiments. While no reactivity could be observed for the nuclease III specific
antiserum (data not shown), the nuclease II specific antiserum clearly reacted with the recombinant histidine fusion protein of sak_0220 of *S. agalactiae* (Figure 7.6).

**Figure 7.6** Western Immunoblot of recombinant Sak_0220 protein of *S. agalactiae* (computed molecular weight 38.9 kDa). The coding region of the *S. agalactiae* Sak_0220 was subcloned into the His-Tag vector pET21a, resulting in a C-terminal translational fusion with the histidine-Tag. Following expression of the construct in *E. coli* strain BL21(DE3). The recombinant protein was purified over a Nickel column under native conditions. 5 μg of protein in each lane were separated by SDS-PAGE electrophoresis blotted onto nitrocellulose membrane and probed with anti-HisTag antibody (lane 1) or a polyclonal anti *S. agalactiae*-nuclease antibody (lanes 2-4). The following protein samples were tested for reactivity: Purified protein native conditions, lanes 1+4; *E. coli* BL21(DE3) total protein lysate prior to (lane 2) and after IPTG induction (lane 3).

### 7.3.4 Expression analysis for sak_0220 and sak_0814

To investigate the expression of sak_0814 and sak_0220 under different growth and environmental conditions, the promoter regions of both genes were introduced into the EGFP expression vector pBSU409 as detailed in 7.2.8 (Materials and Methods). Both plasmids were transferred into strain BSU 6, generating the strains BSU 738 and BSU 741 that carry reporter gene constructs for sak_0814 and sak_0220, respectively. Initial expression analysis was carried out at different growth phases and displayed that the highest expression of sak_0814 can be observed in the late logarithmic growth phase (OD$_{600}$ = 0.8) shortly before the bacterial cells enter into stationary phase (Figure 7.7A). In contrast to these findings sak_0220 expressions shows an early peak at an OD$_{600}$ of 0.2 drops in mid-logarithmic phase and shows highest values in the overnight culture.
To analyse whether the nuclease genes of *S. agalactiae* are under carbon catabolite control, expression was also measured under increasing glucose conditions. For this purpose the strain BSU 738 and BSU 741 were grown overnight in THY broth supplemented with 0 to 1% of glucose. For both genes high glucose conditions resulted in a significantly reduced reporter gene activity (Figure 7.7B). Furthermore the expression of both genes following contact with DNA was determined. For these experiments strains BSU 741 and BSU 738 were grown to mid-logarithmic phase and exposed to 1 μg/ml and 2 μg/ml of DNA. Under these conditions a moderate increase of reporter gene activity was observed for the *sak_0814* gene, whereas no significant difference could be observed for *sak_0220* (Figure 7.7C).

**Figure 7.7.** Expression analysis of *sak_0220* and *sak_0814* by FACS analysis. The strain BSU738 (carrying the promoter region of *sak_0814* upstream of EGFP) and the strain BSU 741 (carrying the promoter region of *sak_0220* upstream of EGFP) were analyzed. To determine gene activity in different growth phases, reporter gene activity was measured at an OD₆₀₀ of 0.2, 0.4, 0.6, 0.8 and in overnight cultures (A). To investigate the effect of glucose on the expression of both genes, strains were grown overnight in regular THY broth or THY broth supplemented with 0.25, 0.5 and 1% of glucose (B). To determine the effect of external DNA supplementation, bacteria were grown to mid-logarithmic phase, washed and incubated in the concentration of DNA as indicated. Shown are mean values of 5 independent experiments (C).
7.3.5 The effect of nucleases for survival of *S. agalactiae* in human blood and granulocyte assays

To characterize the role of nuclease activity for the ability of *S. agalactiae* to survive or multiply in fresh human blood, heparinized blood samples (1 ml) were inoculated with $10^4$ bacterial cells of the nuclease deficient strain BSU 775 and the nuclease producing strain BSU 4. Quantification of bacterial survival after 1, 2 and 4 hours showed a significantly reduced survival of strain BSU 775, that was most pronounced after 4 hours of incubation (Figure 7.8). Survival rates for the nuclease deficient strain BSU 775 could be restored to some extent by supplementation of the assay with 10 μl of sterile culture supernatant from an overnight culture of the nuclease producing strain BSU 4. As shown in the semi-quantitative analysis of nuclease activity the sterile filtered culture supernatant of *S. agalactiae* displays strong nuclease activity.

![Figure 7.8](image.png)

**Figure 7.8** Growth curves of the nuclease producing *S. agalactiae* strain BSU 4 and the double nuclease mutant strain BSU 775 in whole human blood. 1 ml of blood was inoculated with $10^4$ cfu per strain and viable bacterial counts were determined after 1, 2 and 4 hours of incubation by sub-culturing on agar plates. Experiments were repeated 5 times, displayed are the results of one representative assay.

To substantiate our observation of the role of *S. agalactiae* nucleases, the nuclease producing strain BSU 4 (positive control) and the nuclease deficient strain BSU 775 were tested for their ability to survive the presence of granulocytes (Figure 7.9). Freshly isolated human granulocytes were infected with a MOI of 1 and 10 with both strains and survival was quantified after 2 and 4 hours incubation time. For both time-points and both MOIs, the survival of the nuclease mutant strain was significantly reduced in comparison to the nuclease producing strain. The results of these assays show that bacterial nuclease activity plays a very prominent role for the survival of *S. agalactiae* in the contact with granulocytes.
Figure 7.9 Human granulocytes were infected with *S. agalactiae* strain BSU 775 carrying a double nuclease mutation at a multiplicity of infection (MOI) of 1 and 10 as indicated. After 2 and 4 hours of incubation detection of surviving bacteria was performed by subculture on blood agar plates. Depicted are the measurements of five independent experiments. *S. agalactiae* BSU 4 was used as a positive control.

### 7.4 Discussion

Nuclease activity of *S. agalactiae* has first been described and characterized in detail as early as 1980. Three bacterial nucleases were identified and biochemically characterized (Ferrieri *et al.*, 1980). These nucleases are secreted into the culture supernatant, display DNase as well as RNase activity and nuclease activity could be observed in close to 100% of the *S. agalactiae* strains examined. Bacterial nucleases are produced by many microbial pathogens and renewed interest arose from the observation that bacterial nucleases are essential for the escape of bacterial pathogens from neutrophil extracellular traps. This novel mechanism for the elimination of microbial pathogens through the innate immune system was first detected and described in 2004 by Brinkmann *et al.* (Brinkmann *et al.*, 2004). However, despite numerous genome sequencing projects resulting in the identification of up to 20 different putative nucleases in *S. agalactiae*, the genetic basis of the biochemically characterized nuclease activity of *S. agalactiae* remained unclear for more than three decades. To identify the genes responsible for the nuclease activity of *S. agalactiae*, we screened a previously generated and well established mutant library (Spellerberg *et al.*, 1999a) for clones displaying a reduced nuclease activity on DNA-methyl green agar plates and in semi-quantitative nuclease assays. Our screen led to the identification of two candidate genes for the previously described nuclease II and nuclease III of *S. agalactiae* that are annotated as *sak_0220* and *sak_0814* in the nucleotide sequence deposited in GenBank (strain A909, accession nr. CP000114.1). Interestingly comparing the amino acid sequence of *sak_0814* with the GenBank database high homologies can only be observed for other *S. agalactiae* sequences. For all of the
deposited S. agalactiae genomes identities of sak_0814 at the protein level are extremely high (99-100%). Comparing the amino acid sequences of Sak_0814 from different S. agalactiae strains in a phylogenetic analysis (Figure 7.5) confirmed the high conservation of this gene. On average only 2.4 nucleotide substitutions are observed. Interestingly the sak_0814 genes of strains from the clonal complex 17 and 19 appear to be quite distinct from the other alleles of this gene, which cluster together (Figure 7.5). In contrast to this observation, the alleles of sak_0220 are much more heterogeneous and in some of the strains insertions as well as premature truncations occurred. Close homologues of sak_0220 are found in other streptococcal species and, especially in S. pyogenes, this gene plays an important role for the nuclease activity (Sumby et al., 2005). Surprising is the lack of close homologues for sak_0814 in other streptococcal species from the pyogenic group; Especially in view of the fact that the function, nuclease activity, with the ability to destroy neutrophil traps appears to be highly conserved in other streptococci (Beiter et al., 2006; Buchanan et al., 2006; Sumby et al., 2005).

Based on the sizes of the identified DNase genes, the 38.9 kDa protein encoded by sak_0220 corresponds best to the nuclease II of S. agalactiae. This finding could be substantiated with Western blot experiments showing that previously generated antisera against nuclease II displayed a positive reaction with the His-Tag fusion protein of Sak_0220 (Figure 7.6). The Sak_0814 protein displayed a computed size of 29.4 kDa that fits with the predicted size of the nuclease III protein. Our screen did not lead to the identification of a gene encoding the nuclease I of S. agalactiae, which does however not contradict the existence of a third nuclease. In fact, the double nuclease mutant (BSU 775) generated displays some residual DNase activity surrounding the colonies on DNA-methyl green agar (Figure 7.4) supporting the existence of nuclease I or other still unidentified. The amount of loss of nuclease activity in strain BSU 775 does however indicate that the genes sak_0814 and sak_0220 encode the major nucleases of S. agalactiae. Taken together, these data confirm the publication of Ferrieri and co-authors, which identified three distinct nucleases in S. agalactiae over 30 years ago and provide for the first time the genetic basis of nuclease II and nuclease III.

Expression analysis of the two DNase genes revealed quite different expression profiles (Figure 7.7A). While the sak_0814 gene expression gradually increases over the logarithmic growth phase until it reaches maximum levels at an OD_{600} of 0.8, the sak_0220 gene has an early peak at an OD_{600} of 0.2, is low during early logarithmic phase and shows maximum levels in overnight cultures. This data shows that in mid-logarithmic phase, the main nuclease activity is provided by sak_0814 and that both genes are active in late logarithmic phase. In streptococci, as well as other bacteria, many virulence factors are under carbon catabolite control. In a recent investigation the negative regulation of many S. agalactiae virulence factors under high glucose conditions was reported (Di Palo et al., 2013). Nucleases as potential virulence factors may also be
controlled through glucose. A down regulation of the transcription could be observed for sak_0220 as well as for sak_0814 with increasing glucose conditions (Figure 7.7), supporting the hypothesis that both nucleases are under a carbon catabolite control mechanism.

The effect of the *S. agalactiae* nucleases for surviving exposure to the innate immune system was investigated in experiments determining the bacterial survival in whole human blood and in the encounter with human granulocytes. In both settings the double nuclease mutant strain BSU 775 displayed a significantly reduced ability to survive (Figures 7.8 and 7.9). A partial rescue of this effect could be achieved by providing 1% of sterile filtered bacterial culture supernatant from a nuclease producing strain. *S. agalactiae* culture supernatants contain a very high nuclease activity, which may presumably help the double nuclease mutant strain to escape from NETs. Even though we did not visualize the formation of NETs in the whole blood survival assay, bacteria are known to lead to an effective stimulation of NET formation in neutrophils (Fuchs *et al.*, 2007). The time-point at which we observed the biggest differences in survival between the nuclease producing wild-type strain and the nuclease deficient mutant strain BSU 775 in whole blood is very well compatible with the formation of bacterial NETs. NET formation usually occurs after 2 hours of stimulation with a NET inducing stimulus like for example PMA (Fuchs *et al.*, 2007). Results of the granulocyte assays support the data from the whole blood survival assays, since at the identical time-points a severely diminished survival of strain BSU 775 was evident (Figure 7.9). These results are of special interest in the context of neonatal infections. It has just recently been shown that neonatal granulocytes are impaired in their ability to form NETs, in connection with the strong nuclease activity of *S. agalactiae* provided through the genes sak_814 and sak_0220 this may contribute to the high susceptibility of neonates for invasive *S. agalactiae* infections.

In summary our investigation led to the identification of the genes encoding the previously described nuclease II and III of *S. agalactiae*. Infection experiments support their role in the escape of innate immunity and with our data we could show that both nucleases are a major determinant for survival of *S. agalactiae* in human blood.

### 7.5 Acknowledgements and Transparency Declaration

The work of JD and BS was supported by DFG Grant GSC 270 International Graduate School in Molecular Medicine Ulm. PhD fellowship SFRH/BD/48231/2008 and DAAD grant to CF. Conflicts of interest: None.
Evaluation of the DNase activity in clinical strains of *Streptococcus agalactiae* of human and bovine origin

**Florindo C. Damião V, Gomes JP, Nunes B, Santos-Sanches, I; Borrego MJ, and Spellerberg B**

Manuscript in preparation

*Author contributions*
CF, BS and MJB designed the study; CF and VD performed the all the molecular characterization and DNase assays; CF, JGP performed and analysed the qRT-PCR data; CF, ISS, MJB and BS evaluated the molecular data; BN performed the statistics; BS, MJB and ISS supervised the study; CF and VD wrote the first draft; MJB performed the draft revision

*Note*
Part of the results were included in the Thesis of Vera Damião for Master Degree in Molecular Genetics and Biomedicine, entitled *Produção de DNases extracelulares em estipes de Streptococcus agalactiae de origem humana e bovina*, FCT/UNL, November 2012.
Abstract

*Streptococcus agalactiae* is the leading cause of neonatal pneumonia, sepsis and meningitis, and emerging infection disease among adults with underlying medical conditions. Extracellular DNases contribute to the spread of pathogenic bacteria through the evasion of the host innate immunity; however, only recently its role in the pathogenesis of *S. agalactiae* has been clarified. The main objective of this study was to evaluate the production of extracellular DNases by *S. agalactiae* clinical strains and to perform a correlation of the DNase phenotype with other epidemiological variables, such as, capsular type, genetic lineage, clinical origin (colonization and infection) and host (human or bovine) in order to better understand the virulence potential of particular clones. A collection of 345 *S. agalactiae* clinical strains was extensively characterized by capsular typing, MLST, PFGE, Multiple-Locus Variant-Repeat Assay, antibiotic resistance profiling, detection of mobile elements and surface proteins, and the evaluation of their DNase activity by qualitative and quantitative assays. All the bovine *S. agalactiae* strains (N = 60) and 86% of the human *S. agalactiae* strains (N = 285) showed DNase activity. Of note, all the *S. agalactiae* strains without DNase activity belonged to the same genetic lineage: CC19 (capsular types II, III-1 and V). Genetic and transcriptomic analysis of the DNase encoding genes revealed important clues that may contribute to the absence of DNase activity among the majority of CC19 strains. In order to identify a particular nonproducing DNase CC19 clone, an extensive sub-characterization was implemented, including PFGE, MLVA and the detection particular mobile genetic elements. GBS11 was mostly found in ST28 strains whereas the mobile element IS1548 was identified in the remaining CC19 strains. All strains, except one carried the *rib* gene. Five PFGE restriction patterns were identified indicating that the majority of the CC19 strains was clonal. The present study highlights the fact that a *S. agalactiae* strain belonging to the CC19 has a high probability to display a DNase (-) phenotype.

**Keywords:** *Streptococcus agalactiae, extracellular DNases, DNase activity assays*
8.1 Introduction

Classically, two strategies by which neutrophils serve as a first line of defense against invading pathogens are understood: the secretion of antimicrobial peptides (degranulation) and the engulfment of bacteria (phagocytosis). More recently, Brinkmann and co-authors (Brinkmann et al., 2004) characterized neutrophil extracellular traps (NETs) as a novel additional antimicrobial function of these specialized leukocytes. Neutrophils produce NETs in response to gram-positive and other pathogens and are thought to kill microbes by exposing them to high local concentrations of antimicrobial effectors (Brinkmann et al., 2004). The structure of NETs, held together by the DNA backbone, is critical for their antimicrobial function. A shared mechanism of bacterial escape from NET entrapment by means of extracellular DNase production has now been described for S. pyogenes (Buchanan et al., 2006; Sumby et al., 2005), S. pneumoniae (Beiter et al., 2006), S. aureus (Berends et al., 2010), and most recently for S. agalactiae (Derré-Bobillot et al., 2013; Dick et al. – Chapter VII), promoting neutrophil resistance and the spread of infection in vivo. Whereas each pathogen deploys a different suite of virulence factors, encoded by unique sets of genes and possessing unique chemical structures, the cumulative effect of these features provides each pathogen significant resistance to phagocyte recruitment and activation, opsonophagocytosis, bacterial entrapment and uptake, and the microbicidal activities of key host defense factors. For example, in S. pyogenes, the acquisition of the potent bacteriophage-encoded DNase Sda1 may have been a critical step in the evolution of the hypervirulent MIT1 clone that has disseminated globally as a leading agent of severe invasive infections (Walker et al., 2007). In S. agalactiae, the ST17 lineage strains are also disseminated worldwide which causes significantly more meningitis in neonates than strains of other lineages and hence is considered as a highly virulent clone (Héry-Arnaud et al., 2005; Manning et al., 2009; Martins et al., 2007). In addition, ST17 strains more frequently cause meningitis than sepsis, and late onset disease than early onset disease (Manning et al., 2009; Tazi et al., 2010). These data support the idea that serotype III ST17 strains have a particular ability to invade the central nervous system of the neonates. This hypothesis is consistent with studies highlighting genetic variations in virulence genes between S. agalactiae clonal groups (Brochet et al., 2006; Florindo et al., 2011 – chapter V; Manning et al., 2009). Therefore, we aimed to correlate the DNase activity of S. agalactiae clinical strains with several epidemiological variables, which may elucidate the virulence potential and/or host tropism of particular genetic lineages, such as ST17.
8.2 Material and Methods

8.2.1 Strain collection

*S. agalactiae* reference strains belonging to different genetic lineages were used in this study: [2603V/R (genotype: V/ST110); COH1 (genotype: III/ST17); NEM316 (genotype: Ia/ST23); O90R (genotype: Ia/ST25)], as well as clinical strains of bovine and human origin.

Bovine *S. agalactiae* strains (N = 60) were isolated from cases of subclinical mastitis diagnosed in Portugal between 2002 and 2003, which were subject to prior characterization (Rato et al., 2012). Briefly, the following genotypes were identified: III-3/ST23 (N = 1); V/ST2 (N = 14); New 1/ST2 (N = 6); New 2/ST61 (N = 11); New 3/ST61 (N = 4); New 3/ST554 (N = 20); New 4/ST61 (N = 1) and New 5/ST2 (N = 3). All the strains displaying “New” capsular types were nontypeable by serology. The capsular types were designated “New” due to lack of identity with the alleles included in the *cpsD-E-F* database (Florindo et al., 2010 – chapter III; Rato et al., 2012).

*S. agalactiae* strains isolated from humans (colonization and infection) were calculated from a stratified sample, determined by statistical analysis using the computer program SPSS (SPSS software, version 16.0, IBM; Fisher Exact test, significance of 5%), based on the following criteria: *i*) distribution of capsular types among 100 colonizing *S. agalactiae* strains isolated in Lisbon region (Florindo et al., 2010 – chapter III; *ii*) correlation between capsular type and the production of DNases (data not shown) within the strain collection mentioned in *i*; and *iii*) hypothetical higher DNase activity among invasive *S. agalactiae* strains in comparison to carriage strains. Since the biological role of extracellular DNases in *S. agalactiae* may be related to the evasion of the host immune system and consequently, the establishment of an infection, this hypothesis was tested.

Based on the sampling criteria, we selected 157 *S. agalactiae* colonizing strains isolated from the anogenital exudates of pregnant women (last trimester of gestation); on the other hand, we selected 128 *S. agalactiae* invasive strains isolated from sterile body fluid (peripheral blood, cerebral spinal fluid and pleural fluid) from infants (age < 3 months) which developed *S. agalactiae* sepsis or meningitis.

*S. agalactiae* colonizing strains were isolated from the routine laboratory diagnosis of *S. agalactiae* at National Institute of Health of Lisbon, INSA, between 2005 and 2008; Invasive strains were isolated from pediatric hospitals and affiliated microbiological laboratories in Germany in the period between 2001 and 2003 (Table 8.1) (Fluegge et al., 2011).
Table 8.1 Collection of *S. agalactiae* clinical strains used in the present study.

<table>
<thead>
<tr>
<th>Capsular serotype/genotype</th>
<th>Number of colonizing strains (n=157)</th>
<th>Number of invasive strains (n=128)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Ib</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>II</td>
<td>39</td>
<td>17 a)</td>
</tr>
<tr>
<td>III-1</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>III-2</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>IV</td>
<td>10</td>
<td>3 a)</td>
</tr>
<tr>
<td>V</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

a) The number of invasive strains previously defined in the sampling was not achieved due to the low prevalence of serotype in neonatal infection (Fluegge et al., 2011).

In order to relate the DNases activity of *S. agalactiae* with phenotypic and molecular epidemiological data, we constituted the following work plan (Figure 8.1):

![Figure 8.1 Algorithm for the study of *S. agalactiae* strains.](image-url)
8.2.2 *S. agalactiae* identification and antimicrobial susceptibility profile

*S. agalactiae* was isolated in accordance with the CDC guidelines (Verani *et al*., 2010). Subcultures on 5% sheep-blood agar plates were performed, and *S. agalactiae* strains were identified by standard criteria on the basis of colony morphology, Gram staining, and group B latex-agglutination assay (Becton Dickinson). Antimicrobial susceptibility testing (penicillin G, erythromycin, clindamycin and vancomycin) was executed by E-test according to Clinical and Laboratory Standards Institute guidelines (CLSI, 2009), and the presence of macrolide resistance-associated genes (*ermTR, ermB* and *mefA*) was analysed by PCR amplification, as described elsewhere (Gygax *et al*., 2006; Sutcliffe *et al*., 1996).

8.2.3 Capsular genotyping, MLST, MLVA

Capsular genotyping was carried out by PCR and DNA sequencing of the *cpsD-cpsE-cpsF* region, as documented previously (Florindos *et al*., 2010 – chapter III).

For the MLST method (Jones *et al*., 2003), PCR fragments (~500 bp) of seven housekeeping loci were amplified and sequenced. Alleles of all loci were examined on an MLST database (http://pubmlst.org/sagalactiae/) and the combination provided an allelic profile or ST. Clonal complexes (CCs) comprising isolates sharing six or seven identical alleles were defined.

The analysis of the variable number of tandem repeats (VNTR) present in *SAG2, SAG3, SAG4, SAG7, SAG21* and *SAG22* gene according to nucleotide sequences of the reference strains 2603V/R, NEM316, and A909 was performed as previously (Haguenoer *et al*., 2011). The determination of the allelic profile was based on the visualization of the bands present on the agarose gel. The number of repetitions for each VNTR was deduced from the size of the amplicon, when compared to the reference strain for which the number of repetitions was known (Haguenoer *et al*., 2011). The number assigned to each allele corresponds to the number of repeats in this gene, allowing the definition of an MLVA genotype for each strain (allelic profile).

8.2.4 Alpha-like protein (Alp) family and MGEs

The molecular characterization included the study of a major antigen, the Alp family, which was analysed by multiplex PCR for direct identification of the *alpha-C, rib, epsilon* and *alp2–alp4* genes (Gherardi *et al*., 2007). The presence of two mobile genetic elements (MGEs), IS1548 and GBS1I, within the *scpB–lmb* intergenic region was evaluated by PCR, as described previously (Al Safadi *et al*., 2010). In the absence of MGEs, the presence of the flanking genes (*scpB* and *lmb*) was evaluated.
8.2.5 DNase activity, qualitative assays

DNase production was assessed qualitatively by inoculation of all strains (human or bovine origin) of the *S. agalactiae* collection on DNA-methyl green agar plates (Oxoid). *S. agalactiae* strains O90R or NEM316 were used as a positive control in order to validate the test. The interpretation of the results was done after 24 hours of incubation at 37°C in an atmosphere of 5% CO₂. Strains were considered DNase producers when displaying transparent halos around colonies of *S. agalactiae*.

8.2.6 DNase activity, semi-quantitative and quantitative assays

The semi-quantitative assessment of the activity of the DNases was performed for all *S. agalactiae* strains (bovine and human) in order to confirm the results of the qualitative tests on DNA-methyl green agar, taking into account the limitations of their interpretation / visualization. The quantitative evaluation of the activity of DNases was only performed for *S. agalactiae* strains defined as DNase (-) based on the results obtained from qualitative and semi-quantitative methods. Semi-quantitative assays were based on methods described by Sumby and co-authors (Sumby *et al.*, 2005), with some modification. Thus, we proceeded to the inoculation of *S. agalactiae* strains into 5 ml of Todd-Hewitt broth supplemented with 0.5% yeast extract - THY (Oxoid). Culture supernatant of *S. agalactiae* from stationary phase was achieved by centrifugation (10 min., 3000 rpm) and syringe filtration (0.2 μm). Subsequently, 1 μg of double stranded DNA (*atr* amplicon, Jones *et al.*, 2003) was incubated at 37°C with different volumes of culture supernatant in the presence of 1x M buffer (Roche), for four time points: 1 h, 2 h, 4 h and "overnight" (~ 17h) in order to evaluate the digestion of DNA in a final volume of 50 μl. Nuclease reaction was stopped with EDTA (0.5 M, pH 8.0) at 4°C. The samples were analyzed visually by 1% agarose gel electrophoresis for DNA digestion. A negative control consisting on a reaction mixture without supernatant was used in all experiments.

The quantitative DNase assays were performed by measuring the amount of DNA present in each sample, after the end of each incubation period. For this purpose, the fluorescent PicoGreen dye (Invitrogen) was used to quantify the dsDNA according to manufacturer's instructions. Briefly, 1 ml of 1x Quant-iT PicoGreen was added to an equal volume of each sample previously diluted in 1x TE buffer. After 5 minutes of incubation at room temperature, in the dark, we proceeded to fluorescence measurement (Fluorimeter - Anthos Zenith 3100) by using 96 well microtiter plates (Corning 96 Well Clear Flat Bottom Polystyrene Black TC). To calculate the concentration of DNA in each sample, we determined a standard curve with four solutions of phage Lambda DNA of known concentrations (1, 10, 100 and 1000 ng/ml): \( y = 3675x \) (x = concentration of dsDNA ng / ml, y = fluorescence); Correlation Factor \( (R^2) = 0.9998 \). Each fluorescence value obtained was
subtracted the value of fluorescence of the blank solution (PicoGreen 1x + 1x TE buffer at a ratio of 1:1). The final results were based on three independent experiments.

8.2.7 Expression of DNase genes by qRT-PCR

The transcriptomic level of DNase genes sak_0220 and sak_0814 from S. agalactiae 2603V/R and NEM316 was evaluated during three time-points of the logarithmic growth phase at 37°C, OD<sub>600</sub> = 0.2, OD<sub>600</sub> = 0.5 and OD<sub>600</sub> = 0.8. Briefly, S. agalactiae were grown in THY in 5% CO<sub>2</sub> at 37°C overnight as standing cultures. Dilutions of 1:50 of these cultures were used to inoculate triplicate cultures of fresh THY broth (50 ml) that were allowed to incubate without shaking at 37°C with 5% CO<sub>2</sub>. Cell growth of S. agalactiae strains 2603V/R and NEM316 was monitored by optical density at 600 nm (OD<sub>600</sub>). At each time-point, bacterial cells were collected by centrifugation, resuspended in PBS and incubated for 1 h at 37°C in the presence of 2 μg/ml of DNA (stimulus, calf thymus DNA). Then, 1 ml of each bacterial culture was collected, homogenized and immediately subjected to RNA isolation by using the RNeasy mini kit (Qiagen) according to manufacturer’s instructions. Residual contaminant DNA was removed by using 30 U RNase-free DNase (Qiagen), and RNA elution was done in 40 μl of RNase-free water. Extracted RNA was finally stored at -80°C until use.

cDNA was generated from 20 ng of each RNA sample collected at each time-point, by using TaqMan RT reagents (Applied Biosystems), as previously described (Florindo et al., 2012 – chapter VI). The qRT-PCR was performed by using ABI 7000 SDS, SYBR Green chemistry and optical plates (Applied Biosystems). The following primers were used to amplify sak_0220 and sak_0814, respectively: Forward, 5’- CAG TAG TGC TGT GAT GTT TG; Reverse, 5’-TTG ATT TAA CGC TTC TTG; Forward, 5’- GTC TTC CAA CGC GCC GCA AA; Reverse, 5’- AAC ACC CGA TAG TAC ATG CTG. The qRT-PCR reagents consisted of 1× SYBR Green PCR Master Mix (Applied Biosystems), 400 nM of each primer and 5 μl of cDNA. in a final volume of 25 μl. All samples were run in duplicate and ‘no template controls’ (NTC) and ‘no-RT’ controls were included in all runs to exclude potential DNA contamination. Thermocycling amplification consisted of an initial denaturation at 95°C for 10 min followed by 40 cycles of 95°C/15 s and 60°C/1 min. The gene expression was determined from the respective standard curves by conversion of the mean threshold cycle (Ct) values. Finally, raw qRT-PCR data was normalized against the transcript level of the stable gene recA, as previously demonstrated (Florindo et al., 2012 – Chapter VI). The final expression results were based on three independent experiments for prototype strains NEM316 and 2603V/R.
8.3 Results

8.3.1 Identification of the genetic lineages within the capsular types

Five genetic lineages have been identified: CC1, CC12, CC17, CC19 and CC23 within our \textit{S. agalactiae} human collection. Different STs were identified, including in \textit{S. agalactiae} strains of the same capsular type (Table 8.2) In particular, serotypes II and IV demonstrated greater heterogeneity, presenting distinct STs belonging to clonal complexes CC1, CC12, CC19, CC23 and CC1, CC12, CC23, respectively.

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Capsular type & Total strains, N & Colonizing strains, N & Invasive strains, N & \multicolumn{2}{c|}{Percentage of strains displaying DNase activity; [STs]; CCs; (frequency)} \\
\hline
\textit{Ia} & 40 & 20 & 20 & \multicolumn{2}{c|}{Colonization} \\
& & & & \textbf{100\% (20/20)} & \textbf{[ST23 (17/20); ST144 (2/20); ST24 (1/20)]} \\
& & & & \textbf{CC23 (20/20)} & \textbf{[ST23 (19/20); ST24 (1/20)]} \\
\hline
\textit{Ib} & 20 & 10 & 10 & \multicolumn{2}{c|}{Infection} \\
& & & & \textbf{100\% (10/10)} & \textbf{[ST8 (5/10); ST12 (2/10); ST1 (1/10); ST10 (1/10); ST563 (1/10)]} \\
& & & & \textbf{CC1 (2/10); CC12 (8/10)} & \textbf{[ST10 (5/10); ST8 (4/10); ST12 (1/10)]} \\
\hline
\end{tabular}
\caption{Evaluation of the DNase activity by qualitative and semi-quantitative assays among \textit{S. agalactiae} of human origin (colonization vs infection).}
\end{table}
83% (32/39)
[ST25 (10/32); ST12 (8/32); ST44 (4/32); ST10 (3/32); ST12 (2/32); ST43 (1/32); ST154 (1/32); ST249 (1/32); ST347 (1/32); ST472 (1/32)]
CC19 (17/32); CC12 (12/32); CC1 (2/32); CC23 (1/32)

48% (14/29)
[ST19 (8/14); ST27 (2/14); ST106 (2/14); ST286 (1/14); ST369 (1/14)]
CC19 (14/14)

59% (17/29)
[ST19 (17/29)]
CC19 (17/17)

100% (29/29)
[ST17 (28/29); ST1287 (1/29)]
CC17 (29/29)

100% (10/10)
[ST196 (3/10); ST2 (2/10); ST1 (1/10); ST3 (1/10); ST162 (1/10); ST23 (1/10); ST10 (1/10)]
CC1 (7/10); CC23 (2/10); CC12 (1/10)

90% (18/20)
[ST2 (13/18); ST1 (5/18)]
CC1 (18/18)

8.3.2 Qualitative and semi-quantitative DNase assays

All bovine *S. agalactiae* strains (60/60) and 86% (245/285) of *S. agalactiae* strains of human origin displayed DNase activity. Among *S. agalactiae* collection of human origin, DNase activity varied with capsular type; thus, all strains belonging to types Ia, Ib, III-2 and IV displayed DNase activity, independently of their clinical origin (colonization or infection). In contrast, the
percentage of non-producing strains among types II, III-1 and V was 82 to 88%, 48 to 59% and 90%, respectively, depending on the clinical origin. Nevertheless, no statistical association between DNase production and infection was observed \( (P > 0.05) \).

Genetic lineages defined by MLST allowed the identification of the CC19 lineage within the strains harboring capsular types II / III-1/ V presenting a DNase (−) phenotype \( (P<0.01; 0 \leq \text{Odds Ratio} \leq 0.024, 95\% \text{ CI}) \). However, although the DNase (−) phenotype is related to CC19 strains, we also identified CC19 members displaying DNase activity.

In order to confirm the DNase (−) phenotype previously determined by qualitative assays, all \( S. agalactiae \) strains lacking DNase activity were subjected to semi-quantitative assays, as exemplified in the Figure 8.2. Qualitative and semi-quantitative assays yielded identical results.

![Figure 8.2](image)

**Figure 8.2** Semi-quantitative assay in gel electrophoresis for the evaluation of DNase activity. 1 µg of DNA (amplicon atr) incubated with 10 µl of \( S. agalactiae \) culture supernatant (NEM316 vs CC19 clinical strain) for 1h, 2h, 4h, overnight at 37°C. Negative control: without culture supernatant. M, Molecular weight Ladder.

### 8.3.3 Quantitative analysis of DNases

In order to confirm the absence/residual DNase activity of some CC19 \( S. agalactiae \) strains, we calculated the amount/percentage of the remaining DNA after incubating (1 h, 2 h, 4 h and 17 h) a known quantity of DNA with culture supernatant of \( S. agalactiae \) collected from the stationary growth phase at 37°C (Figure 8.3).
Figure 8.3 Quantitative DNase assays displaying differential DNase activity between *S. agalactiae* strains over time. Independent experiments were performed in duplicate with mean values displayed. Error bars show ± standard deviation. Col, Colonization; Inf, Infection.

Quantitative DNase assays confirmed a null or residual DNase activity over time within a population of CC19 strains. In accordance with these results, prototype strain 2603V/R showed a disability to degrade DNA, whose amount remained nearly constant after 17 hours (Figure 8.3). In contrast, *S. agalactiae* NEM316 and the clinical strains belonging to other genetic lineages, such as CC17, CC23, CC61 and revealed a substantial digestion of the DNA, reflecting a high production of extracellular DNases. Except ST61 strains of bovine origin, the profiles of DNA digestion were quite similar within strains of the same ST. Once more, the clinical origin did not seem to influence the DNase production. However, more strains should be included to test these findings.

### 8.3.4 Sub-characterization of CC19 strains

In order to identify the putative genetic determinants associated with the production / non production of DNases, all CC19 strains (N = 97; 55 colonizing strains and 42 invasive strains) were subjected to additional molecular and phenotypic characterization. Based on MLVA, two profiles/genotypes were identified within CC19 strains. All DNase producing and not producing CC19 strains, from colonization or infection, showed the profile 33 (3,3,3,5,0,2) except ST28 strains, in which the profile was identified 32 (3,3,1,5,0,2).
Regarding the antibiotic susceptibility testing, none CC19 strain was resistant to penicillin (MIC between 0.032 and 0.125 µg/ml) and vancomycin (MIC between 0.25 and 1 µg/ml). In addition, none invasive CC19 strains displayed resistance to macrolides, contrasting to colonizing CC19 strains, which have shown a macrolide resistance rate of 31% (17/55; III-1, 16 strains; II, one strain). Nine of these 17 strains (16.4% of total CC19 colonizing strains) showed simultaneous resistance to erythromycin (MIC ≥ 256 mg/ml) and clindamycin, attributable to the presence of \textit{ermB} gene. The remaining CC19 strains were only resistant to erythromycin, \(2 \leq \text{MICs} \leq 6\) µg/ml, which presented \textit{ermA} (N = 8) and \textit{mefA} (N = 1). However, a correlation between antibiotic susceptibility profile and DNase activity was not possible to establish.

The screening of MGEs within the \textit{scpB-lmb} intergenic region and the study of \textit{alp} genes also failed to distinguish among CC19 DNase producers and non-producers. In fact, 78% of CC19 strains (76/97) carried the \textit{IS1548}, whereas 99% (96/97) displayed the \textit{rib} gene.

### 8.3.5 Expression of DNase genes

To analyze if there were differences in the expression of DNase genes between \textit{S. agalactiae} strains to correlate with the virulence potential, we used qRT-PCR to determine the mRNA levels at three logarithmic growth time-points. We aimed to find DNase genes showing differences in the highest mRNA levels during the growth cycle (peak of expression) or in the variation of mRNA levels throughout the growth (profile of expression) between \textit{S. agalactiae} strains. We detected substantial differences of expression only for \textit{sak_0814}; this DNase showed a middle/late cycle gene profile of expression, in which mRNA levels were evident for \textit{S. agalactiae} NEM316 but only vestigial for \textit{S. agalactiae} 2603V/R. The In addition, in both \textit{S. agalactiae} strains the expression of \textit{sak_0220} was residual which helped to explain in part the major role of \textit{sak_0814} in strain NEM316 and the lack of DNase activity of strain 2603V/R.
Figure 8.4 Transcription profile of *sak_0814* of *S. agalactiae* NEM316 and 2603V/R during three logarithmic time points. DNA (2 µg/ml) of was tested as a stimulus. Data are displayed as the mean (based on 3 independent experiments) fold change in transcription of the *sak_0814* relative to the control gene *recA*.

Figure 8.5 Transcription profile of *sak_0220* of *S. agalactiae* NEM316 and 2603V/R during logarithmic growth phase (early, mid and late log). DNA (2 µg/ml) of was used as a stimulus. Data are displayed as the mean (based on 3 independent experiments) fold change in transcription of the *sak_0220* relative to the control gene *recA*.

8.4 Discussion

We have provided evidence that the great majority of *S. agalactiae* produces DNases, that is, it was found that 100% of bovine strains and 86% of strains of human origin displayed DNase activity. Our results corroborated earlier findings published by Ferrieri and co-authors (Ferrieri et al., 1980) which found exactly 86% of DNase producing strains of human origin. Despite the
limitations of our bovine *S. agalactiae* collection (limited genetic diversity and absence of colonizing isolates), to our knowledge this is the first study reporting DNase activity in bovine *S. agalactiae*. Moreover, quantitative assays proved the high activity of DNases of bovine *S. agalactiae*, in particular belonging to ST61 lineages, in DNA digestion.

Importantly, we verified that DNase production may vary with the capsular type and genetic lineages. In fact, a percentage of strains displaying a DNase (-) phenotype was found among capsular types II, III-1, and V, and it was restricted to a single genetic lineage, CC19. Although we speculated about the possible existence of a higher production of extracellular DNases among invasive *S. agalactiae* strains due their pathogenic potential and therefore their greater need for acquisition of virulence factors, no significant differences on DNase activity were found between strains of colonization and infection, suggesting that the production of extracellular nucleases is independent of the clinical origin.

Although the DNase (-) phenotype displayed by some CC19 strains may be suggestive of the absence of genes encoding DNases, we confirmed the presence of *sak_0220* and *sak_0814* in all CC19 strains (data not shown) by PCR and sequencing. In fact, the nucleotide sequences of both DNase genes were 100% identical to the corresponding genes in *S. agalactiae* 2603V/R. Our transcriptomic findings on *sak_0814* revealed that the peak of expression occurs in the later stages of the growth cycle, which may suggest that is not involved in the first host-pathogen interactions. This expression profile of *sak_0814* was in agreement with the expression profile of the DNase genes in *S. pyogenes*, whose transcript levels increased upon entry of the bacteria into the stationary phase of cell growth (Sumby *et al.*, 2005).

The putative inactivation of *sak_0220* due to the presence of several premature stop codons (truncated protein) and the vestigial expression of *sak_0814* (so far, the principal DNase of *S. agalactiae*) during the logarithmic growth phase, may justify in part the lack of DNase activity within CC19 subpopulation. We may speculate that the absence of DNases, as virulence factors, may contribute to the previously described association between CC19 strains and the colonization status (Martins *et al.*, 2007). This hypothesis is consistent with studies highlighting genetic and transcriptomic variations in virulence genes between *S. agalactiae* clonal groups. For example, although ST19 lineage strains can also cause infection, they are less likely to cause meningitis than ST17 strains, even though these two STs share the same *cps* type (serotype III) (Jones *et al.*, 2006; Lin *et al.*, 2006; Luan *et al.*, 2005; Manning *et al.*, 2009). Therefore, our data suggest the variability of DNase activity and consequently DNase gene expression in the various lineages that compose the *S. agalactiae* anogenital population, which might be one of the factors that explain the different roles of strains of each lineage in infection. Nevertheless, we do not yet understand the exact participation of each DNase gene in the evasion from the host immune system, neither if the level of activity of DNases influences the escape from the NETs.
In an attempt to identify the subpopulation of CC19 strains lacking DNase activity, we implemented several molecular techniques presenting high discriminatory power, such as MLVA and studied some epidemiological markers. However, the analysis of the polymorphism of tandem repeats in different loci only reinforced the clonality of CC19 strains, as demonstrated by MLST. Although Haguenoer et al. (Haguenoer et al., 2011) define the MLVA as an epidemiological tool displaying higher discriminatory power than MLST, and thus, allowing the sub-typing of homogeneous genetic lineages, such as CC17, it did not allow the separation of CC19 strains in clones associated with the production or non-production of extracellular DNases.

Of note, the antimicrobial susceptibility tests focusing CC19 strains revealed that strains resistant to macrolides not displayed DNase activity. However, this feature was only observed for the CC19 colonizing strains and, in addition, we verified the existence of non-producing DNase susceptible to macrolides, precluding the establishment of an association between the DNase (-) phenotype and the antibiotic resistance, in particular to macrolides. Also, the study of the presence of mobile genetic elements in the intergenic region scpB-ilm not provided new evidences to distinguish the CC19 subpopulations concerning the DNase activity.

A comparative analysis of epidemiological markers used did not allow the establishment of correlations with the production of DNases. Recent studies (Sorensen et al., 2010) warn that difficulty, since they failed to predict the virulence of the strains based on epidemiological markers. According to Sorensen and co-authors (Sorensen et al., 2010) there is no single phylogenetic model to evaluate the evolution of *S. agalactiae*, whose genome is the result of multiple recombination events, during which some clones emerged and spread globally. Thus, according to these authors, *S. agalactiae* is composed of a rather heterogeneous population, provided with an almost unlimited set of genes, according to the concept of "open pan-genome", described by Tettelin et al. (Tettelin et al., 2005). This genomic architecture could justify the lack of correlation between capsular type, genotype, tropism, and other properties as a result of frequent recombination, resulting in a dynamic genomic structure described by Tettelin and Brochet (Brochet et al., 2006, 2008; Tettelin et al., 2005).

In the near future, a comparative genomics, transcriptomics analysis, and infection assays involving reference strains and clinical strains of *S. agalactiae* CC19 from different hosts, will be crucial to better identify molecular features associated to DNase activity.
Chapter IX

Concluding Remarks
9. Concluding Remarks

Group B *Streptococcus* are common asymptomatic colonizers of the digestive and genitourinary tracts of healthy humans that emerged as the leading cause of bacterial neonatal infections in Europe and North America during the 1960s. The reasons for this emergence are unknown, but recent genomic analysis and phylogenetic reconstructions suggest that the expansion of human *S. agalactiae* clones present in the maternal anogenital tract, which may thus infect neonates, were preceded by the insertion of MGEs conferring tetracycline resistance, due to extensive use of this antibiotic from 1948 onwards (Da Cunha et al., 2014).

*S. agalactiae* can subvert host defenses and, consequently, cause opportunistic invasive infections and tissue damage that may lead to serious morbidity in survivors or cause death. Like many pathogenic bacteria, *S. agalactiae* encodes a myriad of virulence factors that are crucial for its ability to cause disease. In order to reduce *S. agalactiae* infections, epidemiological studies were conducted, prenatal screening programs were evaluated, colonizing and antimicrobial resistance rates were assessed and the distribution of serotypes worldwide was somehow monitored through the publication of various studies in several countries. In addition, the identification of specific allelic variants of known virulence factors and the discovery of new ones has been done to better understand the pathogenicity of *S. agalactiae* and also to perform careful drug target selection during vaccine design studies. Despite these advances, the biological role of the extracellular DNAses, which play a prominent role in other pathogenic streptococci is not fully deciphered and a special attention to them was proposed regarding the *S. agalactiae* pathogenesis, in particular their contribution to the host immune evasion.

9.1 Insights into the carrier state

Maternal colonization is recognized as the major risk factor of neonatal *S. agalactiae* infections through vertical transmission in pregnancy. Since 1996, the widespread implementation of screening guidelines for women during pregnancy or at delivery coincided with a decline of the EOD *S. agalactiae* infections in several countries. Also, in Portugal, a study based on *S. agalactiae* neonatal disease in Portugal reported a reduction of approximately 40% in the incidence of disease, as well as of the case-fatality rates (Neto, 2008). However, in chapter II, we observed that the efficacy of the screening-based approach on its capacity to predict *S. agalactiae* colonization status at the time of labor may be influenced by several variables. Although the antenatal *S. agalactiae* screening is practiced in the Portuguese hospitals, which seem to follow the protocol for screening and prevention of *S. agalactiae* issued by the Portuguese Society of Pediatrics (Almeida et al., 2004; based on CDC guidelines), we verified the existence of multiple laboratory methodologies for *S. agalactiae* diagnosis that may originate discordance, including false-positive results that possibly will culminate in antibiotic overtreatment increasing the risk for antibiotic resistance and
increasing the incidence of neonatal infections caused by pathogens other than *S. agalactiae* (e.g. *E. coli*) (Schrag *et al.*, 2002). This scenario is not unique to Portugal, because the antenatal *S. agalactiae* screening policies and the use of standardized laboratory methods are still under debate and evaluation in Europe, in contrast to the United States where the *S. agalactiae* screening guidelines are well defined and implemented by CDC since 1996 with periodic updates. Nevertheless, in the United States, a recent retrospective analysis of labor and prenatal records of mothers of neonates with EOD showed that 57.9% of the cases had one or more implementation errors (Verani *et al.*, 2014). In fact, some authors found that about 60 to 80% of all EOD cases occurred in neonates with negative maternal screening for *S. agalactiae* during pregnancy (Van Dyke *et al.*, 2009). Efforts to reduce missed opportunities of prevention should include the appropriate clinical management of women whose *S. agalactiae* colonization status is unknown, particularly those at risk of preterm delivery; moreover, it is crucial to identify the factors that contribute for false negative screening results.

Although the majority of European countries, except Bulgaria, Denmark, Greece, Norway, and the United Kingdom offer universal antenatal screening for *S. agalactiae* between 35 and 37 weeks gestation, the existing guidelines depend on the country-specific health professional body. These data reinforce the need to improve and to harmonize the antenatal *S. agalactiae* screening in Europe as well as the microbiological procedures for *S. agalactiae* identification and characterization, as shown by the results from the first international multicenter external quality assessment studies for laboratory identification and typing of *S. agalactiae*, which involved 14 European countries (Portugal not included) (Afshar *et al.*, 2011). Hence, novel prevention approaches such as improved intrapartum assays and vaccines are also needed.

### 9.2 Molecular epidemiology and antibiotic resistance

The genetic diversity of our human *S. agalactiae* collection was assessed by using different methods including the capsular typing, PFGE, MLST and MLVA. In chapters III and IV, we verified the predominance of capsular serotypes/genotypes Ia, Ib, II, III and V among colonized women, in accordance to other studies held in Lisbon area and in other European and North American countries (Ippolito *et al.*, 2010; Martins *et al.*, 2007). Importantly, our work revealed the putative emergence of *S. agalactiae* strains belonging to serotype IV, whose frequency is considered rare worldwide, and described the novel association between serotype IV and the CC17 hypervirulent *S. agalactiae* genetic lineage. Several genomic clues suggested the putative emergence and the clonal expansion of a novel epidemic clone, IV/ST291, similar to the emergence of serotype V *S. agalactiae* strains in the 1990s. This scenario may become risky when the emergence of particular capsular types combines with antibiotic resistance. Indeed, in our studies, we observed an increasing of the resistance rates to macrolides and lincosamides from 2005 to 2012, which may be related to particular genotypes such as III-1/ST19 and V/ST1. Although we
did not detect resistance to penicillin or to vancomycin, *S. agalactiae* clinical strains displaying reduced susceptibility to penicillin have been isolated in Sweden, Japan and in the United States and very recently resistance to vancomycin has been described in a study in the United States, involving two type II/ST22 invasive strains; these data represent a major threat to prophylaxis and treatment of *S. agalactiae* infections (Dahesh et al., 2008; Kasahara et al., 2010; Nagano et al., 2008; Srinivasan et al., 2014). Antimicrobial prophylaxis may have unwanted long-term effects due to increased antimicrobial use and alternative prevention strategies are focused on the development of vaccine formulations. Thus, it is important to continuously monitor the capsular serotype distribution as it has direct implications for the polysaccharide vaccine development. However, the identification of a *S. agalactiae* strain lacking the entire capsular loci may represent a drawback of the polysaccharide vaccines, because it does not recognize these organisms that retain the ability to colonize the anogenital tract (Creti et al., 2012).

Publications performed in the course of the present PhD Thesis (Florindo et al., 2010, 2011, 2014b – chapters III to V) should constitute a contribution to a better understanding of the spread of antibiotic resistance among *S. agalactiae*, which may also represent an important reservoir of resistance genes for other species, in particular other streptococci.

So far, the epidemiology of *S. agalactiae* in most developed countries is well documented, but remains sparse in low/middle income countries, with serious consequences for the prophylaxis of *S. agalactiae* infections and for the implementation of a vaccine covering the local serotypes. In chapter IV, we identified and characterized the causative genetic lineages of *S. agalactiae* of childhood meningitis in Luanda, Angola, providing the MLST data of invasive *S. agalactiae* strains from Africa (Florindo et al., 2011 – chapter V). Despite the limitations of our study, namely the low number of *S. agalactiae* strains and the lack of colonizing strains, we observed a high proportion of the hypervirulent CC17 lineage, in accordance to a contemporary study held in Nairobi, Kenya (Huber et al., 2011). Both studies demonstrate the global dissemination of CC17 strains and their ability to also cause disease in African children and adults.

While these studies only represent an insignificant part of Africa, both suggest that the population structure of local invasive *S. agalactiae* may overlap those described in the United States and in Europe. In fact, a subsequent CRISPR1 locus analysis of our Angolan strains belonging to CC17 (N = 18) (data not shown; obtained from a collaboration with Philippe Glaser from Institute Pasteur, France) revealed that only one strain matched the CC17 African subgroup defined by Lopez-Sanchez, and colleagues (Lopez-Sanchez et al., 2012), composed by CC17 strains mostly isolated in Africa, namely in Senegal, Madagascar and Central African Republic. This data confirms the existence of sub-lineages among the CC17, which may reflect different evolutionary states and distinct ability to cause infection.
9.3 Study of Extracellular DNases

Nuclease activity in *S. agalactiae* has been demonstrated in 1980s (Ferrieri et al., 1980), but the functional characterization on the genetic background and the importance of this phenomenon concerning the *S. agalactiae* evasion mechanisms remained unclear until the beginning of the work presented in chapter VII. The screening of DNase-deficient mutants allowed us to identify two DNase encoding genes, *sak_0220* and *sak_0814*, corresponding to the nucleases II and III, respectively, described by Ferrieri and co-authors (Ferrieri et al., 1980). Whole genome comparison between *S. agalactiae* and *S. pyogenes* revealed high degree of nucleotide identity between *sak_0220* and *spd3*; and *sak_0814* with *sda1*, respectively. Spd3 and Sda1 activity has been shown to promote *S. pyogenes* escape from phagocytic killing with DNA-based NETs generated by neutrophils to capture and eliminate bacteria at tissue foci of infection (Buchanan et al., 2006; Sundby et al., 2005; Walker et al., 2007).

The existence of other nucleases in *S. agalactiae* was supported by the residual DNase activity displayed by the double-DNase mutant described in chapter VII; moreover, the analysis of the *S. agalactiae* genomes revealed the presence of several genes encoding putative extracellular DNases (Derré-Bobillot et al., 2013). These data are in accordance to *S. pyogenes* which produces at least four DNases that promotes the bacterial spreading from the site of the initial infection (Bisno et al., 2003; Miyakawa et al., 1985). In vitro infection experiments confirmed the involvement of Sak_0220 and Sak_0814 of *S. agalactiae* in the escape from innate immunity, promoting the course of infection. The production of multiple and distinct DNases would preserve the ability to degrade DNA when *S. agalactiae* infects a host with enzyme-inhibiting antibodies to one DNase. It is also possible that production of multiple DNases with different substrate cleavage specificities or other characteristics (e.g., pH profile and temperature requirement) would provide a survival advantage by enhancing the range of conditions across which the DNase activity is functional. A third possibility is that possession of multiple DNase genes, including chromosomally encoded and prophage-encoded, enhances the probability that DNase activity will be made at distinct phases of the infection cycle.

Besides the NETs digestion, other biological functions of extracellular DNases have been described. Interestingly, it has been proposed that extracellular DNases might also penetrate inside eukaryotic cells and attack chromosomal DNA (Bonsor et al., 2008). An emerging view is that extracellular DNases also modulate biofilm formation through the degradation of extracellular DNA (eDNA), a key component of the biofilm matrix, which has been reported as a putative nutrient source for the pathogen (Blokesch and Schoolnik, 2008; Kiedrowski et al., 2011; Seper et al., 2011; Steichen et al., 2011). Recently, it was demonstrated the ability of DNase Sda1 of the hyperinvasive MIT1 *S. pyogenes*, to alter its own extracellular CpG rich DNA fragments, modifying the TLR9-mediated recognition by the host innate immune cells. This constitutes a
novel mechanism of bacterial immune evasion involving extracellular DNases based on autodegradation of a key pattern-recognition molecule (Uchiyama et al., 2012).

In chapter VIII, we demonstrated that DNase activity of human *S. agalactiae* strains varied with capsular type and genetic lineage, but was independent of the clinical origin. Curiously, all *S. agalactiae* strains lacking DNase activity belonged to a particular clonal complex, CC19. So far, no particular clone within the CC19 lineage could be identified in association to DNase production or nonproduction. Genetic clues of *sak_0220* among DNase-deficient CC19 strains may justify in part this phenotype, namely the existence of premature termination codons, generating a non-functional truncated protein. Moreover, our preliminary data obtained by qRT-PCR (chapter VIII) on *S. agalactiae* strain 2603/V/R (CC19) revealed vestigial transcript levels of both *sak_0220* and *sak_0814* during three time-points of growth cycle, which reinforced their DNase (-) phenotype. In addition, our preliminary *in silico* data (data not shown) from strain 2603/V/R on the presence of a DNase stimulus revealed that other genes putatively encoding DNases were not differentially expressed. These findings are in agreement with previous studies suggesting that among the serotype III strains, the ST19 is mostly associated with carriage (Lin et al., 2006; Martins et al., 2007). Despite multiple virulence factors have been involved in *S. agalactiae* pathogenesis, they are not equally distributed among *S. agalactiae* strains, which should influence the invasive success of some genetic lineages, namely the worldwide disseminated hypervirulent ST17 clone. In fact, one important distinguishing feature of the globally-disseminated hypervirulent MIT1 *S. pyogenes* clone compared to less pathogenic *S. pyogenes* strains is the acquisition of a prophage encoding a potent secreted DNase, Sda1 (Aziz et al., 2004).

In conclusion, this Ph.D Thesis provides new insights into *S. agalactiae* molecular epidemiology, antimicrobial susceptibility and virulence profiling either from colonization or symptomatic infection. The transition from colonization to infection in *S. agalactiae* remains to be understood, but certainly it should involve extracellular DNases.
9.4 Future perspectives

In the near future it would be crucial to improve and standardize the clinical diagnosis of *S. agalactiae* during pregnancy to prevent *S. agalactiae* infections. Although the neonatal mortality associated to *S. agalactiae* infection has been declining, long-term outcomes of *S. agalactiae* meningoitis are similar to those reported 25-30 years ago with approximately one-half of children having some degree of impairment. Thus, new genetic methods, such as the PCR should be considered, not for the replacement of the prenatal culture, but as candidates for rapid patient intrapartum *S. agalactiae* testing to determine whether women in labor are colonized with *S. agalactiae*, especially those women with unknown *S. agalactiae* status at time of delivery.

Further studies are necessary to identify the origin and mode of acquisition of the resistance mechanism, along with the clinical effect. Continuous monitoring of invasive *S. agalactiae* disease including antimicrobial susceptibility and serotype determinations will impact plans for prophylaxis regimens and vaccine design. The discovery of *S. agalactiae* strains resistant to vancomycin or displaying reduced susceptibility to penicillin emphasize the importance of continued surveillance of antibiotic resistance among *S. agalactiae* strains and it may be important to establish susceptibility breakpoints for penicillin and vancomycin in *S. agalactiae*. In fact, resistance to both antibiotics mentioned above may be emerging.

It will be important to continue the identification of the functional role of extracellular DNases in *S. agalactiae* pathogenesis, namely in the evasion from the innate immunity during host-pathogen interaction and in the formation of biofilms, which may contributor for host-cells adherence. Correlation of invasiveness of *S. agalactiae* with biofilm formation and DNase activity might be two of the factors that explain the leading role of ST17 strains in neonatal meningitis. In fact, has recently been shown that serotype III ST17 strains are the major producers of biofilm (D’Urzo et al., 2014).

*In silico* and qRT-PCR analyses will be essential to identify other DNase genes including those which may have been introduced by prophages and are likely to have contributed to the fitness and to the virulence of *S. agalactiae*. Also, it will be interesting to confirm the major role of Sak_0814 in infection assays, by using *S. agalactiae* strains from distinct hosts and STs, in particular among the major serotype III lineages, ST17 and ST19.

The development of antimicrobial therapeutics inhibiting nuclease activity, the induction or stabilization of NET formation is of utmost interest to avoid antibiotic overtreatment and resistance. This novel therapeutics should support host immune defense and help to improve the outcome of bacterial infections caused by *S. agalactiae* and other common pathogens.
The recent development of a pioneer molecular imaging approach for the specific, non-invasive detection of *S. aureus* based on the activity of its secreted nuclease (Hernandez *et al.*, 2014), also reflects the importance to continue the research on extracellular DNases, namely in the field of diagnostics.


References


References


References


References


References


129
References


References


References


References


References


Annex
Annex – supplementary data – Chapter VI

**Table A.1. Streptococcus agalactiae-specific qRT-PCR primers for candidate reference genes.**

<table>
<thead>
<tr>
<th>ORF&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Gene identification (Gene description)</th>
<th>Primers</th>
<th>Primer sequence (5' to 3')</th>
<th>Primer location</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gbs0156</td>
<td>RNA polymerase beta unit (rpoB)</td>
<td>rpoB_F1</td>
<td>CACAATTCCATGGACCAACACAA</td>
<td>1415-1437&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rpoB_R1</td>
<td>GGGCTTTTGTCGACAAATTCT</td>
<td>1465-1446&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs0268</td>
<td>Transketolase (tkt)</td>
<td>tkt_F1</td>
<td>CTGCAATGCAACTGACATGTA</td>
<td>1496-1515&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tkt_R1</td>
<td>AAAACAAAGCATGTTGGTGGACAT</td>
<td>1546-1526&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs0518</td>
<td>Glucose kinase (gicK)</td>
<td>gicK_F1</td>
<td>AATTGACCTTCGAGAAGACGA</td>
<td>21-40&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gicK_R1</td>
<td>CCCTCAAGCGTCGAAGACCAACCC</td>
<td>71-51&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs0948</td>
<td>DNA gyrase A (gyrA)</td>
<td>gyrA_F1</td>
<td>TGGGCCATTAGACAGGCGTAG</td>
<td>2235-2254&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gyrA_R1</td>
<td>ACCATGATGTCCTCAATTCATT</td>
<td>2285-2262&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs1806</td>
<td>Glutamine synthetase (glnA)</td>
<td>glnA_F1</td>
<td>CTGGGCGAGGTCGCAATC</td>
<td>941-958&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>glnA_R1</td>
<td>CTCGAGATGCGGAGGACAAATCA</td>
<td>991-968&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs1814</td>
<td>30S ribosomal protein S12 (rpsL)</td>
<td>rpsL_F1</td>
<td>TCCAGGTATCGGCACACACTTG</td>
<td>240-261&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rpsL_R1</td>
<td>CCAGGGGACACCCACACTG</td>
<td>290-269&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs2029</td>
<td>60 KDa chaperonin (groEL)</td>
<td>groEL_F1</td>
<td>CTGTTAAAGCGGCTGAGTTG</td>
<td>818-838&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>groEL_R1</td>
<td>TATCTTCCACATGGCTTACGA</td>
<td>868-846&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs2047</td>
<td>Recombinase A (recA)</td>
<td>recA_F1</td>
<td>ATTCAGCGCGATTTGATTAGTT</td>
<td>440-462&lt;sup&gt;b&lt;/sup&gt;</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>recA_R1</td>
<td>TCAATTCGAGCAGGAGAATCA</td>
<td>506-486&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbs2105</td>
<td>Serine dehydratase (sdiA)</td>
<td>sdiA_F1</td>
<td>TGGTAAAGCGGATGCTTGGT</td>
<td>660-679&lt;sup&gt;b&lt;/sup&gt;</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sdiA_R1</td>
<td>TGGAATTGGCGATTACATACCA</td>
<td>741-720&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>gbsa001</td>
<td>16S ribosomal RNA (16S rRNA)</td>
<td>16S_F1</td>
<td>GCAAAGCGGAAGAACCCTTACC</td>
<td>927-945&lt;sup&gt;b&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td>gbs004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gbs007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gbsr010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gbsr013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gbsr016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gbsr019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Open reading frame (ORF) numbers are based on the NEM316 strain genome annotation (GenBank No. NC_004368.1).

<sup>b</sup> Based on the gene sequence of NEM316 strain.