



**Universidade do Estado do Rio de Janeiro**  
Instituto de Biologia Roberto Alcântara Gomes  
Programa de Pós-Graduação em Biociências

Pamella Silva Lannes Costa

**Sequenciamento, montagem e comparação genômica entre  
linhagens de *Streptococcus agalactiae* ST-17**

Rio de Janeiro

2019

Pamella Silva Lannes Costa

**Sequenciamento, montagem e comparação genômica entre linhagens de  
*Streptococcus agalactiae* ST-17**

Tese apresentada, como requisito parcial para obtenção do título de Doutora, ao Programa de Pós-graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Orientadora: Prof.<sup>a</sup> Dra. Prescilla Emy Nagao Ferreira

Coorientador: Prof. Dr. Artur Luiz da Costa da Silva

Rio de Janeiro

2019

CATALOGAÇÃO NA FONTE  
UERJ/REDE SIRIUS/CBA

C837 Costa, Pamella Silva Lannes.  
Sequenciamento, montagem e comparação genômica entre  
linhagens de *Streptococcus agalactiae* ST-17 / Pamella Silva Lannes  
Costa – 2019.  
161 f.

Orientadora: Prescilla Emy Nagao Ferreira.

Coorientador: Artur Luiz da Costa da Silva.

Tese (doutorado) – Universidade do Estado do Rio de Janeiro,  
Instituto de Biologia Roberto Alcântara Gomes. Programa de Pós-  
graduação em Biociências.

1. *Streptococcus agalactiae* – Aspectos genéticos - Teses. 2. Ilhas  
Genômicas. 3. Proteínas Hemolisinas. 4. Genoma Bacteriano. 5.  
Células endoteliais da veia umbilical humana. I. Ferreira, Prescilla  
Emy Nagao. II. Silva, Artur Luiz da Costa da. III. Universidade do  
Estado do Rio de Janeiro. Instituto de Biologia Roberto Alcântara  
Gomes. IV. Título.

CDU 575.111:579

Bibliotecária: Kalina Silva CRB7/4377

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial  
desta tese, desde que citada a fonte.

---

Assinatura

---

Data

Pamella Silva Lannes Costa

**Sequenciamento, montagem e comparação genômica entre linhagens de  
*Streptococcus agalactiae* ST-17**

Tese apresentada, como requisito parcial para obtenção do título de Doutora, ao Programa de Pós-graduação em Biociências, da Universidade do Estado do Rio de Janeiro.

Aprovada em 11 de julho de 2019.

Coorientador: Prof. Dr. Artur Luiz da Costa da Silva

Universidade Federal do Pará

Banca Examinadora:

---

Prof.<sup>a</sup> Dra. Prescilla Emy Nagao Ferreira (Orientadora)  
Instituto de Biologia Roberto Alcântara Gomes – UERJ

---

Prof.<sup>a</sup> Dra. Lenise Arneiro Teixeira  
Universidade Federal Fluminense

---

Prof.<sup>a</sup> Dra. Verônica Viana Vieira  
Fundação Oswaldo Cruz

---

Prof. Dr. André Luiz Mencialha  
Instituto de Biologia Roberto Alcântara Gomes - UERJ

Rio de Janeiro

2019

## AGRADECIMENTOS

Antes de eu almejar algo, Ele estava destinando todo o meu caminho, agradeço a Deus por me dar o Dom da vida e por me dar forças para continuar mesmo quando encontro obstáculos para desistir. A Nossa Senhora por interceder por mim até aqui!

Agradeço imensamente a minha família, pelo cuidado e apreço a mim dedicado todos esses anos. Em primeiro lugar, a minha amada mãe, Maria de Fátima (Fatinha), pelo amor, carinho, atenção, preocupações, enfim pela sua vida dedicada as filhas, pela ajuda em todos os momentos da vida e por estar ao meu lado sempre. A Priscilla Lannes, uma companheira para todas as horas, uma amiga que Deus colocou no mundo para que eu a chamasse de irmã, ou melhor, “ismã”, obrigada por tudo.

A minha avó-dinda, que sempre me ajuda, obrigada pela sua existência e por sempre está disponível. Aos meus avós que não estão aqui presentes, mas sinto que estão felizes por eu estar ultrapassando mais esta fase e a todos os meus tios (as) e primos (as) que também se alegram e compartilham comigo deste momento, principalmente a minha tia “China” que sempre demonstra sua felicidade pela minha escolha, amo toda a minha família. Não posso deixar de mencionar o meu cunhado, Pablo, que sempre insistiu para a família se reunir e não ficarmos sem nos ver, mesmo eu tendo um computador sempre ao lado para escrever, *rsrsrsrs*.

Peço obrigada ao meu marido, Alex, pela torcida, força, por todo amor a mim dedicado, por embarcar comigo nesta etapa da minha carreira acadêmica e ser o maior incentivador desse momento. Agradeço pela compreensão dos dias e noites em que passei acordada estudando, dos momentos em que não me dediquei inteiramente a nossa relação, á você, meu amor, devo minha alegria diária, meu ânimo. “TE AMO, meu lindo”. Obrigada por permanecer ao meu lado!

Jamais deixaria de agradecer, ao Marcos Temperini (Marquinho) pela minha indicação ao laboratório, pela amizade cultivada a cada dia, pelas caronas, pelas comidas, por toda ajuda e aprendizado tanto no laboratório quanto na vida, afinal somos concunhados. A minha cunhadinha, Alexsandra, pelos jantares e conversas de todos os dias, aos meus sobrinhos, Tabata e Mateus agradeço a confiança. A minha sogra, Josefa, que sabendo do meu empenho estava sempre pronta a me ajudar com as tarefas de casa para que eu pudesse permanecer estudando.

A Prescilla Emy Nagao Ferreira, minha orientadora, pela oportunidade de realizar o meu sonho de seguir uma vida na pesquisa, por sua generosidade e atenção na

minha iniciação científica, no decorrer do meu mestrado e do doutorado. Além da confiança depositada em mim para continuar ao seu lado a ajudando, saiba que jamais esquecerei de todo aprendizado ao longo desses anos.

As eternas meninas do laboratório, Bia e Glenda, agradeço pela amizade construída nesses anos, pelas conversas, conselhos, discussões (Não é, Bia?), pela ajuda, saibam que sou muitíssimo grata e que poderão sempre contar comigo. A Priscilete, pela relação de amizade que se consolidou, pelas viagens (até em casa), pelos papos e apoio nesses anos e por todos os momentos compartilhados. A Jessica M., Day e Melissa obrigada por toda contribuição, pela companhia e pelos nossos encontros, a amizade que continua.

Agradeço a Gabi todo ensinamento, parceria e ajuda, principalmente nas contas, *rsrsrs*. Aos integrantes do LBMFE em geral obrigada por compartilharem o ambiente de trabalho comigo e torná-lo mais ameno nas dificuldades. A tia de todos no departamento, Mariléia, por toda ajuda e pela alegria que irradia o departamento.

Não poderia esquecer o meu coorientador, Artur da Silva, que me recebeu de braços abertos em seu laboratório para que eu pudesse aprender sobre um novo assunto que desenvolvi durante meu doutorado, além de me apresentar a “anjos” que me ajudaram nesse período, Rommel, Vinícius e Rafael Baraúna. O meu agradecimento especial vai para ele, o revisor dessa tese, Rafael, obrigada de coração por toda ajuda e por todo apoio ao longo desses anos, pela paciência em me ensinar e pelo seu despreendimento em passar o conhecimento a quem não fazia ideia de onde estava entrando. Desculpe minhas importunações, saiba que serei muito grata a você por ter conseguido chegar até o final.

À Capes e ao CNPq, que financiaram todo o meu doutorado nesses quatro anos de luta e a pós-graduação que faço parte. Ao PPGB por todo suporte nesse período. E a todos que direta ou indiretamente contribuíram para que essa tese fosse realizada, mas que por descuido esqueci de citar. Meu muito obrigada!

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Há verdadeiramente duas coisas diferentes: saber e crer que se sabe. A ciência consiste em saber; em crer que se sabe reside a ignorância.

*Hipócrates*

## RESUMO

COSTA, Pamella Silva Lannes. **Sequenciamento, montagem e comparação genômica entre linhagens de *Streptococcus agalactiae* ST-17**. 2019. 161 f. Tese (Doutorado em Biociências) - Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

*S. agalactiae* são importantes bactérias patogênicas que causam infecções graves em humanos, especialmente em neonatos. Os métodos sorológicos têm limitações, pois podem não identificar uma amostra devido à falta ou baixa expressão de polissacarídeo capsular nas condições experimentais. Recentemente, vários métodos para tipagem de *S. agalactiae* baseados em genes alvos de PCR no operon *cps* foram publicados como substitutos para sorologia. A implantação da tipagem capsular por PCR multiplex no Laboratório de Biologia Molecular e Fisiologia de Estreptococos (LBMFE) foi realizada com sucesso neste trabalho. Amostras humanas não oncológicas de *S. agalactiae* apresentaram o tipo capsular (TC) III como o mais prevalente (41,86%), enquanto a TC Ia foi frequentemente isolada (44%) em amostras de pacientes com câncer. Em amostras de *S. agalactiae* isoladas de mastite bovina, o TC II (48,28%) foi predominante. O mecanismo pelo qual *S. agalactiae* ST-17 causa mais infecções invasivas do que outros STs não é bem compreendido. Neste estudo, sequenciamos o primeiro genoma de uma amostra ST-17 de *S. agalactiae* (GBS90356) isolada no Brasil. *S. agalactiae* GBS90356 ST-17 do TC III foi isolada de um caso fatal de meningite neonatal. O genoma apresentou um tamanho de 2,03 Mbp e um conteúdo de G + C de 35,2%. *S. agalactiae* GBS90356 possui 706 genes em seu genoma central e um pangenoma aberto com um tamanho de 5.020 genes, sugerindo alta plasticidade genômica. O software GIPSy foi utilizado para identificar 10 PAIs que correspondiam a 17% do tamanho do genoma. O IslandViewer4 corroborou com a previsão de cinco dessas PAIs. As PAIs mostraram importantes genes de virulência para *S. agalactiae* (*neu*, *cps*, *dlt*, *fts*, *cfb*, *lmb*). SignalP detectou 20 proteínas com peptídeos sinal entre as 352 proteínas encontradas nas PAIs, dos quais 60% estavam localizados na SagPAI\_5. As SagPAI\_2 e 5 foram parcialmente detectadas em amostras ST-17 (n=6) estudadas. Além disso, identificamos genes únicos pertencentes a amostra ST-17, que podem estar relacionados à virulência durante as interações patógeno-hospedeiro. Nossos dados forneceram evidências de alta plasticidade genômica em *S. agalactiae*, embora isso não seja comum em bactérias patogênicas com genomas pequenos. As PAIs detectadas compreenderam uma grande porção do genoma, indicando que a maioria dos genes está envolvida no processo de patogenicidade do *S. agalactiae*. A inibição da atividade β-hemolítica do GBS90356 reduziu os perfis de aderência e invasão, sugerindo que a β-hemolisina foi necessária no processo de disseminação da amostra GBS90356 em HUVEC. Alterações morfológicas e sinais de apoptose foram verificados durante a interação GBS90356-HUVEC. Todos esses achados sugerem que a virulência de *S. agalactiae* pode estar relacionada à sua alta plasticidade e à presença de elementos genéticos móveis.

Palavras-chave: Ilhas de patogenicidade. *S. agalactiae*. β-hemolisina. HUVEC. Sequenciamento.



## ABSTRACT

COSTA, Pamella Silva Lannes. **Sequencing, assembling and genomic comparison between strains of *Streptococcus agalactiae* ST-17**. 2019. 161 f. Tese (Doutorado em Biociências) - Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

*S. agalactiae* are important pathogenic bacteria that cause severe infections in humans, especially neonates. Serological methods have limitations, as they may fail to type an isolate due to lack of or low expression of capsular polysaccharide under the experimental conditions. Recently, several methods for typing of *S. agalactiae* based on PCR targeting genes in the *cps* operon has been published as alternatives for serology. The implantation of *S. agalactiae* capsular typing by multiplex PCR in the Laboratory of Molecular Biology and Physiology of Streptococci (LBMFE) was successful. Non-oncological human *S. agalactiae* strains presented capsular type (CT) III as the most prevalent (41.86%), while CTIa was frequently isolated (44%) in strains from cancer patients. In *S. agalactiae* strains recovered from mastitis bovine, CT II (48.28%) was predominant. The mechanism by which *S. agalactiae* ST-17 causes invasive infections than other STs is not well understood. In this study, we sequenced the first genome of a *S. agalactiae* ST-17 strain (GBS90356) isolated in Brazil. *S. agalactiae* GBS90356 ST-17 from CTIII was isolated from a fatal neonatal case of meningitis. The genome presented a size of 2.03 Mbp and a G+C content of 35.2%. *S. agalactiae* has 706 genes and an open pan-genome with a size of 5.020 genes, suggesting a high genomic plasticity of *S. agalactiae*. GIPSY software was used to identify 10 pathogenicity islands (PAIs) which corresponded to 17% of the genome size. IslandViewer4 corroborated with the prediction of five of these PAIs. The PAIs showed important virulence genes for *S. agalactiae* e.g. *neu*, *cps*, *dlt*, *fb*s, *cfb*, *lmb*. SignalP detected 20 proteins with signal peptides among the 352 proteins found in PAIs, which 60% were located in the SagPAI\_5. The SagPAI\_2 and 5 were partially detected in all ST-17 strains (n=6) studied. Moreover, we identified unique genes belonging to ST-17 strains, which can be related with the virulence during host-pathogen interactions. We provide evidence of high genomic plasticity in not usual pathogenic bacteria as *S. agalactiae* with small genome. PAIs comprise a large portion of the genome, indicating that the majority of virulence genes are involved in *S. agalactiae* pathogenicity process. Activity inhibition of  $\beta$ -hemolytic of GBS90356 reduced the adherence and invasion profiles, suggesting that  $\beta$ -hemolysin was required in the process of *S. agalactiae* dissemination in HUVEC. Morphological changes and signals of apoptosis were validated during GBS90356-HUVEC interaction. All these findings suggest that the virulence of *S. agalactiae* could be related to high plasticity and the presence of the mobile genetic elements.

Keywords: Pathogenicity islands. *S. agalactiae*.  $\beta$ -hemolysin. HUVEC. Sequencing.

## LISTA DE ILUSTRAÇÕES

Figura 1 -	Microscopia Eletrônica de Varredura (MEV) de <i>Streptococcus agalactiae</i> .....	15
Figura 2 -	Estágios da infecção neonatal na infecção precoce - transmissão ascendente.....	21
Figura 3 -	Quantidade total de genomas completos e <i>drafts</i> depositados em GOLD, onde o eixo X representa o ano e o eixo Y a quantidade de genomas .....	41
Figura 4 -	Compartimentos do pangenoma.....	45
Figura 5 -	Eletroforese dos produtos da tipagem capsular de <i>S. agalactiae</i> ..	64
Figura 6 -	Árvore filogenômica de 109 amostras de <i>S. agalactiae</i> .....	66
Figura 7 -	Predição do pangenoma de <i>S. agalactiae</i> .....	68
Figura 8 -	Mapa circular dos genomas para comparar a identidade dos nucleotídeos entre as amostras de <i>S. agalactiae</i> e determinar a conservação das PAIs.....	71
Figura 9 -	Análise de sintenia das PAIs 2 (A) e 5 (B) de <i>S. agalactiae</i> .....	73
Gráfico 1 -	Interação da amostra GBS90356 com HUVEC.....	78
Figura 10 -	Análise morfológica da interação da amostra GBS90356 com HUVEC.....	80

## LISTA DE TABELAS

Tabela 1 - <i>Primers</i> para tipagem capsular de <i>S. agalactiae</i> .....	50
Tabela 2 - Tipagem capsular de amostras humanas de <i>S. agalactiae</i> isoladas de pacientes não oncológicos por PCR multiplex.....	61
Tabela 3 - Tipagem capsular de amostras humanas de <i>S. agalactiae</i> isoladas de pacientes oncológicos por PCR multiplex.....	62
Tabela 4 - Tipagem capsular de amostras bovinas de <i>S. agalactiae</i> por PCR multiplex .....	63
Tabela 5 - Métricas do sequenciamento na plataforma Illumina HiSeq 2500.....	65
Tabela 6 - Genes das PAIs 2 e 5 encontrados em amostras de <i>S. agalactiae</i> ST-17 .....	74
Tabela 7 - Análise funcional das PAIs de <i>S. agalactiae</i> .....	76

## LISTA DE ABREVIATURAS E SIGLAS

AAP	<i>American Academy of Pediatrics</i>
ACOG	<i>American College of Obstetricians and Gynecologists</i>
ALT	Ácido lipoteicóico
BHI	Infusão cérebro coração
CAMP	Christie, Atkins e Miunch – Peterson
CC	Complexo clonal
CDC	Centro de Controle e Prevenção de Doenças
CDS	<i>Coding sequencing</i>
CE	Células endoteliais
CP	Cápsula polissacarídica
CPS	Polissacarídeos capsulares
<i>cps</i>	Gene referente a capsula polissacarídica
DNA	Ácido desoxirribonucléico
DO	Densidade ótica
EDTA	Ácido etileno amino acético
EGB	Estreptococos do grupo B
ELISA	Enzyme-linked immunosorbent assay
EUA	Estados Unidos da América
FAPESPA	Fundação de Amparo à Pesquisa do Estado do Pará
FAPEMIG	Fundação de Amparo à Pesquisa no Estado de Minas Gerais
FBS A	Fibrinogenbindingprotein A, proteína de ligação a fibrinogênio A
FBS B	Fibrinogenbindingprotein B, proteína de ligação a fibrinogênio B
FBS C	Fibrinogenbindingprotein B, proteína de ligação a fibrinogênio C
G+C	Conteúdo guanina + citosina
GEIs	Ilhas genômicas
GenBank	Banco genético
GO	Gene orthology
hBMEC	Célula endotelial da microvasculatura cerebral humana
HUVEC	Célula endotelial da veia umbilical humana
ICAM-1	Molécula de adesão intercelular-1

Ig	Imunoglobulina
Lmb	<i>Laminin binding protein</i> ; proteína ligada a laminina
MEC	Matriz extracelular
MEV	Microscopia eletrônica de varredura
MLEE	<i>Multi Locus Enzyme Electrophoresis</i> ; Eletroforese de isoenzima
MLST	<i>Multi Locus Sequence Typing</i> ; tipagem de sequência multilocus
NCBI	National Center for Biotechnology Information
NGS	<i>Next generation sequencing</i>
NT	Não tipáveis
ORF'S	<i>Open read frames</i>
pb	Pares de base
PAI	Ilha de patogenicidade
PBS	<i>Phosphate buffer solution</i> ; solução de tampão fosfato
PCR	<i>Polymerase Chain Reaction</i> ; reação da cadeia em polymerase
PCA	Peptídeos catiônicos antimicrobianos
pH	Potencial hidrogeniônico
RGMG	Rede Genoma de Minas Gerais
rRNA	RNA's ribossomais
tRNA	RNA's transportadores
RPGP	Rede Paraense de Genômica e Proteômica
ST	Sequência tipo
SDS	Duodecil sulfato de sódio
SFB	Soro fetal bovino
Srr	Glicoproteína rica em serina
TBE	Tris/ácido bórico/EDTA
TC	Tipo capsular
TE	Tris/EDTA
THE	Transferência Horizontal gênica
TLR-2-6	Receptor Tool Like-2 e 6
UFC/mL	Unidades formadoras de colônia por mililitro
UPGMA	<i>Unweighted Pair Group Method with Arithmetic Mean</i>

## SUMÁRIO

	<b>INTRODUÇÃO</b> .....	<b>14</b>
<b>1</b>	<b>REVISÃO DA LITERATURA</b> .....	<b>15</b>
1.1	<b><i>Streptococcus agalactiae</i>: características gerais</b> .....	<b>15</b>
1.2	<b>Epidemiologia</b> .....	<b>17</b>
1.3	<b>Infecções ocasionadas por <i>S. agalactiae</i></b> .....	<b>19</b>
1.4	<b>Fatores de virulência</b> .....	<b>23</b>
1.5	<b>Endotélio e a interação com <i>S. agalactiae</i></b> .....	<b>31</b>
1.6	<b>Diversidade genética de <i>S. agalactiae</i></b> .....	<b>34</b>
1.7	<b>Aspectos da genômica e do sequenciamento</b> .....	<b>38</b>
1.8	<b>Montagem do genômica</b> .....	<b>42</b>
1.9	<b>Genômica comparativa</b> .....	<b>44</b>
<b>2</b>	<b>OBJETIVOS</b> .....	<b>47</b>
2.1	<b>Objetivo geral</b> .....	<b>47</b>
2.2	<b>Objetivos específicos</b> .....	<b>47</b>
<b>3</b>	<b>METODOLOGIA</b> .....	<b>49</b>
3.1	<b>Cultura bacteriana</b> .....	<b>49</b>
3.2	<b>Oligonucleotídeos (<i>Primers</i>)</b> .....	<b>49</b>
3.2.1	<u>Para tipagem capsular</u> .....	<b>49</b>
3.3	<b>Extração de DNA</b> .....	<b>50</b>
3.3.1	<u>Para tipagem capsular</u> .....	<b>50</b>
3.3.2	<u>Para Sequenciamento de alto desempenho</u> .....	<b>51</b>
3.4	<b>PCR-multiplex na tipagem capsular de <i>S. agalactiae</i></b> .....	<b>51</b>
3.4.1	<u>PCR convencional na tipagem capsular de <i>S. agalactiae</i></u> .....	<b>52</b>
3.5	<b>Eletroforese horizontal em gel de agarose</b> .....	<b>52</b>
3.6	<b>Sequenciamento do genoma da amostra GBS90356</b> .....	<b>53</b>
3.7	<b>Avaliação da qualidade das <i>reads</i>, montagem e anotação do genoma da amostra GBS90356</b> .....	<b>53</b>
3.8	<b>Obtenção das sequências dos genomas completos de <i>S. agalactiae</i> em banco de dados</b> .....	<b>54</b>
3.9	<b>Ferramentas de bioinformática</b> .....	<b>54</b>

3.9.1	<u>Análise filogenômica e cálculo do Pangenoma de <i>Streptococcus agalactiae</i></u> .....	54
3.9.2	<u>Predição das Ilhas de Patogenicidade (PAI) e profagos</u> .....	55
3.9.3	<u>Sintenia das PAIs e comparação nucleotídica</u> .....	55
3.9.4	<u>Enriquecimento funcional dos genes das PAIs e predição do secretoma</u> .....	56
3.10	<b>Cultura de células endoteliais da veia umbilical humana (HUVEC)</b> ..	56
3.11	<b>Ensaio de interação entre <i>S. agalactiae</i> e HUVEC</b> .....	57
3.12	<b><i>Time-Lapse</i> da interação celular – HUVEC X <i>S. agalactiae</i></b> .....	59
4	<b>RESULTADOS</b> .....	60
4.1	<b>Tipagem capsular de <i>Streptococcus agalactiae</i> por PCR multiplex</b> .....	60
4.2	<b>Características genômicas do <i>S. agalactiae</i> GBS90356</b> .....	64
4.3	<b>Análise comparativa de <i>S. agalactiae</i>: GBS90356 ST-17 no Brasil</b> ....	65
4.3.1	<u>Análise filogenômica e predição de profagos</u> .....	65
4.3.2	<u>Avaliação do pangenoma de <i>S. agalactiae</i></u> .....	67
4.3.3	<u>Predição e análise das PAIs de <i>S. agalactiae</i></u> .....	69
4.3.4	<u>Análise de sintenia das PAIs de <i>S. agalactiae</i> e predição do secretoma</u>	72
4.3.5	<u>Análise funcional dos genes das PAIs</u> .....	75
4.4	<b>Papel da hemolisina na interação de <i>S. agalactiae</i> GBS90356 com células endoteliais (HUVEC)</b> .....	77
4.5	<b>Análise morfológica do perfil de interação celular (HUVEC X <i>S. agalactiae</i> GBS90356) por <i>time-lapse</i> na <i>Biostation</i><sup>®</sup></b> .....	79
5	<b>DISCUSSÃO</b> .....	81
	<b>CONCLUSÕES</b> .....	92
	<b>REFERÊNCIAS</b> .....	94
	<b>APÊNDICE A</b> - Informações dos 109 genomas de <i>S. agalactiae</i> utilizadas nesse estudo.....	118
	<b>APÊNDICE B</b> - Lista das CDSs detectadas dentro das PAIs do genoma da <i>S. agalactiae</i> GBS90356. Ilhas de patogenicidade são abreviadas com <u>SagPAI</u> .....	123
	<b>ANEXO</b> – Formato do 1º artigo submetido .....	142

## INTRODUÇÃO

O Laboratório de Biologia Molecular e Fisiologia de Estreptococos (LBMFE), coordenado pela Dra. Prescilla Emy Nagao Ferreira no Departamento de Biologia Celular da Universidade do Estado do Rio de Janeiro (UERJ) vem atuando no estudo dos mecanismos de virulência de amostras de *Streptococcus agalactiae*, bem como realizando a caracterização fenotípica e genotípica de amostras bacterianas isoladas de materiais clínicos diversos de origem humana e bovina, como foco de todas as linhas de pesquisa.

O *S. agalactiae* ocupa o terceiro lugar entre os patógenos causadores de meningites bacterianas em neonatos, apresentando danos neurológicos permanentes em 25-50% dos sobreviventes. Além disso, o *S. agalactiae* é reconhecido como agente etiológico de mastite bovina clínica e subclínica, sendo atualmente um dos maiores problemas para a indústria de laticínios. A circulação de amostras epidêmicas e/ou invasoras reforça a necessidade de análises relativas às propriedades biológicas e genotípicas deste microrganismo. A diversidade de nichos hospedeiros (humanos, bovinos) colonizados pelo *S. agalactiae* indica a eficiente capacidade adaptativa deste microrganismo, bem como a apropriada expressão de fatores de virulência. A crescente emergência de cepas resistentes à clindamicina e eritromicina reforça a importância do *S. agalactiae* como problema de Saúde Pública, bem como para indústria alimentícia.

O número exíguo de grupos especialistas em interação de *S. agalactiae* com células endoteliais e epiteliais respiratórias no Rio de Janeiro e no país faz com que análises relativas aos aspectos de virulência e patogênese das infecções causadas por *S. agalactiae* sejam de grande relevância.

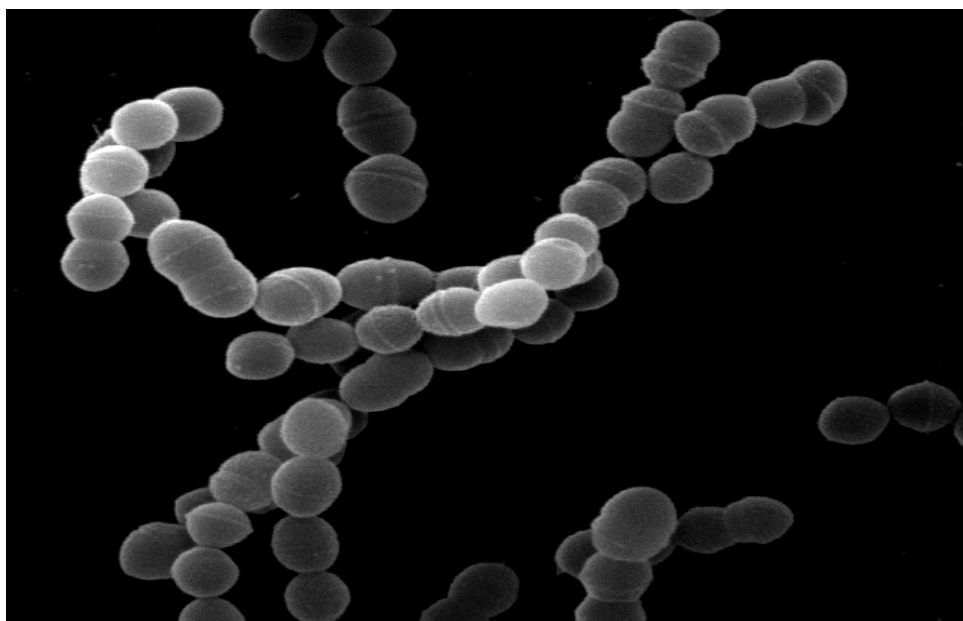


## 1 REVISÃO DA LITERATURA

### 1.1 *Streptococcus agalactiae*: características gerais

O *Streptococcus agalactiae*, denominado também de estreptococos do grupo B (EGB) pertence ao reino monera, família *Streptococcaceae* apresentando morfologia de cocos agrupados em cadeias (Figura 1) e envoltos por uma cápsula polissacarídea. São microrganismos  $\beta$ -hemolíticos, catalase negativa, Gram, CAMP e hipurato positivos (NIZET; RUBENS, 2002; SHABAYEK; SPELLERBERG, 2018).

Figura 1 - Microscopia Eletrônica de Varredura de *Streptococcus agalactiae*



Fonte: MIRANDA, 2018.

Em 1887, o *S. agalactiae* foi isolado pela primeira vez no leite bovino, sendo descrito por Lehmann e Neumann em 1896 e reconhecido em 1920, como agente etiológico da mastite bovina, Nocard e Mollereau (1887 *apud* BISCHARAT *et al.*, 2004). Somente na década de 30, o *S. agalactiae* foi detectado em secreções vaginais de pacientes assintomáticos e associado com sepse, pneumonia puerperal e doenças em neonatos (LANCEFELD, 1934; RAJAGOPAL *et al.*, 2009). O primeiro

caso fatal de infecção por *S. agalactiae* em humanos foi relatado por Fry em 1938. Desde então, o *S. agalactiae* vem emergindo como patógeno responsável por morbidade e mortalidade em humanos (DUTRA *et al.*, 2004). Este microrganismo foi também descrito colonizando outros animais como cachorros, cavalos, peixes, cobaias, golfinhos entre outros (VANDAMME, 1997).

O *S. agalactiae* possui antígenos carboidratos específicos (polímeros de glicose, galactose, ramnose, N-acetilglicosamina e ácido siálico) que permitiram a classificação em tipos sorológicos (LANCEFIELD, 1934; KONG *et al.*, 2008). Atualmente, eles são classificados em dez tipos capsulares - Ia, Ib, II, III, IV, V, VI, VII, VIII e IX - de acordo com as variações no número de unidades repetidas dos açúcares capsulares e o tipo de ligação glicosídica (CIESLEWICZ *et al.*, 2005; SLOTVED *et al.*, 2007; SPELLERBERG, 2000; KONG *et al.*, 2008). O TC IX, proposto por Slotved *et al.* (2007) parece ter evoluído a partir da mutação e/ou recombinação entre os TC V e VII (BERTI *et al.*, 2014).

Este patógeno é reconhecido como membro da microbiota anfibiótica normal do ser humano, podendo ser encontrado nas vias aéreas superiores, trato intestinal, vaginal e tubo digestivo de cerca de 30-40% de indivíduos saudáveis (BEITUNE *et al.*, 2005; MARTINEZ *et al.*, 2000; KORIR *et al.*, 2016; OSTER *et al.*, 2014). Apesar de ser considerado um organismo oportunista comensal, o *S. agalactiae* pode causar infecções invasivas, como a endometrite, bacteriúria, amnionite, osteomielite, endocardite, pneumonia, meningite e sepse em neonatos, adultos, idosos e indivíduos imunocomprometidos como diabéticos, alcoólatras e pacientes com câncer (MARTINS, 2010; PIMENTEL *et al.*, 2016).

Durante os últimos anos, clones hipervirulentos e reguladores que controlam a expressão de genes dos fatores virulência de *S. agalactiae* vem sendo caracterizados (SHABAYEK; SPELLERBERG, 2018). No entanto, a mudança de comensal para patógeno invasivo ainda permanece pouco conhecida. A análise molecular deste patógeno pode contribuir para o melhor entendimento da epidemiologia e patogênese do *S. agalactiae*.

## 1.2 Epidemiologia

A distribuição dos tipos capsulares (TC) de *S. agalactiae* associados a doenças no mundo tem apresentado diferenças relacionadas a região geográfica e origem étnica (LE DOARE; HEATH, 2013). De forma geral, os TC Ia, Ib, II, III, IV e V estão associados a doenças em humanos e animais (MARTINEZ *et al*, 2000), sendo o tipo III responsável por 30-70% dos casos de sepse e meningite neonatal; os tipos Ia, Ib, III e IV detectados em gestantes assintomáticas e o TC V em infecções em idosos e adultos imunocomprometidos (FIOLO *et al*, 2012; HO *et al*, 2013).

A colonização de *S. agalactiae* em gestantes tem sido verificada de forma variada nos diferentes continentes. Na Ásia foram detectadas gestantes com colonização variando de 4,7% (Índia) a 21% (Japão) (AILA *et al.*, 2010; MATSUBARA *et al.*, 2013). Na Europa esse índice variou de 15% na Lituânia a 36% na Dinamarca (CAPANNA *et al.*, 2013; MIYATA *et al.*, 2012). Enquanto no continente africano, na Etiópia, a prevalência da colonização materna por *S. agalactiae* foi de 19% (CAPANNA *et al.*, 2013). Na América verificou-se 23,96% nos Estados Unidos e 25,4% no Brasil (FIOLO *et al.*, 2012; ROCCHETTI *et al.*, 2011). A taxa de transmissão ao neonato foi verificada em 60% e 21% no Irã e Japão, respectivamente (MATSUBARA *et al.*, 2013; NAMAVAR *et al.*, 2008). Enquanto na Europa, a incidência de infecções em neonatos foi de 6% em Portugal (AREAL *et al.*, 2010; SLOTVED; HOFFMANN, 2017).

No continente asiático os TC Ib, III, IV, VI, VII e VIII foram prevalentes (RIVERA *et al.*, 2015), especificamente no Japão com predomínio dos TC III e Ib em isolados clínicos neonatais (WEISNER *et al.*, 2004) e TCs VI e VIII em gestantes (LIU *et al.*, 2018; MENGIST; KANNAN; ABDISSA, 2016). Contudo, nos Emirados Árabes e no Egito os TC prevalentes foram IV e V, respectivamente (AMIN; ABDULRAZZAQ; UDUMAN, 2002; SHABAYEK *et al.*, 2014). O TC III foi o mais frequente entre mulheres grávidas e neonatos na China, Tailândia e Austrália (EDMOND *et al.*, 2012; KANNIKA *et al.*, 2017; LIU *et al.*, 2018; WANG *et al.*, 2013).

Estudos na África do Sul mostraram que 551 mulheres colonizadas por *S. agalactiae* apresentaram isolados pertencentes aos TC III (37,3%), Ia (30,1%), II (11,3%), V (10,2%), Ib (6,7%) e IV (3,7%) (DOARE; HEATH, 2013) e o TC III foi o mais invasivo em recém-nascidos com até uma semana de vida (57,7%) e em

crianças com até 90 dias (84,3%) (MADZIVHANDILA *et al.*, 2011). Na Europa os TC Ia, II e III englobam 85% dos isolados clínicos neonatais (ZALEZNIK *et al.*, 2000), onde amostras dos TC IV, VI, VII e VIII foram raramente isoladas nestas regiões (WEISNER *et al.*, 2004). O TC IX foi primeiramente isolado na Dinamarca (SLOTVED *et al.*, 2007). Estudo realizado na Inglaterra relatou que dos 4878 isolados de *S. agalactiae* analisados 0,3% pertenciam ao TC IX (LAMAGNI *et al.*, 2013). Dados da Austrália demonstraram a presença de 0,5% do TC IX das 814 mulheres grávidas analisadas (FURFARO *et al.*, 2019).

Estudos na América demonstram que os TC Ia e III se apresentaram como mais comuns (LIU *et al.*, 2018), no Peru e na Argentina foi evidenciado que os TC Ia (32%), III (20%), II (15%) e V(12%) foram frequentemente isolados em gestantes (COLLINS *et al.*, 1998; PEREZ *et al.*, 2004). Dados dos Estados Unidos também demonstraram que os TC Ia, II e III foram os mais prevalentes entre isolados clínicos neonatais (KONG *et al.*, 2008; ZALEZNIK *et al.*, 2000), assim como na República Dominicana e Panamá os TC mais frequentes em neonatos foram Ia e III, respectivamente (RIVERA *et al.*, 2015). Abortos e nascimentos prematuros, em sua grande maioria, não são investigados, dificultando a aquisição de dados epidemiológicos a nível mundial. Além disso, alguns países não exigem o rastreio bacteriológico do *S. agalactiae* em seu protocolo nacional obrigatório, incluindo o Brasil (SLOTVED; HOFFMANN, 2017). Desta forma, limitados são os estudos sobre a distribuição dos TC, caracterização e epidemiologia do *S. agalactiae* nos países em desenvolvimento (DOARE *et al.*, 2013; JOHRI *et al.*, 2013), onde são encontradas as maiores taxas de mortalidade por infecções neonatais (sepsis e pneumonia) (LIU *et al.*, 2015).

No Brasil, Palmeiro *et al.* (2010) demonstraram que das 168 amostras brasileiras de *S. agalactiae* adquiridas de pacientes acometidos por bacteremia, artrite, meningite, osteomielite, infecção puerperal, infecção na pele e infecção do trato urinário (isolados reto/vaginal e de trato urinário) no Hospital da Universidade Federal do Paraná (Curitiba), no período de 2006 a 2008, o TC Ia foi predominante (38,1%), seguido dos TC II (16,1%), IV (13,1%) e V (12,5%). Dados do nosso grupo revelaram que a maioria das amostras isoladas da cidade do Rio de Janeiro-RJ foram dos TC III (38,46%), V (25,64%), Ia (10,26%), II (7,69%), Ib (5,13%) e 12,8% de isolados não tipáveis (SOARES *et al.*, 2013). Contudo, estudo posterior na cidade do Rio de Janeiro, demonstrou a prevalência de colonização dos TC Ia (37,3%) e II

(19,9%), seguidos dos TC Ib (11,1%), V (9,1%), III (6,8%) e IV (3,5%) em mulheres grávidas (BOTELHO *et al.*, 2018).

Apesar dos dados relatados anteriormente, no Brasil ainda são escassos os estudos sobre a distribuição e/ou prevalência dos TC de *S. agalactiae* na população, de acordo com a área geográfica (SOARES *et al.*, 2013; TEIXEIRA *et al.*, 2008). Além disso, não há estimativas nacionais para criação de legislação específica para introdução de normas para diagnóstico e tratamento de recém-nascidos infectados por este microrganismo. Devido ao desprovimento de dados brasileiros na literatura sobre a ocorrência deste fenômeno, especificamente de dados regionais, torna-se difícil a comparação com dados de outras regiões, bem como a elaboração de estratégias profiláticas para o controle do *S. agalactiae* (HEATH, 2016). Neste contexto, a caracterização molecular de amostras de *S. agalactiae* é relevante podendo contribuir para a epidemiologia e possíveis tratamentos das infecções causadas pelo patógeno.

### 1.3 Infecções ocasionadas por *S. agalactiae*

*S. agalactiae* está associado à mastite em bovinos, sendo capaz de colonizar as glândulas mamárias de ruminantes e de sobreviver por longos períodos de tempo, de acordo com Nocard e Mollereau (1887 *apud* BISCHARAT *et al.*, 2004). A mastite é um processo inflamatório das glândulas mamárias, podendo ser classificada como clínica ou subclínica (MARTINS *et al.*, 2010; MIRANDA *et al.*, 2018). Independente da classificação, a mastite bovina apresenta riscos à saúde do homem e bezerros, pois o leite pode ser um transmissor de microrganismos patogênicos e enterotoxinas (ZAFALON *et al.*, 2009).

A mastite é uma doença economicamente importante, pois causa prejuízos aos pecuaristas e indústrias de laticínios (AZEVEDO *et al.*, 1997; CARVALHO-COSTA *et al.*, 2017; KEEFE, 2012), devido as alterações do leite como as modificações do pH, diminuição no extrato seco total (caseína, gordura, lactose), aumento de minerais como sódio e cloro, diminuição do cálcio e fósforo, aumento de algumas enzimas (catalase, fosfatase ácida, arilesterase, entre outras) e aumento de células somáticas, especialmente os leucócitos (SILVA, 2014). Esta é uma

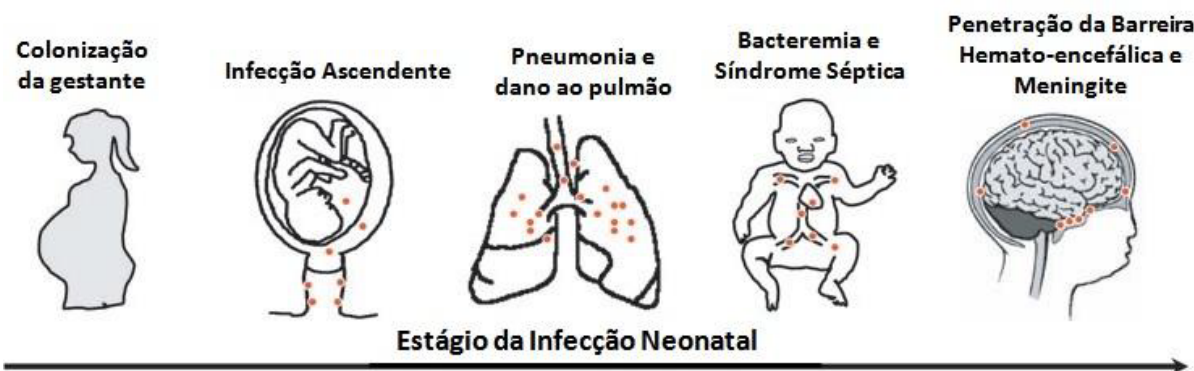
bactéria altamente contagiosa e adaptada à glândula mamária de bovino (KEEFE, 2012), podendo também colonizar e causar infecções invasivas em outros animais como cachorros, cavalos, porcos guiné (BISHOP *et al.*, 2007; VANDAMME, 1997), sapos (AMBORSKI *et al.*, 1983), golfinhos (EVANS *et al.*, 2008), macacos (GARCIA *et al.*, 2008) e peixes, como tilápias (DANGWETNGAM *et al.*, 2016; MIAN *et al.*, 2009; WANG *et al.*, 2017).

Em humanos o *S. agalactiae* pode causar desde quadros clínicos leves (infecção vaginal e urinária) até infecções graves como celulite e fascite em neonatos e em adultos (FARLEY *et al.*, 2001; SCHUCHAT, 1999). O *S. agalactiae* permanece como a principal causa de sepse e meningite neonatal e está associado com significativa morbidade e mortalidade (LIBSTER *et al.*, 2012; MITHAL *et al.*, 2017). Cerca de 20-30% das mulheres são portadoras de *S. agalactiae* de forma permanente, intermitente ou transitória (VERANI; SCHRANG, 2010). Durante a gravidez, muitos neonatos, particularmente os prematuros, nascidos de mães colonizadas pelos *S. agalactiae* são infectados ainda no útero, podendo nascer com baixo peso corporal ou natimorto (YANCEY *et al.*, 1994; AREAL *et al.*, 2010). Nos homens pode se multiplicar nos órgãos reprodutivos masculinos como, uretra e próstata, podendo levar a bacteremia (FREY *et al.*, 2011). O contato sexual parece não influenciar significativamente na aquisição do *S. agalactiae*, sendo a transmissão vertical seu principal meio de transferência (HONIG *et al.*, 2002).

Baker e Edwards (2001) demonstraram que a patogênese da infecção em neonatos tem início através da colonização assintomática do trato genital feminino, geralmente nos três primeiros meses de gravidez. O risco de transmissão deste patógeno em neonatos é de aproximadamente 70% (BODASZEWSKA-LUBAS *et al.*, 2013). Neste caso pode ocorrer a infecção ascendente, onde a bactéria atravessa a membrana placentária e a infecção ocorre ainda no útero (DORAN *et al.*, 2004) (Figura 2) ou através da aspiração do líquido amniótico ou fluidos vaginais contaminados com *S. agalactiae* de mãe portadora (transmissão vertical), seguida da translocação bacteriana através do epitélio respiratório do neonato podendo levar a infecção sistêmica, com frequentes manifestações respiratórias como, por exemplo, pneumonia (ESCOLANO *et al.*, 2014; SHABAYEK; SPELLERBERG, 2018). Sendo assim, as infecções de início precoce que acontecem na primeira semana de vida têm como principais manifestações clínicas a pneumonia e a sepse

(DORAN; NIZET, 2004; SCHUCHAT, 1998), atingindo 0,5-3% em cada 1000 crianças nascidas vivas (KAAMBWA *et al.*, 2010).

Figura 2 - Estágios da infecção neonatal na infecção precoce - transmissão ascendente



Fonte: adaptado de DORAN *et al.*, 2004.

Estudos realizados nos Estados Unidos demonstraram que 10-36% das mulheres grávidas portadoras de *S. agalactiae* transmitiram o patógeno por transmissão vertical (PRESENTATION *et al.*, 2002; ZALEZNIK *et al.*, 2000). Na Índia a taxa de transmissão vertical foi de 53-56% (PATIL *et al.*, 2013). Desta forma, mais de 95% dos casos de infecções neonatais estão relacionados a vários fatores de risco como: grau de colonização materna, tempo entre a ruptura da placenta e o parto, febre materna acima de 38°C, infecção urinária e prematuridade (EDWARD; BAKER, 2001; GIBBS; SCHRAG; SCHUCHAT, 2004; HICKMAN *et al.*, 1999; LAW *et al.*, 2005). A transmissão vertical, aquisição nasocomial e a prematuridade são reconhecidas como os fatores de risco das infecções causadas pelo *S. agalactiae* (MELIN; EFSTRATIOU, 2013; RAJAGOPAL, 2009).

A infecção de início tardio ocorre em neonatos entre a primeira semana e os 2-3 meses de idade, sendo a sepse e a meningite, as manifestações clínicas mais frequentes que atingem cerca de 50% dos casos (EDWARDS; BAKER, 2001; ESCOLANO *et al.*, 2014). Estima-se que a principal fonte envolvida neste processo seja resultado de infecção hospitalar ou adquirida na comunidade. O *S. agalactiae* está entre os principais patógenos responsáveis pela meningite neonatal bacteriana, englobando 70-80% dos casos (DONDEERS *et al.*, 2016; HEAT *et al.*, 2003). O TC III é responsável pela maioria (~70%) das infecções tardias (BODASZEWSKA-LUBAS

*et al.*, 2013; ESCOLANO *et al.*, 2014; PRESENTATION *et al.*, 2002; VERANI *et al.*, 2010).

Em relação aos adultos, o *S. agalactiae* está associado com infecções cutâneas, pneumonia, infecção urinária, meningite, endocardite, bacteremia, entre outras (LANOTTE *et al.*, 2013). Geralmente a infecção por *S. agalactiae* em adultos acomete idosos e pessoas com baixa imunidade e/ou doenças concomitantes, como diabetes *mellitus*, doenças hepáticas e câncer (MUÑOZ *et al.*, 1997; PALMEIRO *et al.*, 2010). Poucos estudos estão disponíveis na literatura descrevendo infecções por *S. agalactiae* em adultos imunocomprometidos, principalmente quando se refere a portadores de neoplasias. Um panorama das infecções deste patógeno em pacientes com câncer no Rio de Janeiro (2010-2014) demonstrou um número significativo de casos de infecções invasivas em adultos de 20-59 anos (36%) e idosos com idade entre 60 a 79 anos (46%), onde 34,2% foram a óbito (PIMENTEL *et al.*, 2016).

A infecção bacteriana tem sido a principal causa de morbidade e mortalidade em pacientes com câncer (VELASCO *et al.*, 2000). As infecções adquiridas em hospitais são responsáveis por mais de 200.000 pacientes por ano nos Estados Unidos (DIEDRICK *et al.*, 2010). De acordo com dados publicados, as infecções por *S. agalactiae* são muito invasivas e possuem elevada taxa de mortalidade (SLOTVED; HOFFMANN, 2017). Desde a introdução dos programas de profilaxia em mulheres grávidas (BALLARD *et al.*, 2016; SLOTVED; HOFFMANN, 2017), as infecções por *S. agalactiae* de início precoce em neonatos têm sido reduzidas e a incidência da doença provocada por esse patógeno tem diminuído (HEATH, 2016). Em contraste, a infecção de início tardio ainda é um problema importante no mundo em desenvolvimento e presumivelmente um problema subestimado (HEATH, 2016).

Nos últimos anos, países desenvolvidos têm demonstrado interesse crescente na incidência de infecções invasivas de *S. agalactiae* com políticas de prevenção implantadas em alguns países houve uma redução significativa da incidência da doença neonatal precoce pelo *S. agalactiae* (BALLARD *et al.*, 2016; COSTA *et al.*, 2008; SLOTVED; HOFFMANN, 2017). Contudo, as autoridades em saúde pública do Brasil não aderiram às recomendações sugeridas pelo *Centers of Disease Control and Prevention* (CDC), *American College of Obstetricians and Gynecologists* (ACOG) e a *American Academy of Pediatrics* (AAP) para prevenção das infecções de início precoce por *S. agalactiae* (BARROS; SOUZA; BAPTISTA, 2016). Desta



forma, não há dados nacionais referindo-se as taxas de infecção por *S. agalactiae* durante a gravidez e nem a incidência da infecção neonatal de início precoce em nosso país, tornando-se importante estudo sobre a patogênese do *S. agalactiae*.

#### 1.4 Fatores de virulência

*S. agalactiae* apresentam uma série de estratégias para infectar o hospedeiro, através da secreção de moléculas no ambiente extracelular ou expressando estruturas em sua superfície (BARON *et al.*, 2005). Estes fatores de virulência estão envolvidos principalmente na aderência, invasão e sobrevivência do patógeno na célula hospedeira (MAISEY *et al.*, 2008). Genes de virulência são responsáveis pela adaptação e sobrevivência do microrganismo no hospedeiro durante a infecção, sendo relatados em ilhas de patogenicidade (AGUIAR *et al.*, 2016). *S. agalactiae* é um microrganismo invasivo com capacidade de se adaptar às diversas mudanças ambientais, sendo necessária a expressão adequada de fatores de virulência em resposta ao ambiente extracelular (MILLER; MEKALANOS, 1990; MOUSLIM *et al.*, 2004).

*S. agalactiae* em sua maioria são recobertos pela cápsula polissacarídica (CP, antígeno tipo-específico) que promove a aderência do microrganismo às superfícies epiteliais, além de inibir a fagocitose pelos macrófagos e neutrófilos (CIELSLEWICZ *et al.*, 2005; NIZET; RUBENS, 2002). A CP desempenha um papel crítico na evasão imune através de seu mimetismo com epítomos de carboidratos do hospedeiro, interferindo com as vias de defesa dependentes de complemento e facilitando a internalização e sobrevivência intracelular em células dendríticas (NIZET *et al.*, 1996). Além disso, foi relatado que a CP medeia a formação de biofilme de *S. agalactiae* na presença de plasma humano (SHABAYEK; SPELLERBERG, 2018).

Atualmente mudanças têm ocorrido neste panorama, onde amostras de *S. agalactiae* naturalmente sem ou com baixa expressão de CP demonstrou virulência exacerbada, fenótipo hiperemolítico, aumento da persistência intracelular, capacidade de resistir ao tratamento com antibióticos e maior capacidade para penetrar a barreira hematoencefálica (GENDRIN *et al.*, 2018).

Todos os genes responsáveis pela síntese e ligação dos polissacárideos capsulares à parede celular de *S. agalactiae* (CPS) estão agrupados no operon *cps*. Este operon é composto por 16-18 genes, cujas sequências diferem entre os TC (BERTI *et al.*, 2014; SLOTVED; HOFFMANN, 2017). O operon *cps* é dividido em três regiões principais (CIELSLEWICZ *et al.*, 2005; TONIOLO *et al.*, 2015), sendo a porção inicial constituída pelos genes *cpsA-D* que são conservados entre todos os TC de *S. agalactiae* e apresentam ortólogos em outras espécies de estreptococos, permitindo a identificação deste gênero bacteriano. A parte central do operon (*cpsE-L*) determina o TC e compreende genes que codificam glicosiltransferases e uma polimerase; enquanto a porção final do operon (*neu-neuD*) codifica enzimas que sintetizam o ácido siálico (TONIOLO *et al.*, 2015).

O ácido siálico, também conhecido como ácido N-acetilneuramínico é um dos componentes da CP do *S. agalactiae* que inibe a ativação da via do complemento, impedindo a deposição de C3b à superfície do patógeno, dificultando assim, o reconhecimento do *S. agalactiae* pelas células do sistema imune (BASSET *et al.*, 2003; CARLIN *et al.*, 2009; MAISEY *et al.*, 2008; RAJAGOPAL, 2009). O gene *neu* localizado no operon *cps* é responsável pela produção do ácido siálico e sialilação da cápsula (POYART *et al.*, 2007; TONIOLO *et al.*, 2015). Estudo realizado por nosso grupo demonstrou que amostras de *S. agalactiae* pertencentes aos TC Ia, III e V, quando tratadas com neuraminidase e incubadas com macrófagos murinos residentes não apresentaram viabilidade intracelular, indicando a importância deste oligossacarídeo para sobrevivência do microrganismo em macrófagos (MONTEIRO *et al.*, 2004)

Destacam-se como fatores de virulência responsáveis pela disseminação sistêmica do *S. agalactiae* no hospedeiro, as toxinas formadoras de poros como a  $\beta$ -hemolisina/citolisina e o fator de Christie Atkins Munch Peterson (CAMP) (RAJAGOPAL, 2009). Em especial, a  $\beta$ -hemolisina é uma exotoxina pluripotente associada à membrana que desempenha um papel fundamental na virulência do *S. agalactiae*. A  $\beta$ -hemolisina promove a invasão de *S. agalactiae* nas células hospedeiras, permitindo que atravesse barreiras como as epiteliais, endoteliais e hemato-encefálica (DORAN *et al.*, 2003). Além disso, pode induzir a apoptose, liberar citocinas relacionadas com a invasão celular e resistência a fagocitose (DORAN; NIZET, 2004).

Uma das principais características fenotípicas é a zona de  $\beta$ -hemólise que circunda as colônias quando crescidas na presença hemácias (AYERS; RUPP, 1922; ROSA-FRAILE *et al.*, 2014). Portanto, a  $\beta$ -hemólise tem sido amplamente utilizada em laboratório para a identificação presuntiva do *S. agalactiae* (FACKLAM, 2002; SPELLERBERG *et al.*, 2011). Os determinantes genéticos que codificam a hemolisina e o pigmento do *S. agalactiae* foram previamente descritos, onde o operon *cyl* é regulado pelo sistema CovS/R que coordena a expressão de diversos fatores de virulência de *S. agalactiae*. O operon *cyl* composto por 12 genes (*cylX*, *cylD*, *cylG*, *acpC*, *cylZ*, *cylA*, *cylB*, *cylE*, *cylF*, *cylI*, *cylJ*, *cylK*) é exclusivo do *S. agalactiae* e foi inicialmente ligado à atividade hemolítica (SPELLERBERG *et al.*, 1999, 2000). Pesquisas demonstraram que apenas a mutação do gene *cylE*, resultou em fenótipo não hemolítico (PRITZLAFF *et al.*, 2001; ROSA-FRAILE *et al.*, 2014; SIGGE *et al.*, 2008).

A análise genética do operon *cyl* é essencial na investigação da transformação fenotípica da hemólise de *S. agalactiae*. O sistema CovS/R pode inibir a expressão do operon *cyl* de acordo com a estimulação ambiental e a deleção completa do sistema CovS/R foi capaz de tornar o *S. agalactiae* altamente hemolítico e com menor expressão de CP (CHOU *et al.*, 2018; JIANG *et al.*, 2005). Desta forma, variantes fenotípicas de *S. agalactiae* podem ser essenciais em análises futuras da patogênese molecular deste patógeno (CHOU *et al.*, 2018).

O fator CAMP (correspondente as iniciais, Christie, Atkins e Munch-Petersen) é codificado pelo gene *cfb* (KAYANSAMRUJAJ *et al.*, 2014). Foi observado que o fator CAMP pode se agrupar e formar poros discretos nas membranas de células alvo (LANG; PALMER, 2003), bem como interagir com a porção Fc das imunoglobulinas humanas da classe M e G tornando ineficaz a opsonização bacteriana (JURGENS *et al.*, 1987; RAJAGOPAL, 2009). Além disso, esta toxina pode se ligar a proteínas ancoradas a glicosilfosfatidilinositol, sugerindo que esta ligação a receptores presentes na célula hospedeira possa ser necessária para a lise celular, como observado em outras toxinas bacterianas (LANG; PALMER, 2003). O fator CAMP é amplamente utilizado no diagnóstico laboratorial presuntivo de *S. agalactiae* através da intensificação da hemólise (ponta de seta) promovida por amostra de *Staphylococcus spp.* produtora de  $\beta$ -lisina em placas de agar sangue contendo 5% de hemácias de carneiro (BORGER, 2005).

Outro importante fator para a patogênese do *S. agalactiae* é a hialuronidase, uma proteína de 110kDa codificada pelo gene *hylB* (GASE *et al.*, 1998; KAYANSAMRUJ *et al.*, 2014). A hialuronidase é uma endoglicosidase responsável por clivar cadeias de glicosaminoglicanas, principalmente a N-acetilglicosamina que compõe o ácido hialurônico, principal componente do tecido conjuntivo e do sistema nervoso (AKHTAR; KRISHNAN; BHAKUNI, 2006; HYNES; WALTON, 2000; RAJAGOPAL, 2009). A hialuronidase facilita a disseminação da infecção, sendo encontrada em altas concentrações nos tecidos placentários, fetais e líquido amniótico (RAJAGOPAL, 2009).

Para a evasão do sistema imune, a resistência dos patógenos às espécies reativas de oxigênio (ROS) produzidas pelo hospedeiro é essencial (HAMILTON *et al.*, 2006; RAJAGOPAL, 2009). *S. agalactiae* possui o gene *sodA* que codifica uma superóxido dismutase responsável pela resistência às espécies reativas de oxigênio e evasão imune (RAJAGOPAL, 2009). A superóxido dismutase converte oxigênio simples ou ânions de superóxidos ( $O_2^{\cdot-}$ ) em molécula de oxigênio ( $O_2$ ) e de peróxido de hidrogênio ( $H_2O_2$ ), que é subsequentemente metabolizado por catalases e peroxidases (RAJAGOPAL, 2009). Consequentemente, essas enzimas permitem que o *S. agalactiae* possa resistir ao estresse oxidativo durante a infecção no hospedeiro, pois reduzem a capacidade de reconhecimento e ativação dos mecanismos de defesa do hospedeiro (PALMEIRO, 2009; POYART *et al.*, 2001).

O *S. agalactiae* também apresenta como fatores de virulência proteínas com função de adesinas, como a proteína C, formada pelos antígenos alfa e beta, independentemente expressos (SPELLERBERG, 2000). A subunidade alfa é resistente ao tratamento com protease tripsina e a beta sensível, ambas as subunidades conferem resistência a fagocitose e à morte intracelular por leucócitos polimorfonucleares na ausência de anticorpos específicos (TRABULSI *et al.*, 2008). A proteína C diminui a fagocitose por opsonização por ser capaz de fazer a ligação inespecífica à porção Fc da imunoglobulina A humana (HICKMAN *et al.*, 1999; JERLSTROM *et al.*, 1996). Além disto, Baron *et al.* (2004) descreveram a interação da proteína C com glicosaminoglicanos presentes na superfície de células epiteliais, promovendo o rearranjo dos filamentos de actina, mediando a translocação do patógeno pela barreira epitelial e facilitando, assim, a invasão do patógeno.

A C5a peptidase é uma serino-protease altamente conservada de 128kDa presente na superfície dos *S. agalactiae*, capaz de clivar e inativar o componente

quimiotático C5a do complemento, colaborando para uma resposta imunológica ineficiente nos tecidos infectados (BROWN *et al.*, 2005; RAJAGOPAL, 2009). Dados publicados reportaram a capacidade de *S. agalactiae* se ligar, através da C5a peptidase, às células epiteliais humanas e proteínas de matriz extracelular (MEC) (CHENG *et al.*, 2002), agindo como uma adesina e promovendo a invasão bacteriana nas células epiteliais através de alterações no citoesqueleto da célula hospedeira. Contudo, ainda é escasso o entendimento real do papel da C5a peptidase na virulência do *S. agalactiae* (BECKMANN *et al.*, 2002).

Diversas bactérias patogênicas aderem às células hospedeiras por meio de proteínas de superfície que se ligam à MEC (SCHUBERT *et al.*, 2004). A MEC de tecidos de mamíferos é composta por glicoproteínas, como colágeno, laminina, fibronectina e fibrinogênio, formando uma estrutura macromolecular subjacente às células epiteliais e endoteliais (ARAUJO *et al.*, 2008; SCHUBERT *et al.*, 2004; SUHR *et al.*, 2010). O fibrinogênio é uma proteína de fase aguda sintetizada pelo fígado e tem sua liberação aumentada no processo inflamatório (VECINI *et al.*, 2006), estando presente no plasma sanguíneo e na MEC com importante função no processo de cicatrização (SCHUBERT *et al.*, 2002). Várias pesquisas têm demonstrado interações de *S. agalactiae* com as proteínas da MEC e, para cada uma dessas proteínas, existem receptores específicos. Em *S. agalactiae*, a peptidase C5a possui domínios de ligação à fibronectina (BECKMANN *et al.*, 2002), a proteína Lmb medeia a ligação à laminina humana (SPELLERBERG *et al.*, 1999) e as proteínas FbsA, FbsB e FbsC estão envolvidas na ligação ao fibrinogênio (BUSCETTA *et al.*, 2014; GUTEKUNST; EIKMANN; REINSCHIED, 2004; SCHUBERT *et al.*, 2002).

A aderência deste patógeno aos tecidos do hospedeiro é importante no início do processo infeccioso (RAJAGOPAL, 2009). *S. agalactiae* também pode clivar o fibrinogênio em fibrina, o que impede o reconhecimento e a fagocitose bacteriana, facilitando a sua evasão do sistema imune do hospedeiro (HARRIS, 2003). A ligação de *S. agalactiae* ao fibrinogênio é mediada por três proteínas conhecidas como FbsA, FbsB e FbsC, e curiosamente, seus genes recebem a mesma nomenclatura, *fibrinogen binding protein A, B e C*, que podem se ligar tanto ao fibrinogênio solúvel quanto ao imobilizado de humanos e de bovinos (BUSCETTA *et al.*, 2014; JONSSON *et al.*, 2005; SCHUBERT *et al.*, 2002). Estudos têm demonstrado que a proteína FbsA também tem função de agregação plaquetária, podendo causar outras injúrias teciduais durante a infecção (PIETROCOLA *et al.*, 2005), bem como

evitar a opsonização por macrófagos e neutrófilos (PIETROCOLA *et al.*, 2005; SCHUBERT *et al.*, 2002). Tettelin *et al.* (2005) relataram que FbsA é uma proteína de superfície associada a algumas amostras de *S. agalactiae*, enquanto FbsB é encontrada em todas as amostras de *S. agalactiae*. O gene que codifica FbsA foi detectado em quase todos os TC de *S. agalactiae*, porém o número de unidades repetidas de aminoácidos pode variar de amostra para amostra (SCHUBERT *et al.*, 2002).

Embora FbsB tenha analogia com FbsA em número de aminoácidos (16), as proteínas FbsA e FbsB não exibem similaridade significativa entre si. Além disso, a proteína FbsA carrega o motivo LPKTG e a proteína FbsB é desprovida do mesmo, indicando que a proteína FbsB é secretada ou não covalentemente ligada à parede celular. Contudo, análise da superfície bacteriana por fluorescência mostrou que ao menos uma fração da proteína FbsB permanece ancorada na parede celular do microrganismo (GUTEKUNST; EIKMANN; REINSCHIED, 2004).

A proteína FbsC, última a ser caracterizada na amostra de *S. agalactiae* NEM316, exibe uma sequência de aminoácidos quase que única. Análises genéticas mostraram que a proteína FbsC foi expressada por todos os isolados clínicos de *S. agalactiae* analisados, exceto aqueles pertencentes à linhagem hipervirulenta ST-17. FbsC demonstra ser importante para a ligação do *S. agalactiae* ao fibrinogênio, pois amostras mutantes apresentaram redução na capacidade de se ligar ao fibrinogênio, indicando que FbsC possa ser um importante fator durante a infecção neonatal (BUSCETTA *et al.*, 2014).

A aderência do *S. agalactiae* à laminina pode ser uma etapa crucial para o desenvolvimento da infecção (CHENG *et al.*, 2002). A *laminin binding protein* (laminina; Lmb) é uma glicoproteína multidomínio de 900 kDa da membrana basal, codificada pelo gene *Lmb*, fazendo parte do vasto repertório de fatores de virulência do *S. agalactiae* (BROWN *et al.*, 2002; RAGUNATHAN *et al.*, 2013). Constituída por três cadeias de polipeptídeo distintas: A, B1 e B2 que montam de forma reversível uma estrutura macromolecular, localizada na superfície celular bacteriana (RAGUNATHAN *et al.*, 2013). Isolados de *S. agalactiae* mutantes em *Lmb* apresentaram diminuição da aderência em células endoteliais microvasculares de cérebro humano (SPELLERBERG *et al.*, 1999; TENENBAUM *et al.*, 2007).

O *S. agalactiae* foi capaz danificar o epitélio pulmonar, levando à exposição das estruturas subjacentes da membrana basal (NIZZET *et al.*, 1996). Assim, a

adesão aos componentes de membrana basal pode ser crítico para a colonização bacteriana do epitélio danificado e, conseqüente invasão do *S. agalactiae* para a corrente sanguínea (SPELLERBERG *et al.*, 1999). Desta forma, as proteínas Lmb e FbsA são importantes fatores de virulência relacionados com a aderência do *S. agalactiae* às células do hospedeiro e a invasão da barreira hematoencefálica (SAFADI *et al.*, 2010; TENENBAUM *et al.*, 2007).

O ALT pode facilitar o processo de aderência como primeira etapa na infecção. Apesar do envolvimento do ALT como adesina ainda não seja bem explicado, estudos utilizando células do epitélio vaginal tratadas com ALT, apresentaram diminuição na aderência bacteriana (COX *et al.*, 1993). Porém, se constatou que o ALT foi incapaz de inibir a aderência de *S. agalactiae* às células de epitélio vaginal (TAMURA *et al.*, 1994). O ALT é também um importante indutor de citocinas (TNF- $\alpha$ , IL-1 $\beta$  e IL-6) em células mononucleares, sendo a resposta inflamatória induzida pela ligação do ALT ao receptor *Tool Like* (TLR)-2/6 e dependente da fosforilação de resíduos de tirosina no TLR-2 (HENNEKE *et al.*, 2005; SPELLERBERG, 2000).

O *S. agalactiae* produz uma exotoxina (CM-101) relacionada com a resposta inflamatória e a hipertensão pulmonar (FARLEY, 2001; HELLERQVIST; SUNDELL; GETTINS, 1987). Este polissacarídeo pode ser detectado pela técnica de ELISA em líquido e urina de neonatos com sepse por *S. agalactiae* (BORGER, 2005). Yakes *et al.* (2000) relataram também que a utilização do domínio purificado desta proteína inibiu o crescimento de tumores. Contudo, seu mecanismo de ação permanece pouco conhecido.

*S. agalactiae* codifica apêndices na superfície celular denominadas *pili* (DRAMSI *et al.*, 2006; LAUER *et al.*, 2005) que possibilitam o desenvolvimento de infecções invasivas em seres humanos, contribuindo para a ocorrência de meningite e sepse neonatal (DORAN *et al.*, 2004; MAISEY *et al.*, 2007). Os *pili* são codificados pelos genes *pilA*, *pilB* e *pilC* (MAISEY *et al.*, 2007) localizados em dois clusters, ilha-1 (PI-1) e ilha-2 (PI-2), sendo que este último apresenta duas variantes, a PI-2a e PI-2b (LAUER *et al.*, 2005; RAJAGOPAL, 2009; ROSINI *et al.*, 2006). Pi-2a em *S. agalactiae* tem sido descrito como principal *pilus* encontrado na espécie, mediando à aderência do *S. agalactiae* nas células epiteliais humanas (RINAUDO *et al.*, 2010). Pi-2b também tem sido relacionado com a formação de biofilme, além de mediar à

aderência e invasão do patógeno em células endoteliais da microvasculatura cerebral (hBMEC) (PEZZICOLI *et al.*, 2008; PÉRICHON *et al.*, 2017).

As ilhas PI-1 e PI-2 são constituídas por genes que codificam a proteína de transporte, motivo LPXTG, subunidade principal dos *pili*, duas proteínas acessórias (AP1 e AP2) e dois genes sortase envolvidos na montagem dos *pili* (DRAMSI *et al.*, 2006, MAIONE *et al.*, 2005; ROSINI *et al.*, 2006). As ilhas de *pili* são formadas por três subunidades protéicas: PilA, PilB e PilC e a sua montagem envolve duas classes de proteínas sortases tipo C, StrC3 e StrC4 (KONTO-GHIORGH *et al.*, 2009). As proteínas sortases pertencem a uma família de proteínas de organismos procariotos envolvidas na formação de *pili* (LEMIEUX *et al.*, 2008). Acredita-se que as sortases tipo C possam polimerizar os *pili* pela formação de ligações covalentes entre diferentes subunidades (DRAMSI *et al.*, 2005; MARESSO; SCHNEEWIND, 2008).

A extremidade dos *pili* geralmente apresenta propriedades adesivas, que promovem a ligação bacteriana à MEC e/ou a ligação a receptores celulares do hospedeiro (MARESSO; SCHNEEWIND, 2008). Diversos estudos comprovaram que os *pili* são capazes de mediar à resistência aos peptídeos catiônicos antimicrobianos (PCA) e também facilitar a aderência dos *S. agalactiae* em células do hospedeiro (DRAMSI *et al.*, 2006, MAISEY *et al.*, 2007, 2008; PEZZICOLI *et al.*, 2008;). Embora não seja de total entendimento o mecanismo pelo qual os *pili* permitem que os *S. agalactiae* resistam aos PCA, isolados deste microrganismo que apresentavam deficiência na proteína acessória PilB apresentaram uma diminuição na virulência e um aumento na suscetibilidade à fagocitose e PCA (MAISEY *et al.*, 2007; 2008). Em amostras mutantes, a aderência dos *S. agalactiae* às células epiteliais foi reduzida, evidenciando a relevância dos *pili* na patogênese das infecções ocasionadas pelo patógeno (DRAMSI *et al.*, 2006; SLOTVED; HOFFMANN, 2017).

Diversos dados demonstram os *pili* mediando a aderência e migração transepitelial (DRAMSI *et al.*, 2006; KONTO-GHIORGH *et al.*, 2009; PEZZICOLI *et al.*, 2008; ROSINI *et al.*, 2006) e viabilidade intracelular em células endoteliais microvasculares cerebrais (MAISEY *et al.*, 2007). Sharma *et al.* (2013) demonstraram que mutantes para *pili* (TC Ia, III e VII) diminuíram significativamente a aderência, e o TC VII também reduziu a invasão.

A capacidade de formação de biofilme pela maioria dos isolados de *S. agalactiae* foi confirmada utilizando mutantes, na qual o *pilus* tipo 2a (PI-2a) foi



essencial para formação de biofilme, enquanto os *pili* 1 e 2b não apresentaram o mesmo resultado (RINAUDO *et al.*, 2010). O *pilus* 2b está relacionado à maior sobrevivência intracelular em macrófagos (CHATTOPADHYAY *et al.*, 2011). Este tipo de *pilus* foi recentemente caracterizado em amostras de *S. agalactiae* ST-17, onde a análise do genoma demonstrou que 100% destes isolados apresentaram os genes *Pi-2b* (PÉRICHON *et al.*, 2017). Estudos reforçam o envolvimento dos *pili* do *S. agalactiae* na aderência, invasão e formação de biofilme, no entanto novas funções para os *pili* ainda não foram descritas (PÉRICHON *et al.*, 2017; SLOTVED; HOFFMANN, 2017).

Compreender a expressão de fatores de virulência de *S. agalactiae* é útil para o desenvolvimento de estratégias terapêuticas (SLOTVED; HOFFMANN, 2017). Desta forma, os fatores de virulência são críticos para a capacidade dos *S. agalactiae* em causar doenças, pois favorecem a sobrevivência e colonização bacteriana, estabelecendo um nicho adequado para que o microrganismo possa replicar, disseminar e evadir das defesas do hospedeiro (ALBERTS *et al.*, 2004; PÉRICHON *et al.*, 2017).

### 1.5 Endotélio e a interação com *S. agalactiae*

Células endoteliais (CE) formam uma barreira dinâmica semipermeável entre o espaço vascular dos vasos e os tecidos que os cercam, Jaffe (1987 *apud* JIETANG *et al.*, 2013). Estas células estão envolvidas em modulações da homeostase metabólica, equilíbrio vascular, vasoconstrição e vasodilatação, adesão, coagulação e anticoagulação. Além disso, modulam o tônus vasomotor, regulam a estrutura vascular, controlam o fluxo sanguíneo e medeiam respostas inflamatórias e imunológicas (DAVIDSON *et al.*, 2010). A permeabilidade endotelial tem sido extensivamente investigada no contexto de patologias (FUNAMOTO *et al.*, 2017), onde citocinas pró-inflamatórias, vírus, bactérias e fatores de crescimento podem alterar a função de barreira endotelial (BOGATCHEVA; GARCIA; VERIN, 2002; KELLER *et al.*, 2003).

A ativação das CE decorre de uma resposta adaptativa normal, uma vez que elas estão constantemente reconhecendo e respondendo alterações do ambiente

extracelular. No entanto, a natureza e duração das respostas não dependem apenas do tipo de estímulo, mas também do tipo de sistema; isto é, células endoteliais de artérias e veias podem responder de forma diferente a um mesmo estímulo (AIRD, 2004). Por expressarem diferentes antígenos e receptores em suas superfícies, estas células respondem diferentemente a um mesmo estímulo (SUMPIO *et al.*, 2002).

As CE desempenham um importante papel durante infecções, estando entre as primeiras células do corpo que mantêm contato com patógenos invasores (PETERS *et al.*, 2003; TORRES-ESTAY *et al.*, 2016). A inflamação é um mecanismo vital de defesa do organismo caracterizado pela resposta dos tecidos à presença de microrganismos ou injúria. A lesão tecidual causada por bactérias, traumas, agentes químicos, calor ou qualquer outro fenômeno, determina a liberação de diversas substâncias que produzem alterações secundárias acentuadas nos tecidos lesionados (GLÜCK *et al.*, 2001; RESCHE-RIGON *et al.*, 2006). Um dos primeiros eventos de resposta endotelial aos estímulos inflamatórios é a expressão de moléculas de adesão como a molécula de adesão intercelular-1 (ICAM-1), molécula de adesão celular vascular (VCAM), E-selectina e P-selectina que regulam a entrada dos leucócitos para os tecidos afetados (CAPOSIO *et al.*, 2007; ROH *et al.*, 2011). A transmigração aumentada de monócitos sanguíneos circulantes através do endotélio vascular é considerada uma contribuição importante na patogênese de doenças inflamatórias, como a sepse (MATTILA *et al.*, 1998; SEMERARO *et al.*, 2012; TORRES-ESTAY *et al.*, 2016).

A interação de patógenos com células eucarióticas é um processo altamente dinâmico e selecionado pela evolução, sendo o primeiro passo para o estabelecimento de infecções bacterianas (MOLINARI; CHHATWAL, 1999). O sucesso das infecções por patógenos depende do processo multifatorial que se inicia com a aderência do microrganismo às células hospedeiras, colonização dos tecidos, invasão tecidual, multiplicação ou não do patógeno no hospedeiro, persistência bacteriana intracelular e como resultado final do processo, a instalação e possível disseminação para outros tecidos e órgãos (PIZZARO-CERDA; COSSART, 2006).

Para que todo este processo ocorra com êxito, estruturas do patógeno devem desempenhar importantes funções e mediar a interação. De maneira geral, são as adesinas da superfície bacteriana que interagem especificamente a uma grande

variedade de receptores nas células eucarióticas, incluindo resíduos de açúcar como manose, lipídeos e gangliosídeos ou proteínas como a fibronectina que permitem a colonização do tecido hospedeiro pelo patógeno (KONEMAN *et al.*, 1997; TAMURA *et al.*, 1994). Apenas realizar o processo de aderência com êxito, não garante que o patógeno irá conseguir se estabelecer. Desta forma, frequentemente a bactéria usa a capacidade de aderência e a invasão de células susceptíveis para realizar a evasão do sistema imune, garantindo assim a sua sobrevivência (ROY; VAN DER GOOT, 2003; SHABAYEK; SPELLERBERG, 2018).

Apesar dos avanços nas técnicas da biologia celular e molecular na última década, os mecanismos de aderência e invasão de *S. agalactiae* em células hospedeiras ainda não estão totalmente esclarecidos. Ensaio de interação *in vitro* com *S. agalactiae* vem sendo exaustivamente desenvolvidos com diferentes linhagens celulares como: células do epitélio vaginal (TAMURA *et al.*, 1994), epitélio pulmonar (COSTA *et al.*, 2011; TAMURA *et al.*, 1994), células endoteliais (COSTA *et al.*, 2005; SANTOS *et al.*, 2003), macrófagos murinos e humanos (FETTUCIARI *et al.*, 2011) para tentar esclarecer os mecanismos e as funções específicas de cada estrutura do *S. agalactiae*.

O endotélio vascular contribui consideravelmente para o processo inflamatório, e a magnitude da resposta inflamatória é crucial ao organismo. Assim como a insuficiência nesta resposta (ou imunodeficiência) pode levar à infecção generalizada e a resposta excessiva pode causar prejuízo ao organismo (FALLAVENA *et al.*, 2009; TIZARD, 2002; TRACEY, 2002). Uma cultura primária pode ser estabelecida a partir do crescimento de células oriundas de um fragmento de tecido obtido por desagregação mecânica ou enzimática (MORAES *et al.*, 2007). As células que sobrevivem a este processo aderem e formam a primeira monocamada de células, sendo muito utilizada para estudar o comportamento de determinada célula *in vitro* porque mantém as características genóticas e fenotípicas das células até a renovação tecidual (MORAES *et al.*, 2007). A célula endotelial da veia umbilical humana (HUVEC) é um modelo clássico de cultura primária, por ser de fácil obtenção e manutenção (VALBUENA; WALKER, 2006). O modelo de extração e cultivo das HUVEC se encontra padronizado em nosso laboratório, de acordo com a técnica descrita por Jaffe *et al.* (1973).

## 1.6 Diversidade genética de *S. agalactiae*

A primeira publicação de sequenciamento genômico completo de *S. agalactiae* ocorreu em 2002, onde a amostra NEM316 pertencente ao TC III foi sequenciada demonstrando um baixo percentual de conteúdo G+C (GLASER *et al.*, 2002). A similaridade genômica entre os isolados humanos de *S. agalactiae* com outros *Streptococcus* spp. foi demonstrada por Tettelin *et al.* (2002). Porém, Glaser *et al.* (2002) evidenciaram que as regiões gênicas codificadoras de proteínas relacionadas a elementos móveis (transposons, elementos de inserção e bacteriófagos) são distintas de outras espécies do gênero, como *S. pyogenes* e *S. pneumoniae*. A maioria desses genes parece estar associado, especialmente, a adaptação, capacidade de colonização e infecção do *S. agalactiae* aos diferentes hospedeiros e nichos, podendo ter sido adquiridos por transmissão horizontal de diferentes espécies bacterianas (ZUBAIR *et al.*, 2013).

Genomas de *S. agalactiae* pertencentes aos cinco principais TC (Ia, Ib, II, III e V) isolados de sapo, peixe, camelo, bovino e humano foram sequenciados, descrevendo genes não relatados anteriormente (RICHARDS *et al.*, 2011; ZUBAIR *et al.*, 2013). Os genomas sequenciados de *S. agalactiae* indicam que a espécie pode ser definida pelo seu pangenoma, devido a possibilidade de novos genes serem descritos a cada novo genoma sequenciado da espécie (TETTELIN *et al.*, 2005).

Tettelini *et al.* (2005) descreveram o pangenoma de *S. agalactiae*, estruturado por um genoma central, contendo genes de funções reguladoras e constitutivos que representaram aproximadamente 80% do genoma total, presentes em amostras de *S. agalactiae*. A presença de genoma acessório associado aos elementos móveis e extras cromossomais, sugere que a aquisição de características específicas seja dependente de transferência gênica lateral (TETTELIN *et al.*, 2005). O sequenciamento do genoma de *S. agalactiae* possibilitou uma análise molecular mais detalhada com maior conhecimento da presença de ilhas de patogenicidade, resistência a antimicrobianos, fatores de virulência presentes em áreas distintas do genoma, assim como do sistema regulatório pleiotrópico de dois componentes da CovRS (MAISEY *et al.*, 2008).

O sequenciamento do *cluster cps*, genes referentes à cápsula polissacarídica de *S. agalactiae*, possibilitou a correlação entre a identificação definitiva por sorologia e molecular, reduzindo a quantidade de amostras não tipáveis (NT) por sorologia clássica. Contudo, alguns isolados permanecem ainda não tipáveis (KONG *et al.*, 2008; SLOTVET *et al.*, 2007). A tipagem molecular, baseada na amplificação e sequenciamento gênico, possuem alta reprodutibilidade e potencial discriminativo, permitindo não só a resolução de discrepâncias nos resultados da tipagem por sorologia, mas também na identificação de novos TC (FIOLO *et al.*, 2012; KONG *et al.*, 2002, 2008).

Atualmente, no Brasil são raras as equipes empenhadas em caracterizar o TC dos *S. agalactiae*, tornando-se relevante a implantação de técnicas que identifiquem os diferentes tipos em nossa comunidade, com elevada reprodutibilidade, baixo custo e capacidade discriminativa. O uso da sorologia para a tipagem de *S. agalactiae* era a metodologia laboratorial clássica, sendo insuficiente para discernir isolados clínicos envolvidos em diversas infecções invasivas (BISHARAT *et al.*, 2004). Os métodos sorológicos têm limitações, uma vez que podem falhar no tipo de isolamento devido à falta ou baixa expressão do polissacárideo capsular nas condições experimentais. Além disso, estes métodos são altamente dependentes da qualidade dos anticorpos utilizados e da experiência do laboratório (YAO *et al.*, 2013). A dificuldade em gerar os soros específicos em coelhos, reações cruzadas entre os diferentes tipos capsulares e o elevado número de amostras não tipáveis (NT) eram algumas das desvantagens produzidas pela tipagem por sorologia (AMUNDSON *et al.*, 2005).

O emprego da técnica de reação em cadeia da polimerase (PCR) colaborou para uma identificação mais adequada dos diferentes TC de *S. agalactiae* (BISHARAT *et al.*, 2004; JONES *et al.*, 2003). A amplificação por PCR está entre os melhores métodos genotípicos para a tipagem capsular de *S. agalactiae* tanto pela resolução de discrepâncias em resultados contraditórios da sorologia, especificidade, mas também na identificação de novos TC (IMPERI *et al.*, 2010). A implantação da tipagem molecular por PCR multiplex fornecerá um melhor entendimento do perfil de TC prevalentes nas diferentes regiões brasileiras, assim como ampliar os conhecimentos sobre a importância clínica do *S. agalactiae*.

Análises utilizando *Multi Locus Enzyme Electrophoresis* (MLEE) demonstraram que amostras de *S. agalactiae* isoladas de neonatos do TC III

pertenciam a dois distintos *clusters* evolucionários que refletiam o seu potencial patogênico (JOHRI *et al.*, 2006). O que foi confirmado em amostras oriundas de humanos de diversos locais no mundo por *Multi Locus Sequence Typing* (MLST), baseados na combinação de alelos de sete genes *housekeeping* codificadores de enzimas relacionadas ao metabolismo bacteriano: álcool desidrogenase (*adhP*), fenilalanil tRNA sintetase (*pheS*), aminoácido transportador (*atr*), glutamina sintetase (*glnA*), serina desidratase (*sdhA*), glicose quinase (*glcK*) e transketolase (*tkt*) (JONES *et al.*, 2003; JOHRI *et al.*, 2006).

A técnica de MLST consiste na associação dos perfis de alelos encontrados para os sete genes *housekeeping* gerando um perfil e uma sequência tipo (ST). O fato de existir um banco de dados de MLST, permite uma comparação entre isolados bacterianos sem a necessidade de solicitar troca de material biológico, implementando um modelo simples de expansão clonal e diversidade genética (FEIL *et al.*, 2004; MAIDEN *et al.*, 2006). Consequentemente, cinco principais complexos clonais (CCs), CC1, CC10, CC23, CC19 e CC17 foram identificados, enfatizando a diversidade de *S. agalactiae* em humanos (JONES *et al.*, 2006; KARATAN; WATNICK, 2009). Recentemente, alguns destes complexos clonais (CC1, CC23, CC17 – especificamente a ST-17) foram propostos como biomarcadores genômicos de fatores virulência para *S. agalactiae* e alguns genes como *HvgA*, *Srr1/2*, *pil2b* foram encontrados associados ao ST-17 (LIN *et al.*, 2014), no entanto, estes genes são encontrados em 76% dos casos de meningite tardia em recém-nascidos (D'URZO *et al.*, 2014).

Interessantemente, estudos demonstram que um único clone pertencente ao TC III, designado ST-17 pertencente ao CC17, foi responsável por uma porção significativa de infecções de início precoce e de início tardio (>80%). Desta forma, o CC17 dos *S. agalactiae*, mais especificamente a ST-17 tornou-se um clone emergente que se propagou com sucesso a nível mundial, sendo um importante alvo para estudos epidemiológicos (PÉRICHON *et al.*, 2017; POYAT *et al.*, 2008). Pesquisas associam o aumento da prevalência de infecções invasivas neonatais a partir de 1960 com a propagação da linhagem hipervirulenta ST-17 (BELLAIS *et al.*, 2012; JONES *et al.*, 2006; LAMY *et al.*, 2006). Em 2010, Sorensen e colaboradores, descreveram que o recente surgimento e expansão de isolados altamente virulentos em humanos, pode ser explicado, em parte, pela disseminação mundial da ST-17. Constatou-se que o *S. agalactiae* ST-17 tem uma alta propensão para a persistência

no trato intestinal em elevado número e por períodos prolongados, favorecendo a sua translocação intestinal e posterior sepse (ELLING *et al.*, 2014). *S. agalactiae* ST-17 causam a síndrome de início tardio (BICHARAT *et al.*, 2005; JONES *et al.*, 2006; LIN *et al.*, 2006) e com maior frequência a meningite quando comparada com outras STs (D'URZO *et al.*, 2014). O mecanismo exato pelo qual ST-17 provoca a síndrome de início tardio mais frequentemente do que outros STs não é bem compreendido, embora evidências recentes indiquem que ST-17 exibe uma combinação específica conservada das proteínas secretadas e expostas à superfície (D'URZO *et al.*, 2014; MANNING *et al.*, 2009; TAZI *et al.*, 2010). Um determinante chave da fisiopatologia da síndrome de início tardio associada a ST-17 é a proteína HvgA expressa na superfície ST-17, crítica para a colonização e translocação intestinal de *S. agalactiae* através da barreira hemato-encefálica durante a meningite (KANG *et al.*, 2017; TAZI *et al.*, 2010).

Um estudo realizado por Bellais *et al.* (2012) detectou o clone hipervirulento ST-17 entre as amostras de *S. agalactiae* do tipo IV. A análise da sequência revelou que a mudança de cápsula do tipo III para IV teria ocorrido devido à troca de um fragmento de DNA contendo todo o operon *cps*. A partir deste evento genético, verificou-se que o DNA sequenciado era idêntico aos demais isolados de *S. agalactiae* sugerindo que estas sejam derivadas de um ancestral comum (BELLAIS *et al.*, 2012).

Curiosamente, os isolados ST-17 possuem características únicas independentes do gene *cps* (WANG *et al.*, 2015). O gene *gbs2018*, codifica uma proteína de superfície de *S. agalactiae* específica do clone ST-17, alvo utilizado por PCR para rastreio molecular de ST-17 em amostras de *S. agalactiae*, altamente virulentas (LAMY *et al.*, 2006). O surgimento de clones hipervirulentos de *S. agalactiae* pode ser resultado de eventos de transferência genética horizontal (GLASER *et al.*, 2002). Bisharat *et al.* (2004) realizaram uma análise filogenética que determinou que o clone hipervirulento ST-17 surgiu a partir de uma linhagem de isolados de bovinos.

Desta forma, a ST-17, atualmente, pode estar relacionada com a capacidade de persistir após o tratamento antibiótico ou que possua características moleculares únicas que a torne hiperinvasiva (BRUEGGEMANN *et al.*, 2003; SIX *et al.*, 2015). Pouco se sabe também sobre as relações filogenéticas entre os *S. agalactiae* ST-17 isolados de infecções neonatais e adultas. Embora, Teatero *et al.* (2016)

descobriram que amostras ST-17 isoladas dessas infecções são intercaladas e agrupadas dentro uma mesma árvore filogenética, demonstrando o maior potencial virulento das amostras pertencentes a esta sequência tipo.

Análises comparativas do genoma de *S. agalactiae* revelaram que a glicoproteína Srr2 (Srr), rica em serina, é também uma proteína ancorada na parede celular específica para as amostras ST-17, enquanto que os isolados não ST-17 expressam Srr1 (CUNHA *et al.*, 2014; SIX *et al.*, 2015). Six *et al.* (2015) mostraram ainda que Srr2 aumentou a sobrevivência e a persistência bacteriana em modelo murino de meningite. Srr2 é uma adesina multifacetada usada pelo clone ST-17, que contribui para a disseminação e invasividade bacteriana e, consequentemente, para a meningite (SIX *et al.*, 2015).

### 1.7 Aspectos da genômica e do sequenciamento

A genômica é um recente ramo da ciência, em expansão, que apresenta como atributo a caracterização de genomas completos (POHLHAUS; COOK-DEEGAN, 2008; SCHAEFER *et al.*, 2006). Contextualizando a palavra “genoma”, Miller; Koren; Sutton (2010) descrevem que genoma é todo material genético contido em determinado organismo, incluindo toda a informação necessária para manutenção da vida.

Os laboratórios brasileiros, desde a década de 90, têm conduzido seus experimentos de geração de dados genômicos dentro de grandes redes genômicas. Essa conexão entre centros de pesquisas públicos e/ou privados atingiu vários estados brasileiros como São Paulo, Minas Gerais, Rio de Janeiro, Santa Catarina, Rio Grande do Sul, Pará e Bahia (MILLER; KOREN; SUTTON, 2010). A importância no estabelecimento dessas estruturas, além do gerenciamento centralizado de informações geradas, visou possibilitar o avanço da área (SIMPSON, 2000). A primeira grande rede foi criada em 1997, esta conectou mais de trinta unidades de pesquisa em São Paulo e foi pioneira no sequenciamento de um genoma completo em território nacional, o genoma do fitopatógeno *Xylella fastidiosa* (SIMPSON *et al.*, 2000). Desde então, outras redes genômicas se formaram e forneceram meios para alavancar e projetar o Brasil em um cenário favorável no campo. Entre elas, vale



destacar a Rede Genoma Brasileira que é formada por 25 laboratórios, um centro de bioinformática e três laboratórios centrais distribuídos em todo território nacional. Esta foi responsável pelo sequenciamento do genoma das bactérias *Chromobacterium violaceum*, *Mycoplasma hyopneumoniae* e *Mycoplasma synoviae* (VASCONCELOS *et al.*, 2005).

Outro consórcio que merece destaque é a Rede Genoma de Minas Gerais (RGMG), financiada pela Fundação de Amparo à Pesquisa no Estado de Minas Gerais (FAPEMIG). A RGMG conta com a colaboração de nove instituições distribuídas não só pelo estado de Minas Gerais, mas também nos estados da Bahia, Rio de Janeiro e do Ceará, e esteve envolvida em grandes projetos de sequenciamento (CAPANEMA *et al.*, 2008). Pode-se citar também, a região Norte que possui a Rede Paraense de Genômica e Proteômica (RPGP), a qual foi fundada em 2009 e é financiada pela Fundação de Amparo à Pesquisa do Estado do Pará (FAPESPA) que dispõe de plataformas de sequenciamento de DNA de última geração. Como resultado da criação de redes genômicas e parcerias entre centros que desenvolvem pesquisa na área, diversos organismos estão com seus genomas caracterizados ou são alvos de intenso estudo.

As tecnologias de sequenciamento de nova geração, tecnologias NGS, possibilitaram avanços significativos no campo do conhecimento dos genomas dos organismos. Com a redução dos custos de sequenciamento e aumento na rapidez do processamento das amostras, o número de genomas sequenciados e disponíveis em bancos de dados cresceu exponencialmente. Paralelamente, a enorme quantidade de dados gerados e a falta de recursos humanos capacitados para trabalharem com estas informações representam desafios significativos para a bioinformática (GLENN, 2011). Dessa forma, armazenar racionalmente os dados, gerenciar soluções computacionais e a análise exploratória dos dados viabilizam a extração de informações valiosas a respeito da sequência genômica do organismo (LIU *et al.*, 2012).

É sabido que os dados provêm de diferentes plataformas de sequenciamento, que por sua vez, utilizam diferentes metodologias experimentais, gerando sequências com tamanhos desiguais e em formatos distintos. Além disso, cada plataforma NGS possui tendências associadas à metodologia, tal como, erros de nomeamento de bases (LIU *et al.*, 2012). Para lidar com esses desafios, houve um grande desenvolvimento de programas e ferramentas para análise de dados, desde

a avaliação da qualidade das leituras até a anotação e montagem dos genomas. As plataformas NGS utilizam diferentes metodologias de sequenciamento, mas em comum, elas não precisam passar pela etapa de clonagem dos fragmentos de interesse. Adicionalmente, as tecnologias NGS são capazes de gerar milhões de sequências em uma única corrida (THUDI *et al.*, 2012).

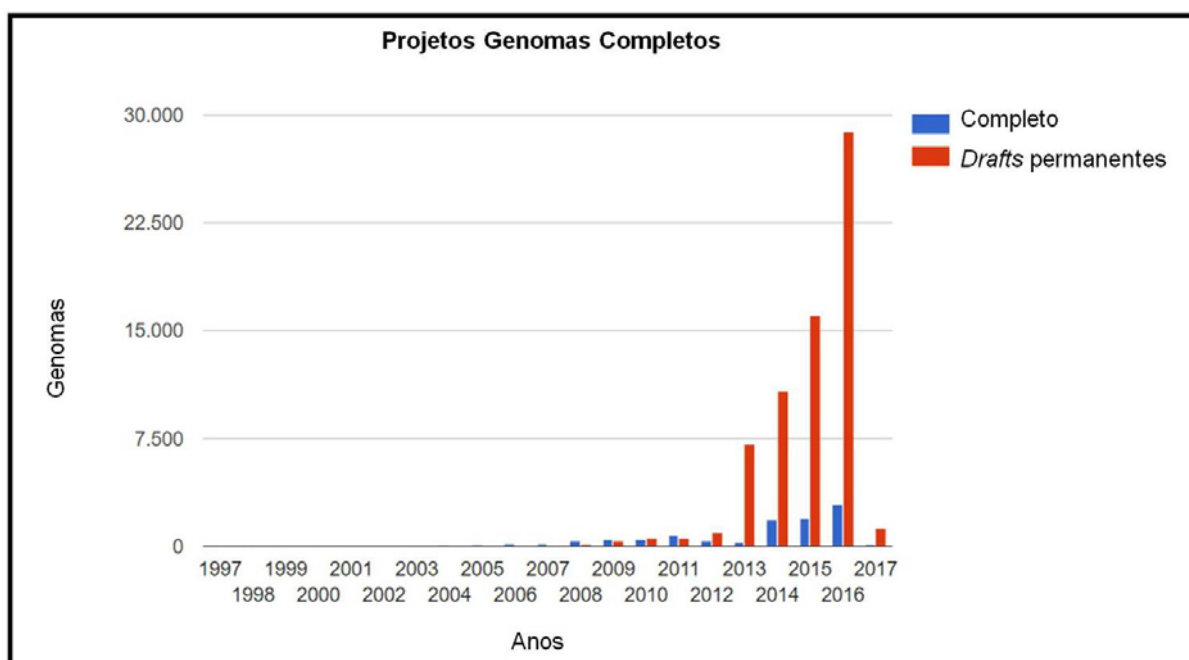
Existem diversas plataformas NGS, dentre elas Illumina, Ion Torrent, Pacbio, Nanopore que estão sendo utilizadas no sequenciamento *de novo*, re-sequenciamento de genoma e análises de genomas completos e de transcriptomas (MOHAMED; SYED, 2013). Dentre essas plataformas podemos salientar as da Illumina (*Genome Analyser*), pois se destacam pelo grande volume de sequências obtidas, grande espectro de utilidade, além de ter o menor custo por base sequenciada do mercado, fazendo com que a empresa venha dominando a indústria de sequenciamento em larga escala (QUAIL *et al.*, 2012). O sistema de sequenciamento HiSeq da Illumina tem alto desempenho e é amplamente adotado. Em termos de dispêndio monetário, o HiSeq reduziu o custo do sequenciamento a níveis incomparáveis (Illumina). O método de sequenciamento adquirido pela Illumina foi uma grande inovação, e atualmente, essa metodologia viabiliza o sequenciamento de fragmentos de 2 X 150 pb na plataforma HiSeq e 2 X 300 pb na MiSeq (GLENN, 2011).

O surgimento de novas tecnologias de sequenciamento possibilitou o sequenciamento de genomas completos em poucas horas (HENSON; TISCHLER; NING, 2012; YI *et al.*, 2014), o que vem contribuindo para os trabalhos nos três domínios da vida, exemplo o domínio *Bacteria* que até janeiro de 2017 apresentou 83.246 projetos relacionados e também favoreceu o crescimento do número de *drafts*, até o final do ano de 2016, com 28.879 projetos e de 2.951 projetos de genomas completos depositados nos bancos de dados públicos (Figura 3). Portanto, podemos observar o avanço e o crescimento impressionante no número de pesquisas envolvendo esta área, tornando-se de extrema importância para o desdobramento de análises mais profundas de organismos até então poucos conhecidos.

Com o aumento da produção de dados e o comprimento das leituras geradas pelas recentes tecnologias de sequenciamento, se tornou necessário desenvolver novas ferramentas e *pipelines* computacionais para a realização das etapas de montagem e fechamento de genomas. Desta maneira, a bioinformática encara

novos desafios, pois cada vez mais aumenta a necessidade de estruturas computacionais mais robustas, capazes de armazenar e processar estes dados (EL-METWALLY *et al.*, 2013). Atualmente, a bioinformática é o ápice na interpretação de dados genômicos e devido o seu surgimento e ao progresso nessa área, foi possível a criação de bancos de dados para gerenciar informações genômicas e o sequenciamento dos primeiros genomas completos (HUTCHISON III, 2007; ZHANG *et al.*, 2011).

Figura 3 - Quantidade total de genomas completos e *drafts* depositados em GOLD, onde o eixo X representa o ano e o eixo Y a quantidade de genomas



Fonte: adaptado de GOLD, 2017.

Atualmente, com o advento das tecnologias de sequenciamento de alta performance, a obtenção de dados genômicos se tornou inquestionavelmente mais rápida, em grande volume e depositado diariamente em bancos de dados genômicos. Com isso, as ciências biológicas têm despendido cada vez mais de ferramentas computacionais. Com o desenvolvimento das plataformas de análise genômica e das ferramentas de bioinformática, hoje, é possível a realização de inúmeras análises e estudos, como por exemplo, o alinhamento e montagem *de novo* de sequências, detecção de polimorfismos e outras variações estruturais em

genomas, análises de transcritomas e metagenomas, desenvolvimento de marcadores moleculares, entre outras aplicações (GIACHETTO; HIGA, 2015). Assim, podemos afirmar que a bioinformática se tornou uma ferramenta fundamental nas pesquisas de biologia molecular (PETREY; HONIG, 2014).

Neste contexto, torna-se indispensável o uso de abordagens computacionais visando à integração dos dados gerados pelo sequenciamento em larga escala de DNA com a informação biológica existente e depositada em bancos de dados. A integração dessas duas áreas de conhecimento, biologia computacional e biologia experimental, permite uma melhor compreensão do organismo biológico e fornece modelos matemáticos que podem direcionar os experimentos de bancada.

### 1.8 Montagem do genômica

Genomas de organismos são sequenciados e depositados nos bancos de dados em todo o mundo, fornecendo informações valiosas dos mais diversos organismos e sistemas. Porém, antes de agregar valor biológico para estas sequências, o processo de montagem das sequências dos genomas é uma etapa fundamental em um projeto de sequenciamento (MILLER; KOREN; SUTTON, 2010). A montagem é definida como um processo onde as leituras são agrupadas por identidade entre as bases, ou seja, leituras de sequências de DNA (*reads*) geradas pelo sequenciador. Esta união é obtida através do alinhamento das sequências e sobreposição das bases de DNA, sendo chamado de *contig*. O consenso é construído com a base mais frequente no alinhamento para cada posição do *contig*, visando à reconstrução do genoma alvo com o objetivo de representar a molécula de DNA que compõe cada cromossomo da espécie estudada (GIACHETTO; HIGA, 2015).

Diversos projetos de genomas não almejam a representação total do cromossomo, mas um mapa global (*draft*) contendo a maior parte da sequência do genoma, sequenciando as regiões que faltam de acordo com a necessidade. Estes mapas globais são formados por vários *contigs* ordenados, a montagem pode ser classificada em montagem por referência ou *de novo* (MILLER; KOREN; SUTTON, 2010). Especificamente, refere-se a métodos utilizados para determinar a sequência

de DNA quando não há nenhuma sequência genômica conhecida anteriormente e disponível para uso como referência. As diferentes estratégias de sequenciamento *de novo* são quase sempre complexas e difíceis, particularmente quando o genoma é grande e o DNA analisado possui sequências que se repetem muitas vezes, causando falhas na montagem (BAKER *et al*, 2012).

A definição de parâmetros e medição da acurácia da montagem *de novo* de um genoma não é tarefa trivial. A tendência é, muitas vezes, otimizar o valor de N50, parâmetro usado para estimar o comprimento dos *contigs* montados. (SALZBERG *et al.*, 2012). Em geral, os montadores *de novo* trabalham com dois algoritmos principais: *overlap-layout-consensus* e grafo de *Bruijn* (POP; SALZBERG, 2008; ZHANG *et al*, 2011). Este é um algoritmo que quebra os segmentos de leitura em *k-mers*, ou seja, quebra das leituras em subleituras de tamanho fixo, antes de montá-los em *contigs*. A abordagem de grafos forma *contigs* ligando dois fragmentos (*k-mers*) com *k* ou mais nucleotídeos sobrepostos (SALZBERG *et al.*, 2012; NAGARAJAN; POP, 2013). Um exemplo de programas que utilizam esta abordagem é o SPAdes (BANKEVICH *et al.*, 2012). Desta forma, a anotação do genoma aponta as principais características da sequência genômica classificada em: automática (estática) ou manual (dinâmica) (STEIN, 2001). Na primeira classificação (automática) são utilizados diversos programas que realizam predições de regiões abertas para leituras - *Open Read Frames* (ORF'S), e localização de regiões codificantes de proteínas- *Coding Sequencing* (CDS), que podem estar presentes em ORF'S preditas *in silico*, além de RNA's ribossomais (rRNA) e transportadores (tRNA). Porém, estes podem não identificar pequenos genes, além disto, *drafts* de genomas depositados podem conter altas taxas de erros de sequenciamento, acarretando em predições equivocadas de genes (MÉDIGUE; MOSZER, 2007). Dependente da etapa automática, a anotação manual pode contribuir para melhorar a qualidade das anotações atribuídas.

O sistema *web* RAST é uma ferramenta que faz a anotação automática dos principais elementos estruturais do genoma (AZIZ *et al.*, 2008). A partir dos dados biológicos obtidos da anotação automática e manual é possível compreender as informações funcionais, ou seja, pode incorporar elementos relacionados às funções e processos biológicos que permitem distribuir os genes em classes funcionais, sendo a anotação um passo crucial. As análises comparativas podem auxiliar no conhecimento de genes alvos de interesse biomédico, ambiental, biotecnológico,

como outros. Através de análises pangenômicas, a qual consiste na comparação de várias linhagens de uma mesma espécie ou de diferentes espécies, possibilita o reconhecimento de dissimilaridades e diferenças existentes entre os genomas e esclarecer meios de patogenicidade do microrganismo.

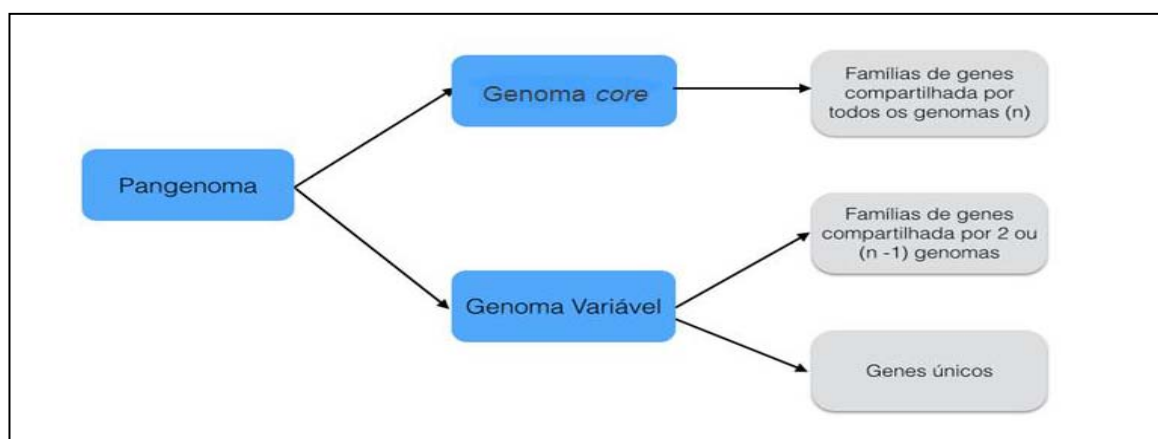
### 1.9 Genômica comparativa

Na filogenia os relacionamentos entre os organismos são representados por árvores filogenéticas compostas por ramos que conectam nós. A árvore filogenética pode ser construída utilizando famílias de genes diferentes (árvores de genes), onde os nós internos correspondem à duplicação gênica, à eventos de especiação ou pela combinação de ambos (WEISS, 2014).

Os métodos de construção de árvores filogenéticas estão divididos em duas categorias: matriz de distância e dados discretos, também conhecidos como métodos de procura de árvore. Como exemplos de métodos de matriz de distância destacam-se UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*), (SNEATH; SOKAL, 1973) *neighbor-joining* (SAITOU; NEI, 1987), *Fitch–Margoliash* (FITCH; MARGOLIASH, 1967); e de dados discretos destacam-se, *maximum parsimony* (DAY, 1987), *maximum likelihood* (FELSENSTEIN, 1985) e o método bayesiano (RANNALA; YANG, 1996). Enquanto os métodos de matrizes utilizam cálculos estatísticos para determinar a distância entre cada combinação de pares de nós para então agrupá-los em uma árvore, os métodos discretos examinam separadamente cada coluna lida de alinhamento buscando a árvore que melhor acomode toda informação (BALDAUF, 2003).

O pangenoma é o conjunto de todos os genes presentes nos genomas de isolados de uma espécie, podendo ser dividido em genoma core, composto pelo conjunto de genes comuns entre todos os genomas do grupo, e genoma variável ou acessório, composto pelo restante dos genes (exceto o genoma core). No genoma variável encontram-se tanto os genes presentes na maioria do conjunto, quanto os genes únicos de certos isolados (Figura 4) (JANNECHEVITZ *et al.*, 2015).

Figura 4 - Compartimentos do pangenoma



Fonte: JANNECHEVITZ *et al.*, 2015.

O pangenoma de uma espécie pode ser caracterizado como “aberto” (possibilidade do aumento de novos genes ao genoma core a medida que um novo genoma é inserido na análise) ou “fechado” (apresenta pouca ou nenhuma aquisição de novos genes) (MUZZI *et al.*, 2007; TETTELIN *et al.*, 2005). Nas análises de pangenoma, o PGAP, é um dos programas criados para avaliar a variação genética e evolutiva. Esta ferramenta possibilita identificar o genoma core dos singletons (genes únicos) (ZHAO *et al.*, 2012). O genoma core de bactérias do gênero *Streptococcus* foi analisado por Gao *et al.* (2014) e evidenciaram que este aumenta de acordo com a inclusão de novos genomas, o que comprova que este gênero possui um pangenoma aberto, ou seja, está em expansão (GAO *et al.*, 2014).

Além disso, o genoma acessório pode estar presente em Ilhas Genômicas (GEIs), potencialmente adquiridas por transferência horizontal (THG), o que leva a plasticidade, evolução e a adaptação de microrganismos, estas GEIs quando carregam genes de virulência são consideradas Ilhas de Patogenicidade (PAIs) (JUHAS *et al.*, 2009; MIRA *et al.*, 2010; SOARES *et al.*, 2012). Este é um importante mecanismo relacionado à grande variação encontrada em populações bacterianas garantindo a capacidade destes microrganismos se adaptarem a uma grande diversidade de ambientes (SOARES *et al.*, 2012). PAIs são responsáveis pela virulência e plasticidade genômica das bactérias patogênicas, propriedade dinâmica do genoma que pode envolver adição, perda ou rearranjo do DNA. Os genes de virulência, principal característica que distingue uma linhagem patogênica de uma avirulenta e estão inseridos de forma frequente nas PAIs (JUHAS *et al.*, 2009).

Atualmente, as PAIs ocupam regiões genômicas relativamente grandes e apresentam diferentes uso de códon, possuem eventos de deleção que ocorrem com frequências distintas, além de serem passíveis de transferência entre outras PAIs (GAL-MOR; FINLAY, 2006). Com intuito de identificar estas ilhas alguns programas são desenvolvidos, um exemplo é o programa GIPSy que determina ilhas de patogenicidade por meio de análises de desvio de uso de códon, conteúdo GC e fatores de virulência entre uma bactéria de espécie patogênica e outra não patogênica (SOARES *et al.*, 2012) e o programa BRIG possibilita a geração de imagens que mostram comparações entre diferentes genomas procariotos (ALIKHAN *et al.*, 2011).

A genômica comparativa pode esclarecer processos evolutivos que influenciam espécies ou populações bacterianas, podendo, por exemplo, identificar componentes do genoma com importantes papéis na virulência e/ou adaptação a diferentes hospedeiros (JANNECHEVITZ *et al.*, 2015). Devido ao pouco conhecimento do genoma deste patógeno, os mecanismos moleculares envolvidos na patogênese, além de genes de virulência, o sequenciamento da amostra GBS90356 pertencente ao TC III, ST-17 de *S. agalactiae* será descrito neste trabalho. Adicionalmente, diversas análises de genômica comparativa e predição de ilhas de patogenicidade foram realizadas utilizando os genomas disponíveis no banco de dados do NCBI.



## 2 OBJETIVOS

### 2.1 Objetivo geral

Implantar a PCR multiplex para tipagem capsular dos *S. agalactiae*, bem como realizar a análise genômica comparativa da amostra ST-17 de *S. agalactiae* (GBS90356) pertencente ao tipo capsular III, isolada de líquido de neonato.

### 2.2 Objetivos específicos:

- a) padronizar e implantar a técnica de PCR multiplex para tipagem de *S. agalactiae*;
- b) executar o sequenciamento do genoma *S. agalactiae* GBS90356 tipo capsular III, ST-17 de *S. agalactiae*;
- c) realizar a análise filogenômica da amostra GBS90356 com os genomas completos das demais amostras de *S. agalactiae* depositadas no NCBI;
- d) analisar o pangenoma da amostra GBS90356 de *S. agalactiae*;
- e) identificar os genes ortólogos e únicos entre a amostra GBS90356 e as demais amostras de *S. agalactiae* com os genomas completos depositados no NCBI;
- f) detectar as prováveis regiões de ilhas de patogenicidade na amostra GBS90356 de *S. agalactiae*;
- g) comparar as bases nucleotídicas dos genomas de *S. agalactiae* mapeando as PAIs através do anel comparativo gerado pelo software BRIG;
- h) realizar a curadoria dos genes preditos dentro das PAIs;
- i) avaliar as classes Gene Orthology dos genes preditos dentro das PAIs;

- j) verificar o papel da hemolisina na interação celular da amostra GBS90356 com célula endotelial humana;
- k) realizar a análise morfológica do perfil de interação celular por *time-lapse* na *Biostation*<sup>®</sup>.

### 3 METODOLOGIA

#### 3.1 Cultura bacteriana

As amostras de *Streptococcus agalactiae* (n=185) utilizadas nesse estudo fazem parte da Coleção do Laboratório de Biologia Molecular e Fisiologia de Estreptococos/IBRAG/DBCel. Realizou-se o cultivo, purificação e grupagem destas amostras conforme descrito por Lancefield (1934), assim como todos os testes presuntivos. Os genomas das amostras COH1 (TCIII/ST-17), NGBS128 (TCIII/ST-17); NEM316 (TCIII/ST-23) e A909 (TCIa/ST-7) (APÊNDICE A) foram utilizados como referência na compração de sintenia das ilhas de patogenicidade (PAIs).

As culturas bacterianas foram armazenadas a -80°C em alíquotas de meio líquido “Brain Heart Infusion” (BHI; Difco Laboratories, Detroit, MI, USA) contendo 20% de glicerol. As amostras foram crescidas em BHI e padronizadas para uma densidade ótica (D.O.) de 0.4 em comprimento de onda de 540 nm ( $\sim 1 \times 10^8$  unidades formadoras de colônias por mililitro [UFC/mL]). Para a interação celular as bactérias foram crescidas em BHI e BHI suplementado com glicose [10mg/mL] e padronizadas para  $\sim 1 \times 10^5$  UFC/mL.

#### 3.2 Oligonucleotídeos (*Primers*)

##### 3.2.1 Para tipagem capsular

Os *primers* específicos de cada TC permitiram a amplificação de fragmentos de diferentes tamanhos facilmente identificados por eletroforese em gel de agarose (Tabela 1) (POYART *et al.*, 2007). A especificidade do par de *primer* foi testada contra as sequências na base de dados GenBank utilizando pesquisas BLAST para verificar a ausência de possíveis semelhanças. Desta forma, foram selecionadas as sequências correspondentes aos tipos Ia, Ib, II, III, IV, V, VI, VII e VIII com base na

temperatura de fusão e na capacidade de gerar tamanhos de *amplicon* distinguíveis, medidos em pares de base (pb).

Tabela 1 - *Primers* para tipagem capsular de *S. agalactiae*

Nome do primer	Sequência (5' - 3')	Tamanho amplicon (pb)	GenBank
Ia-F	GGTCAGACTGGATTAATGGTATGC		AB028896
Ia-R	GTAGAAATAGCCTATATACGTTGAATGC	521	
Ib-F	TAAACGAGAATGGAATATCACAAACC		AB050723
Ib-R	GAATTAACCTCAATCCCTAAACAATATCG	770	
II-F	GCTTCAGTAAGTATTGTAAGACGATAG		AY375362
II-R	TTCTCTAGGAAATCAAATAATTCTATAGGG	397	
III-F	TCCGTACTACAACAGACTCATCC		AF163833
III-R	AGTAACCGTCCATACATTCTATAAGC	1,826	
IV-F	GGTGGTAATCCTAAGAGTGAAGTGT		AF355776
IV-R	CCTCCCAATTTTCGTCCATAATGGT	578	
V-F	GAGGCCAATCAGTTGCACGTAA		AF349539
V-R	AACCTTCTCCTTCACACTAATCCT	701	
VI-F	GGACTTGAGATGGCAGAAGGTGAA		AF337958
VI-R	CTGTCCGACTATCCTGATGAATCTC	487	
VII-F	CCTGGAGAGAACAATGTCCAGAT		AY376403
VII-R	GCTGGTCGTGATTTCTACACA	371	
VIII-F	AGGTCAACCACTATATAGCGA		AY375363
VIII-R	TCTCAAATTCCGCTGACTT	282	

Fonte: POYART *et al.*, 2007.

### 3.3 Extração de DNA

#### 3.3.1 Para tipagem capsular

Foram semeadas de 5 a 10 colônias em tubos falcon de 15 mL de BHI e incubadas a 37°C por 18h. Posteriormente, as culturas foram centrifugadas a 10.000 rpm por 15min a 4°C, o sedimento foi suspenso em 1mL de TE (Tris HCl 0,1M e

EDTA 1mM) e solução de lise (lisozima [20mg/mL] + mutanolisina [10U/ $\mu$ L] + Brij 0,5% + RNase [10mg/mL] + TE) e incubado a 37°C *overnight*. A seguir, foi adicionada a solução de degradação proteica (protease [10mg/mL] + SDS 10%), incubação a 37°C por 1h e adição de acetato de sódio 3M e etanol PA a -20°C por 30min.

Logo, foram centrifugadas e descartada a fase aquosa, o DNA foi ressuscitado com 100 $\mu$ L de água bidestilada e, então, adicionado 1mL de etanol 70% gelado, transferiu-se para tubos eppendorfs de 2 mL e então, foram centrifugados e invertidos os tubos para secagem na estufa a 37°C. Então o DNA foi ressuscitado com água bidestilada estéril (50uL) e estocado a -20°C, até a sua quantificação (Nanovue-GE) e posterior utilização.

### 3.3.2 Para Sequenciamento de alto desempenho

O DNA genômico da amostra GBS90356 foi extraído utilizando o kit *PureLink Genomic DNA Mini Kit* (Invitrogen), conforme as instruções do fabricante, em colaboração com o laboratório de Coleção de Bactérias do Ambiente e Saúde, localizada no Instituto Oswaldo Cruz/Fiocruz.

### 3.4 **PCR-multiplex na tipagem capsular de *S. agalactiae***

A padronização e implantação da técnica de PCR multiplex para tipagem capsular das amostras de *S. agalactiae* foi realizada através de diversas mudanças em temperaturas e ciclagens até a obtenção da padronização ideal. Um total de 35 ciclos foram realizados para desnaturação, anelamento e amplificação. A tipagem capsular por PCR-multiplex foi validada com a execução da técnica de PCR convencional. Uma alíquota de 10  $\mu$ L de cada produto de PCR e seus controles foi separada por eletroforese em gel de agarose para confirmar a amplificação bem sucedida. Porém, a metodologia encontra-se em processo de obtenção de patente não podendo ser totalmente detalhada.

### 3.4.1 PCR convencional na tipagem capsular de *S. agalactiae*

A confirmação da tipagem capsular baseada na PCR Multiplex foi realizada por PCR convencional (POYART *et al.*, 2007). Resumidamente, um volume de 50µL de reação final (200µM cada dNTP, 5u/µL GoTaq polimerase, tampão GoTaq 5x + MgCl<sub>2</sub> [1.5mM] e 50ng DNA) com os determinados *primers* foi colocada em um termociclador (VERITI 96 Well Thermal Cycler, AppliedBio), onde as temperaturas e tempos utilizados foram: desnaturação a 96°C por 2min, anelamento de 48 a 58°C por 1 min, dependendo da temperatura de *Melting* dos *primers*, e amplificação a 72°C de 30seg a 2min de acordo com tamanho dos amplicons, por 35 ciclos. Um total de 10 µL de cada produto de PCR e seus controles foi separado por eletroforese em gel de agarose (1,2-1,5%) para confirmar a amplificação bem sucedida.

### 3.5 Eletroforese horizontal em gel de agarose

Os produtos amplificados na PCR-Multiplex e PCR convencional foram separados por eletroforese em gel de agarose de 1 a 2% (de acordo com tamanho do *amplicon*) diluída em TBE 1x (Tris [89mM]; EDTA [2mM] e ácido bórico [89mM]), corados com SYBR Safe (Green) (Invitrogen) e visualizados como bandas brancas sobre um fundo negro por transiluminação UV. Como padrão de tamanho molecular foi utilizado 1-kb Plus DNA (Invitrogen [1µg/µL]). O controle negativo foi incluído em cada ensaio, constituído pelo mesmo mix de reação, mas com água bidestilada em vez de DNA. A corrida dos géis foi realizada em aparelho de eletroforese Biorad-Power Pac Basic em cuba horizontal 11.14 (Gibco-BRL), contendo tampão de transferência TBE 1x por 90 min a 100 V e temperatura ambiente.

### 3.6 Sequenciamento do genoma da amostra GBS90356

O genoma da amostra GBS90356 foi sequenciado utilizando o sequenciador Illumina HiSeq 2500, na plataforma de sequenciamento de alto desempenho do Instituto Oswaldo Cruz/Fiocruz. O DNA genômico foi fragmentado aleatoriamente em segmentos de aproximadamente 500 pb para a construção da biblioteca utilizando o kit Nextera XT (Illumina). Em ambas as extremidades dos fragmentos de 500 pb foram adicionados os adaptadores (*paired-ends*) usados nas etapas de quantificação, amplificação e sequenciamento.

### 3.7 Avaliação da qualidade das *reads*, montagem e anotação do genoma da amostra GBS90356

A qualidade das *reads* geradas pelo sequenciador Illumina HiSeq 2500 em formato fastq foi analisada através do programa FastQC, versão 0.11.2 (ANDREWS, 2010). As *reads* foram filtradas e tratadas empregando o programa Trimmomatic 0.36 (BOLGER; LOHSE; USADEL, 2014). Esta etapa teve como finalidade a exclusão de adaptadores internos ou sequências de baixa qualidade. Após este processo, as *reads* já filtradas foram utilizadas para a montagem do genoma, utilizando o programa SPADES (BANKEVICH *et al.*, 2012). Os *contigs* (regiões formadas pela sobreposição de dois ou mais *reads*) gerados na montagem dos genomas foram alinhados e ordenados no programa Mauve versão 2.0 (DARLING; MAU; PERNA, 2010), usando o genoma referência de *Streptococcus agalactiae* 2603V/R (GenBank AE009948.1). O melhor alinhamento gerado pelo programa Mauve em formato .fasta foi utilizado para a anotação automática dos genes através da ferramenta *online* RAST (*Rapid Annotation using Subsystem Technology*) versão 2.0 (AZIZ *et al.*, 2008).

### 3.8 Obtenção das sequências dos genomas completos de *S. agalactiae* em banco de dados

A busca de sequências completas dos genomas depositados no NCBI de 108 amostras de *Streptococcus agalactiae* foi realizada no GenBank/NCBI com auxílio do aplicativo Blastn (APÊNDICE A), extraídas em diferentes formatos (*embl*, *fasta*, *gbk*) para as análises posteriores.

### 3.9 Ferramentas de bioinformática

#### 3.9.1 Análise filogenômica e cálculo do Pangenoma de *Streptococcus agalactiae*

O valor de  $\alpha$  da Lei de Heap para caracterizar o nível de plasticidade genômica de *S. agalactiae* foi calculado com base nos resultados do pangenoma obtido em PGAP V.1.11, através da ferramenta *online* PanWeb (PANTOJA *et al.*, 2017). Todos os genomas completos da espécie depositados no GenBank foram extraídos em formato *gbk* e submetidos ao PanWeb. O parâmetro de identidade e de cobertura da sequência foi de 50% e o valor de E utilizado foi de 0,001. Esta análise permite identificar genes compartilhados entre as linhagens, genes acessórios compartilhados por duas ou mais linhagens e genes espécie específicos. O dendograma Neighbor-Joining também foi calculado a partir dos genes do genoma central preditos no PGAP. Cento e oito (108) genomas disponíveis na base de dados NCBI mais o genoma sequenciado neste estudo foram utilizados, sendo 42 genomas de amostras humanas, 55 de peixes, 3 de bovinos, 2 de camelo, 1 de sapo, 1 de cachorro e 5 isolados desconhecidos (APÊNDICE A).



### 3.9.2 Predição das Ilhas de Patogenicidade (PAI) e profagos

As PAIs foram identificadas com o software GIPSy v1.1.2 (SOARES *et al.*, 2015). A predição das PAIs foi confirmada através da ferramenta IslandViewer4 usando todos os quatro métodos: IslandPick, SIGI-HMM, IslandPath-DIMOB and Islander (BERTELLI *et al.*, 2017). Diversas ferramentas têm sido desenvolvidas para detecção *in silico* de potenciais PAIs, incluindo programas que integram múltiplos recursos das ilhas, como composição de sequências atípica e genes associados à transferência gênica horizontal (GAL-MOR; FINLAY, 2006; HSIAO *et al.*, 2005). GIPSy apresentou os melhores resultados globais do que o IslandViewer4 (SILVA-FILHO *et al.*, 2018) e comparado aos demais softwares, o GIPSy é o que apresenta mais de 90% de acurácia na predição destas ilhas (SILVA-FILHO *et al.*, 2018). Assim, os resultados do GIPSy foram considerados neste trabalho. O GIPSy prediz ilhas genômicas utilizando uma metodologia híbrida que inclui viés de conteúdo GC, códon usage e outras características (SOARES *et al.*, 2015). Para predição das ilhas no GIPSy, uma bactéria não patogênica e taxonomicamente relacionada ao *S. agalactiae* foi utilizada como referência, a *S. termophilus* JIM8232 (número de acesso do NCBI: NC\_017581). O valor de E utilizado foi de 0,000001 com cobertura acima de 90% para garantir a alta identidade entre as sequências. A ferramenta de navegação e anotação de genoma, Artemis foi utilizada para revisar as anotações automáticas, assim como anotar possíveis PAIs e os genes analisados no UNIProt (CARVER *et al.*, 2012). Os profagos foram preditos na ferramenta web PHASTER (Phage Search Tool) (ARNDT *et al.*, 2016).

### 3.9.3 Sintenia das PAIs e comparação nucleotídica

As ilhas de patogenicidade da GBS90356 foram comparadas contra as ilhas das linhagens pertencentes ao TCs III (COH1, NGBS128 e NEM316) e Ia (A909) para determinar a sintenia dos genes. Esta análise foi realizada na ferramenta web Simple Synteny (VELTRI *et al.*, 2016) utilizando BLASTn com um *cutoff* de cobertura mínima de 50%. As figuras de sintenia foram exportadas em arquivos jpeg.

Para comparação da sequência de nucleotídeos dos genomas, o software Blast Ring Image Generator (BRIG) foi utilizado (ALIKHAN *et al.*, 2011). Treze (13) genomas completos foram incluídos na comparação: 7 genomas do TC III (incluindo 6 genomas do complexo clonal hipervirulento, ST-17), além de um genoma para cada TC Ia, Ib, II, IV, V e VI. O nível de identidade entre os 13 genomas foi representado por uma série de anéis concêntricos utilizando como referência para o BLASTn a amostra GBS90356. Regiões do genoma com identidade acima de 90% foram representadas por cores mais intensas, enquanto regiões com identidade menor que 70% não apresentaram cor, formando um gap. As sequências das ilhas preditas pelo GIPSy e corroboradas pelo IslandViewer4 foram incluídas na análise.

Para visualizar a conservação dos genes de patogenicidade de *S. agalactiae* GBS90356 nos anéis comparativos calculados, um arquivo gbk contendo o nome dos genes foi adicionado na análise.

#### 3.9.4 Enriquecimento funcional dos genes das PAIs e predição do secretoma

Todos os genes contidos dentro das ilhas preditas no genoma da linhagem GBS90356 foram extraídos utilizando o software Artemis e salvos como um arquivo multi-fasta. Para determinar sua função, estes genes foram analisados no programa Blast2GO (CONESA *et al.*, 2005) para atribuição de termos do Gene Ontology (GO) nas categorias: função molecular, componente celular e processo biológico. Adicionalmente, o software SignalP v.4.1 (BENDTSEN *et al.*, 2004) foi utilizado para predição de peptídeo sinal na sequência de aminoácidos destes genes. Proteínas com peptídeo sinal são alvos para vias de secreção, mas não são necessariamente secretadas (NIELSEN, 2017).

#### 3.10 **Cultura de células endoteliais da veia umbilical humana (HUVEC)**

Células endoteliais da veia umbilical humana (HUVEC) foram extraídas e cultivadas de acordo com a técnica descrita por Jaffe *et al.* (1973) com modificações.

Cordões umbilicais foram adquiridos no Hospital Maternidade Carmela Dutra, até 24 h após o parto e mantidos em tampão fosfato salina 0,1 M (PBS) estéril + glicose 0,1 M, pH 6.5 (PBS-glicose), acrescido de penicilina 500 U/mL, estreptomicina 40 µg/mL (ambas Gibco, BRL, Grand Island, NY, EUA) e fungizona 2,5 µg/mL (Bristol Myers Squibb Brasil, SP, Brasil) até o momento da cultura.

As HUVEC foram obtidas da veia de cordões umbilicais humanos a partir do tratamento com 0,1% colagenase II (Sigma Cehm Co, St Louis, EUA) diluída em PBS-glicose com o cordão submerso por 10 min em salina 0,9% a 37 °C. A reação foi paralisada com a adição de meio 199 (M199 - Sigma Chem Co., St Louis, EUA) contendo 20% de soro fetal bovino (SFB - Cultilab, Campinas, Brasil), L-glutamina 2 mM, bicarbonato de sódio 0,026 mM, fungizona 2,5 µg/mL, penicilina 500 U/mL, estreptomicina 40 µg/mL (M199 suplementado). A suspensão resultante foi centrifugada (120 x g por 10 min) e o sedimento contendo as células foi ressuspensão em M199 suplementado e semeado em garrafas de 25 cm<sup>2</sup> previamente cobertas com gelatina 2% e mantido em atmosfera de 5% CO<sub>2</sub> a 37 °C até atingirem a confluência.

A partir das culturas primárias foram realizados subcultivos até no máximo a segunda passagem, onde as monocamadas confluentes de HUVEC foram lavadas duas vezes com PBS-glicose, tratadas com solução de 0,025% tripsina (Sigma Cehm Co, St Louis, EUA) feita em 0,025%, NaCl 0,14 M, Na<sub>2</sub>HPO<sub>4</sub> 9 mM, KCl 3 mM, vermelho de fenol 0,02% e ácido etileno-diamino-tetra-acético (EDTA) 0,02% até o descolamento das células. A reação foi paralisada com M199 suplementado, seguida de uma nova centrifugação (120 x g por 10 min) e a suspensão do sedimento celular feita em M199 suplementado para cultivo em garrafas 75 cm<sup>2</sup> ou placas de 24 poços cobertas com gelatina 2% para realização dos experimentos (SANTOS *et al.*, 2005).

### 3.11 Ensaios de interação entre *S. agalactiae* e HUVEC

Monocamadas confluentes de HUVEC em placas de 24 orifícios (2x10<sup>5</sup> células por poço) foram lavadas duas vezes com tampão PBS-glicose antes dos ensaios de interação. Após as lavagens, as monocamadas foram infectadas com a

amostra bacteriana, na proporção de 1 célula eucariótica para 25 células bacterianas (MOI 1:25) em meio 199 puro. Para verificar a inibição da atividade da  $\beta$ - hemolisina, previamente detectada por PCR na amostra GBS90356, o meio foi suplementado com glicose [10mg/mL], de acordo com Ulett et al. (2004).

As condições analisadas nesse estudo foram: i) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI; ii) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI + glicose [10mg/mL]; iii) HUVEC em meio 199 + glicose [10mg/mL] infectada com GBS90356 crescidas em BHI + glicose [10mg/mL]. A incubação foi realizada em diferentes intervalos de tempo para todas as condições analisadas (0h, 1h, 2h e 4h) a 37 °C em ambiente de 5% de CO<sub>2</sub>. Após cada período de incubação foi realizada a lavagem das HUVEC infectadas com solução salina estéril 0,9% para remoção dos microrganismos não aderidos.

Para análise do número total de bactérias interagidas (aderidas e internalizadas), as HUVEC foram lisadas com tampão de lise contendo Tris 25 mM pH 7.5, EDTA 5 mM, NaCl 150 mM e Igepal 1% (Sigma Chem Co., St Louis, EUA). O lisado foi diluído em salina 0,9% e plaqueado em meio agar sangue para a contagem do número de unidades formadoras de colônias (UFC/mL).

Paralelamente, em outra série de poços, para a quantificação da invasão do *S. agalactiae* nas HUVEC, após lavagens com salina 0,9% foi feita administração de gentamicina (100  $\mu$ g/mL) e penicilina (5  $\mu$ g/mL) por 2 h a 37 °C em 5% de CO<sub>2</sub> para eliminar as bactérias aderidas. Após esta etapa foi realizada a lavagem da monocamada com salina 0,9% e posterior lise da célula eucariótica com o mesmo tampão de lise descrito acima. O lisado foi diluído em solução salina e plaqueado em meio agar sangue para a obtenção do número de bactérias intracelulares viáveis (PEOTTA *et al.*, 2001).

O número de *S. agalactiae* aderido às células endoteliais foi determinado através da subtração do número total de bactérias interagidas pelo número de bactérias intracelulares viáveis (SANTOS *et al.*, 2003). Todos os ensaios de aderência e invasão foram realizados em triplicata e em pelo menos seis réplicas biológicas independentes. Os dados foram plotados no GraphPadPrism versão 7.0, os percentuais de aderência e invasão foram determinados e desvio padrão calculados.

### 3.12 Análise morfológica por *time-lapse* da interação celular (HUVEC X *S. agalactiae*)

Monocamadas confluentes de HUVEC ( $2 \times 10^5$  células por poço) em placa de cultura HiQ-4 (Nikon), a qual apresenta quatro divisões e possibilita a observação conjunta de até 4 condições distintas, foram lavadas duas vezes com tampão PBS-glicose antes dos ensaios de interação. Após as lavagens, as HUVEC foram infectadas com GBS90356 na proporção de 1 célula eucariótica para 10 células bacterianas (MOI 1:10) em meio 199 puro ou suplementado com glicose [10mg/mL], dependendo da condição analisada.

Durante esse período foi realizado o monitoramento das células pelo tempo de 24h, através da utilização do sistema de imagens *time-lapse* BioStation IM-Q (Nikon). As condições analisadas foram as mesmas descritas no item 2.12. As fotos obtidas foram utilizadas para análise da morfologia celular do perfil de interação de *S. agalactiae* GBS90356 em HUVEC. Os experimentos foram realizados em triplicata.

## 4 RESULTADOS

### 4.1 Tipagem capsular de *Streptococcus agalactiae* por PCR multiplex

Os resultados da tipagem capsular por PCR multiplex das 185 amostras de *S. agalactiae* estão representados nas Tabelas 2, 3 e 4 e Figuras 5 e seção 4.3.1 (Figura 6). Todas as amostras foram testadas para cada par de *primer* pertencentes aos nove tipos capsulares pesquisados. Das 43 amostras de *S. agalactiae* de origem humana não oncológica, 41,86% pertenceram ao TC III, seguido pelos TC V (20,93%), Ia e II (9,3% cada), Ib (4,65%) e VIII (2,33%). Contudo, amostras isoladas de pacientes oncológicos apresentaram perfil diferente, sendo 44% pertencem ao TC Ia (24/55), seguido de 24% do TC V (13/55), 15% do TC II (8/55), 11% do III (6/55), 7,27% do TC IV (4/55), 2% do TC VI (1/55) e 2% do TC VII (1/55). Não foram detectados os TC Ib e VIII.

Com relação as 87 amostras de *S. agalactiae* oriundas de mastite bovina, os TC prevalentes foram: II (48,28%), III (29,89%), V (10,34%) e VII (1,15%). Neste trabalho, o tipo IX não foi investigado, pois encontra-se em fase de testes, podendo explicar o número de amostras de *S. agalactiae* não tipáveis de 7,5% até o presente estudo.

Tabela 2 - Tipagem capsular de amostras humanas de *S. agalactiae* isoladas de pacientes não oncológicos por PCR multiplex

<i>S. agalactiae</i>	Material clínico	Origem	Tipo capsular
GBS90356	LCR <sup>1</sup> /neonatal	CCLCP <sup>2</sup>	III
80340	Vagina/adulto	CCLCP	III
88641	Vagina/adulto	CCLCP	V
90186	Sangue/adulto	CCLCP	V
90222	Urina/adulto	CCLCP	Ia
85147	Orofaringe/adulto	CCLCP	Ia
1501	Urina/gestante	LBMFE <sup>3</sup>	III
1502	Urina/gestante	LBMFE	III
1601	Urina/adulto	LBMFE	VIII
50B	Perianal/adulto	HUPE <sup>4</sup>	Ia
51B	Perianal/ adulto	HUPE	Ia
60B	Perianal/adulto	HUPE	Ib
54A	Vagina/adulto	HUPE	Ib
4401	Vagina/adulto	HUPE	II
4488	Vagina/adulto	HUPE	II
4539	Vagina/ adulto	HUPE	II
82A	Vagina/adulto	HUPE	III
65A	Vagina/adulto	HUPE	III
66A	Vagina/ adulto	HUPE	III
89A	Vagina/adulto	HUPE	III
51	Perianal/adulto	HUPE	III
109B	Perianal/adulto	HUPE	III
41B	Perianal/adulto	HUPE	III
79A	Vagina/adulto	HUPE	III
109A	Vagina/ adulto	HUPE	III
43A	Vagina/adulto	HUPE	III
85A	Vagina/adulto	HUPE	III
72A	Vagina/adulto	HUPE	III
39A	Sangue/neonato	HUPE	III
83A	Vagina/ adulto	HUPE	III
81779	Vagina/ adulto	HUPE	V
80427	Vagina/adulto	HUPE	V
5468	Perianal/adulto	HUPE	V
5467	Vaginal/ adulto	HUPE	V
83174	Vagina/ adulto	HUPE	V
5459	Vagina/ adulto	HUPE	V
4421	Vagina/ adulto	HUPE	V
4538	Vagina/adulto	HUPE	NT
5285	Vagina/adulto	HUPE	NT
6468	Perianal/adulto	HUPE	NT
84A	Vagina/adulto	HUPE	NT
4622	Vagina/adulto	HUPE	NT
1458	Vagina/adulto	HUGG <sup>5</sup>	II

Legenda: <sup>1</sup>Líquido cefalorraquidiano; <sup>2</sup> Coleção de Cultura do Laboratório de Cocos Patogênicos/UFRJ; <sup>3</sup> Laboratório de Biologia Molecular e Fisiologia de Estreptococos; <sup>4</sup> Hospital Universitário Pedro Ernesto; <sup>5</sup> HUGG: Hospital Universitário Gaffrée e Guinle.

Fonte: A autora, 2019.

Tabela 3 - Tipagem capsular de amostras humanas de *S. agalactiae* isoladas de pacientes oncológicos por PCR multiplex

<i>S. agalactiae</i>	Material clínico	Origem	Tipo capsular
1301	Sangue	INCA-HCI <sup>1</sup>	II
1302	Urina	INCA-HCI	Ia
1303	Urina	INCA-HCI	Ia
1304	Urina	INCA-HCI	Ia
1305	Urina	INCA-HCI	Ia
1306	Urina	INCA-HCI	V
1307	Urina	INCA-HCI	Ia
1308	Urina	INCA-HCI	V
1401	Urina	INCA-HCI	Ia
1402	Sangue	INCA-HCI	V
1403	Swab de lesão	INCA-HCI	Ia
1404	Secreção traqueal	INCA-HCI	II
1405	Urina	INCA-HCI	Ia
1406	Urina	INCA-HCI	V
1407	Sangue	INCA-HCI	V
1408	Urina	INCA-HCI	Ia
1409	Urina	INCA-HCI	Ia
1410	Urina	INCA-HCI	III
1411	Urina	INCA-HCI	III
1412	Sangue	INCA-HCI	III
1414	Abscesso pulmonar	INCA-HCI	Ia
1416	Urina	INCA-HCI	IV
1417	Secreção biliar	INCA-HCI	VII
1418	Urina	INCA-HCI	Ia
1419	Urina	INCA-HCI	III
1422	Sangue	INCA-HCI	Ia
1423	Urina	INCA-HCI	II
1425	Sangue	INCA-HCI	Ia
1427	Secreção traqueal	INCA-HCI	Ia
1428	Urina	INCA-HCI	III
1429	Urina	INCA-HCI	Ia
1430	Urina	INCA-HCI	Ia
1431	Não informado	INCA-HCI	IV
1432	Não informado	INCA-HCI	V
1433	Urina	INCA-HCI	Ia
1434	Urina	INCA-HCI	VI
1435	Sangue	INCA-HCI	Ia
1436	Urina	INCA-HCI	V
1437	Urina	INCA-HCI	V
1438	Urina	INCA-HCI	III
1439	Urina	INCA-HCI	Ia
1440	Urina	INCA-HCI	Ia
1441	Sangue	INCA-HCI	II
1442	Secreção traqueal	INCA-HCI	V
1443	Urina	INCA-HCI	V
1444	Urina	INCA-HCI	Ia
1445	Urina	INCA-HCI	II
1446	Secreção biliar	INCA-HCI	II
1501	Urina	INCA-HCI	Ia
1502	Urina	INCA-HCI	Ia
1503	Urina	INCA-HCI	V
1504	Urina	INCA-HCI	V
1505	Urina	INCA-HCI	II
1506	Abscesso cutâneo	INCA-HCI	II
1507	Urina	INCA-HCI	V

Legenda: <sup>1</sup> Instituto Nacional de Câncer.

Fonte: A autora, 2019.



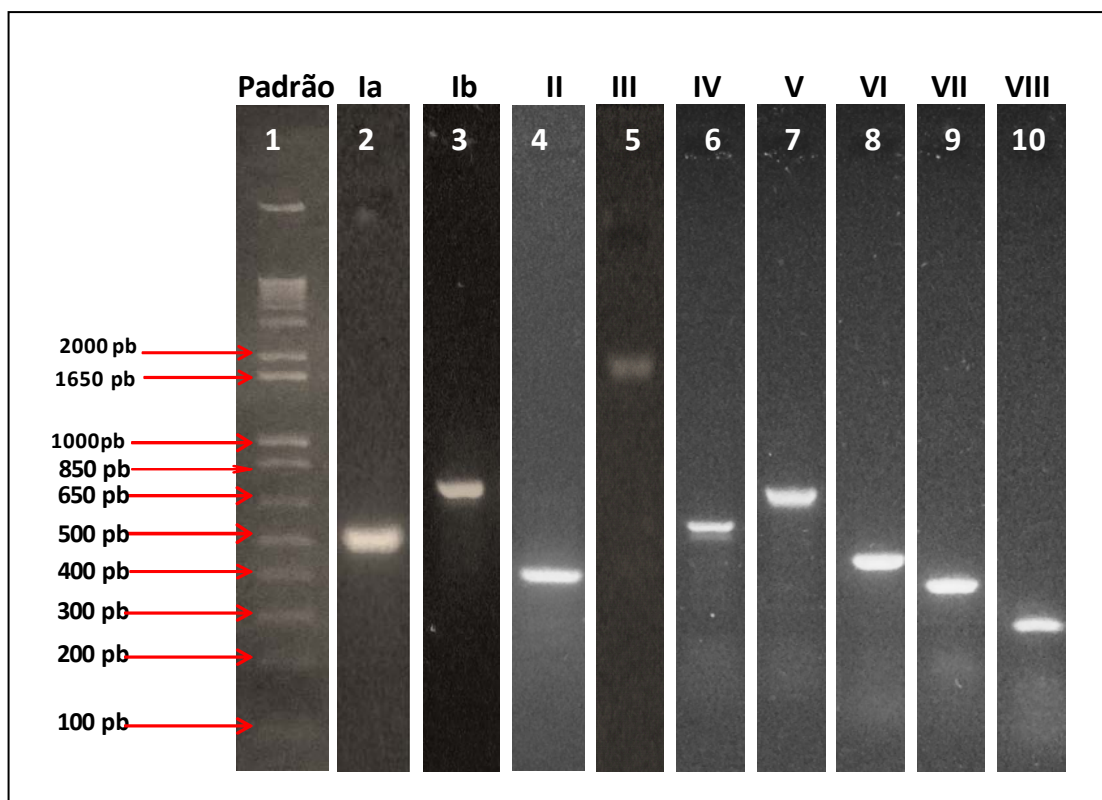
Tabela 4- Tipagem capsular de amostras bovinas de *S. agalactiae* por PCR multiplex

<i>S. agalactiae</i>	MATERIAL CLÍNICO	Origem	Tipo capsular
A1201	Mastite subclínica	DMUFRP <sup>1</sup>	II
A1202	Mastite subclínica	DMUFRP	II
A1203	Mastite subclínica	DMUFRP	II
A1204	Mastite subclínica	DMUFRP	II
A1205	Mastite subclínica	DMUFRP	II
A1211	Mastite subclínica	DMUFRP	II
A1212	Mastite subclínica	DMUFRP	II
A1214	Mastite subclínica	DMUFRP	II
A1215	Mastite subclínica	DMUFRP	II
A1220	Mastite subclínica	DMUFRP	II
A1222	Mastite subclínica	DMUFRP	II
A12802	Mastite subclínica	DMUFRP	II
A1287	Mastite subclínica	DMUFRP	II
A12109	Mastite subclínica	DMUFRP	II
A12110	Mastite subclínica	DMUFRP	II
A12159	Mastite subclínica	DMUFRP	II
A1302	Mastite subclínica	UNESP <sup>2</sup>	V
A1310	Mastite subclínica	UNESP	V
A1311	Mastite subclínica	UNESP	III
A1312	Mastite subclínica	UNESP	II
A1313	Mastite subclínica	UNESP	III
A1314	Mastite subclínica	UNESP	II
A1316	Mastite subclínica	UNESP	II
A1319	Mastite subclínica	UNESP	II
A1320	Mastite subclínica	UNESP	II
A1322	Mastite subclínica	UNESP	II
A1323	Mastite subclínica	UNESP	V
A1324	Mastite subclínica	UNESP	II
A1326	Mastite subclínica	UNESP	II
A1327	Mastite subclínica	UNESP	V
A1328	Mastite subclínica	UNESP	III
A1330	Mastite subclínica	UNESP	II
A1331	Mastite subclínica	UNESP	II
A1334	Mastite subclínica	UNESP	V
A1335	Mastite subclínica	UNESP	II
A1336	Mastite subclínica	UNESP	III
A1337	Mastite subclínica	UNESP	II
A1338	Mastite subclínica	UNESP	II
A1339	Mastite subclínica	UNESP	V
A1340	Mastite subclínica	UNESP	III
A1344	Mastite subclínica	UNESP	V
A1345	Mastite subclínica	UNESP	II
A1351	Mastite subclínica	UNESP	V
A1352	Mastite subclínica	UNESP	V
A1360	Mastite subclínica	UNESP	III
A1366	Mastite subclínica	UNESP	III
A1368	Mastite subclínica	UNESP	II
A1369	Mastite subclínica	UNESP	III
A1370	Mastite subclínica	UNESP	II
A1371	Mastite subclínica	UNESP	III
A1372	Mastite subclínica	UNESP	II
A1373	Mastite subclínica	UNESP	II
A1374	Mastite subclínica	UNESP	II
A1377	Mastite subclínica	UNESP	II
A1378	Mastite subclínica	UNESP	II
A1379	Mastite subclínica	UNESP	II
A1390	Mastite subclínica	UNESP	II
A1391	Mastite subclínica	UNESP	II
A1392	Mastite subclínica	UNESP	II
A1395	Mastite subclínica	UNESP	II
A1396	Mastite subclínica	UNESP	III
A1397	Mastite subclínica	UNESP	VII
A1398	Mastite subclínica	UNESP	III
A18101	Mastite subclínica	UNESP	III
A13102	Mastite subclínica	UNESP	III
A18106	Mastite subclínica	UNESP	III
A13107	Mastite subclínica	UNESP	NT
A18108	Mastite subclínica	UNESP	III
A13114	Mastite subclínica	UNESP	III
A18118	Mastite subclínica	UNESP	III
A13119	Mastite subclínica	UNESP	III
A13120	Mastite subclínica	UNESP	III
A13121	Mastite subclínica	UNESP	III
A13123	Mastite subclínica	UNESP	III
A13124	Mastite subclínica	UNESP	III
A13125	Mastite subclínica	UNESP	III
A13126	Mastite subclínica	UNESP	III
A13128	Mastite subclínica	UNESP	NT
A13129	Mastite subclínica	UNESP	III
A13130	Mastite subclínica	UNESP	III
A13131	Mastite subclínica	UNESP	NT
A13136	Mastite subclínica	UNESP	NT
A13137	Mastite subclínica	UNESP	NT
A13138	Mastite subclínica	UNESP	NT
A13139	Mastite subclínica	UNESP	NT
A13140	Mastite subclínica	UNESP	NT
A13141	Mastite subclínica	UNESP	NT

Legenda: <sup>1</sup> Departamento de Medicina da Universidade Federal Rural de Pernambuco; <sup>2</sup> Universidade Estadual Paulista-Botucatu.

Fonte: A autora, 2019.

Figura 5 - Eletroforese dos produtos da tipagem capsular de *S. agalactiae*



Legenda: Modelo da amplificação dos produtos de cada tipo capsular de *S. agalactiae*, onde na *lane* 1- Padrão (1kb); 2- Ia (521pb); 3- Ib (770pb); 4- II (397pb); 5- III (1,826pb); 6- IV (578pb); 7- V (701pb); 8- VI (487pb); 9- VII (371pb) e 10- VIII (282pb).

Fonte: A autora, 2019.

#### 4.2 Características genômicas do *S. agalactiae* GBS90356

O sequenciamento genômico na plataforma Illumina HiSeq 2500 da amostra de *S. agalactiae* ST-17 isolada na década de 90 no Brasil, GBS90356, resultou em um genoma com 40 contigs, tamanho de 2.029.093 pb, N50 de 124.554, 2.023 CDSs, 52 genes tRNA e um conteúdo de G + C de 35,2% (Tabela 5). Estas características estão de acordo com outros genomas de *S. agalactiae* sequenciados e depositados em bancos de dados públicos. Nenhum profago completo foi identificado no genoma. Após a predição automática utilizando a ferramenta *web* RAST, uma edição manual dos elementos estruturais do genoma foi realizada no programa Artemis. Um total de 1588 (78,5%) CDSs foi classificado em subsistemas

do servidor RAST. As categorias mais representativas foram: Carboidratos; Metabolismo Protéico; Aminoácidos e Derivados; Parede Celular e Cápsula; Metabolismo de DNA; Cofactores, Vitaminas, Grupos Protéticos, Pigmentos; Virulência, Doença e Defesa. As categorias de subsistemas importantes para virulência bacteriana envolveram 14,2% das CDSs preditas como: Parede celular e Cápsula; Virulência, Doença e Defesa; Fagos, Propágios, Elementos Transponíveis e Plasmídeos e Resposta ao estresse.

Tabela 5 - Métricas do sequenciamento na plataforma Illumina HiSeq 2500

	GBS90356
Tamanho	2.029.093
Conteúdo GC %	35,21
Número de RNAs	53
tRNA	48
Números de <i>contigs</i>	40
CDS	2.030
N50	124.554

Fonte: A autora, 2019.

### 4.3 Análise comparativa de *S.agalactiae*: GBS90356 ST-17 no Brasil

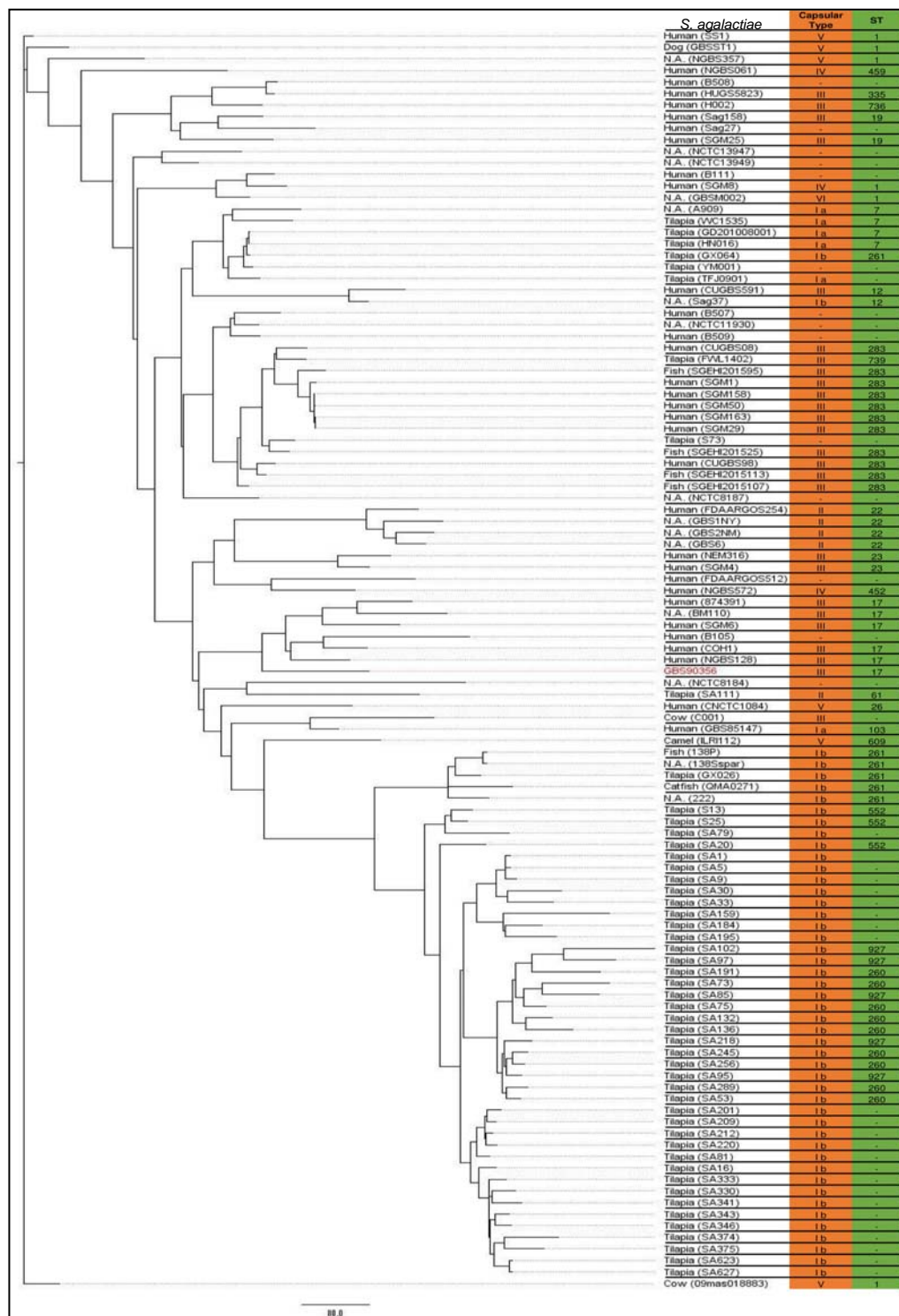
A análise comparativa demonstrou a heterogeneidade da amostra sequenciada, GBS90356, frente às demais amostras de diferentes tipos capsulares com genoma completo depositados no NCBI, assim como entre as amostras pertencentes ao tipo capsular III, apresentados nos tópicos a seguir.

#### 4.3.1 Análise filogenômica e predição de profagos

Para refletir com maior precisão a relação evolutiva dentro das amostras de *S. agalactiae*, uma árvore filogenômica foi construída. Amostras bacterianas da mesma sequência tipo mostraram relações evolutivas próximas, porém, entre os TC, apenas o TC Ib apresentou maior proximidade evolutiva. As ramificações de amostras ST-17 humanas foram agrupadas, indicando um gargalo evolutivo. As amostras bovinas e humanas foram alternadamente distribuídas na árvore evolutiva (Figura 6). A partir

das predições utilizando o PHASTER não foi possível identificar profagos para essa análise.

Figura 6- Árvore filogenômica de 109 amostras de *S. agalactiae*



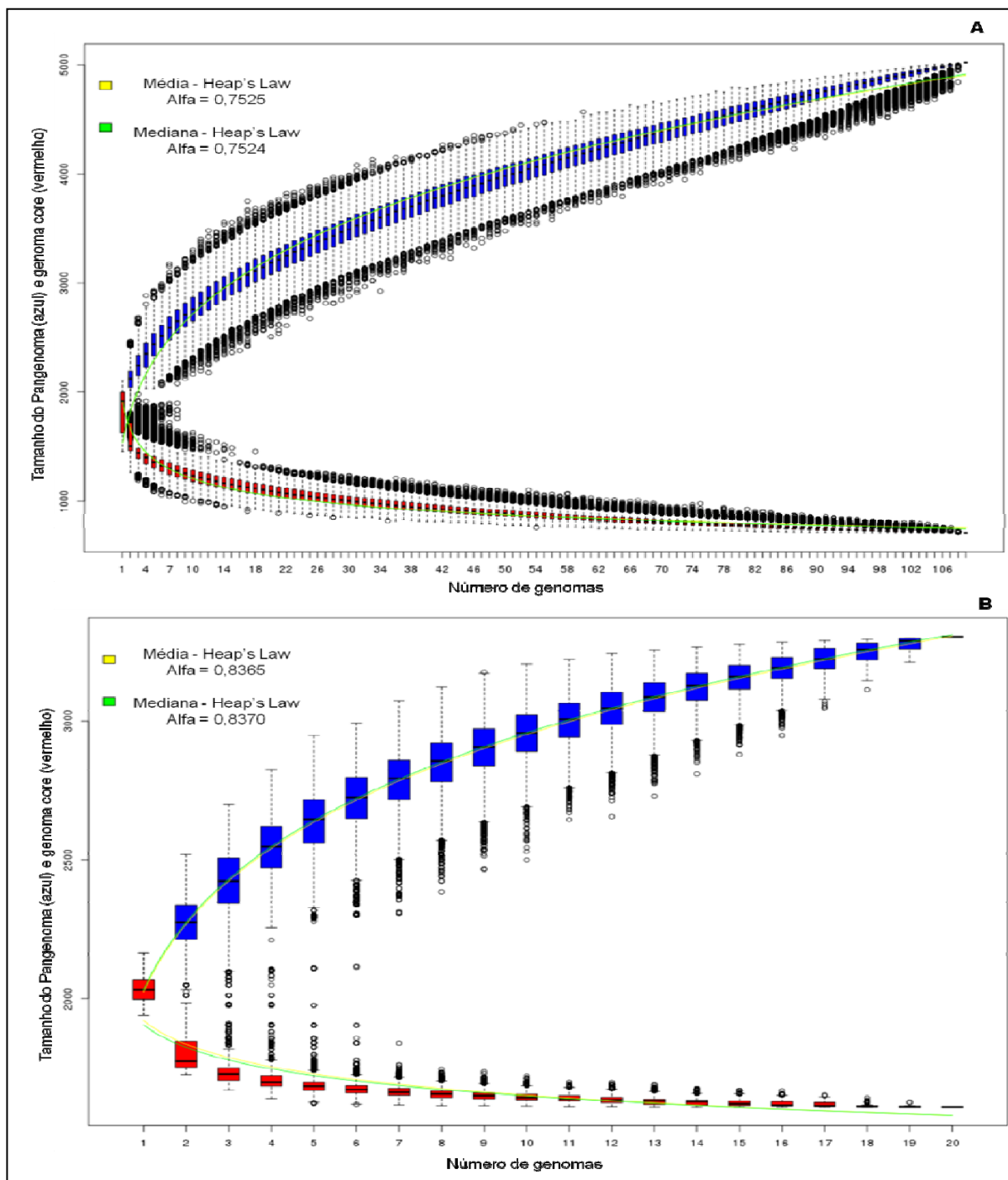
Legenda: Relações filogenômicas entre amostras de *S. agalactiae* de diferentes hospedeiros. À esquerda está a árvore filogenômica de 109 amostras de *S. agalactiae* (APÊNDICE A). À direita estão as informações sobre hospedeiro, tipo capsular e sequência tipo das amostras.

Fonte: A autora, 2019.

#### 4.3.2 Avaliação do pangenoma de *S. agalactiae*

O pangenoma das 109 cepas de *S. agalactiae* apresentou um total de 5.020 genes, dos quais 706 (14%) pertencem ao genoma central e 51 genes são específicos da amostra GBS90356. Com base na lei de Heap, o pangenoma da espécie é aberto ( $\alpha = 0,7525$ ) (Figura 7A).

O pangenoma de 20 amostras de *S. agalactiae* pertencentes ao TC III foi calculado, TC III apresentou um total de 3.304 genes, dos quais 1.610 (48,73%) pertencem ao genoma core, 1.050 (31,78%) ao genoma acessório e 644 (19,49%) genes únicos, sendo 22 genes únicos da amostra GBS90356. O genoma central do TC III (1.610 genes / 48,73% do pangenoma) é maior do que o genoma central calculado para outros TC (706 genes / 14% do pangenoma), sugerindo uma maior similaridade entre os genomas do TC III. Nota-se que o pangenoma de *S. agalactiae* TC III é aberto ( $\alpha = 0,8365$ ), embora apresentem o mesmo TC, esta espécie ainda demonstra uma alta plasticidade, no qual com a inserção de novas amostras, outros genes podem surgir demonstrando a diversidade da espécie (Figura 7B).

Figura 7- Predição do pangenoma de *S. agalactiae*

Legenda: Tamanho do pangenoma (barras gráficas azuis) e tamanho do genoma core (barras gráficas vermelhas). As linhas pontilhadas indicam o desvio padrão e círculos brancos fora das linhas pontilhadas indicam valores discordantes da análise, onde no eixo Y temos o tamanho do pangenoma e genoma core e no eixo X temos o número de amostras de *S. agalactiae* utilizado na análise comparativa realizada no *software* PGAP. (A) Perfil do pangenoma de 109 amostras de *S. agalactiae*; (B) Perfil do pangenoma de 20 amostras de *S. agalactiae* do TC III (APÊNDICE A).

Fonte: A autora, 2019.

#### 4.3.3 Predição e análise das PAIs de *S. agalactiae*

As PAIs contêm um grande número de genes de virulência agrupados em blocos que desempenham papéis importantes na patogenicidade do microrganismo. Em princípio, foram descritas dez ilhas de patogenicidade (APÊNDICE B) com um total de 352 genes preditos para a primeira amostra ST-17 isolada do Brasil, GBS90356, através do software GIPSy e para tal foram consideradas apenas as descritas como normal e forte, descartando assim, as predições fracas. PAIs foram comparadas e analisadas por Blastn e nos dados obtidos foi possível observar que alguns genes pertencentes às ilhas únicas de GBS90356 também estavam presentes no genoma da amostra NGBS128 e COH1, sugerindo que a conservação dessas PAIs é maior do que a observada pelo GIPSy.

A detecção e análise de PAIs no genoma de *S. agalactiae* GBS90356 foi realizada através das ferramentas GIPSy e IslandViewer4 (APÊNDICE B). As PAIs possuem um grande número de genes de virulência agrupados em blocos que desempenham papéis importantes na patogenicidade bacteriana. GIPSy previu 10 PAIs (SagPAI\_1 até SagPAI\_10) que corresponderam a 17% do tamanho do genoma com um total de 352 genes. Por outro lado, o IslandViewer4 detectou apenas 6 PAIs (SagPAI\_2, 4, 5, 8, 9 e 11). Portanto, os resultados previstos pelo GIPSY foram considerados neste trabalho. As PAIs 2 e 5 foram parcialmente detectadas nas amostras do ST-17 (Figura 8) e relacionadas à integrases.

SagPAI\_1 possui 13 genes, incluindo os genes *neu* e *cps* que codificam o ácido siálico e a cápsula polissacarídica, respectivamente. Em contraste, 33 seqüências codificantes em SagPAI\_2 foram identificadas como proteínas não caracterizadas e conservadas em amostras da ST-17 (Figura 8). Apenas 6 genes foram descritos em SagPAI\_3 como os genes *dltS* (regulador da biossíntese do ácido D-alanil-lipoteicóico) e *dltR* (histidina quinase) responsáveis pela caracterização da espécie *S. agalactiae*. Além disso, 4 genes foram identificados como proteínas não caracterizadas.

A análise aprofundada dos genes apresentados na SagPAI\_4 verificou a existência de genes com produtos classificados como proteínas não caracterizadas. Desta forma, torna-se importante realizar estudos futuros para melhor caracterizar as funções biológicas de proteínas não caracterizadas. Na SagPAI\_5, 135 genes foram

identificados, incluindo a proteína de ligação a laminina (gene *lmb*) e o *lacR*, um regulador na ausência de lactose. Os genes relacionados à resistência ao mercúrio foram observados na SagPAI\_6, *merA* (que codifica uma redutase mercúrica) e *merR* (que codifica uma proteína reguladora do operon de resistência ao mercúrio). A álcool desidrogenase induzida por etanol (*adh*) foi localizada na SagPAI\_8, um gene *housekeeping* para caracterização da sequência tipo por MLST. Os genes *vex1-3* e *vncR / S* estavam localizados dentro da SagPAI\_9; embora as funções exatas dos clusters *vex-vnc* continuem a ser elucidadas, demonstrou-se que os genes *vex3* estão associados à resistência à vancomicina, porém a amostra de estudo, GBS90356, foi sensível à vancomicina (dados não mostrados). Por fim, na SagPAI\_10 foi identificado o fator CAMP (sigla para Christie, Atkins e Munch-Peterson), gene codificador, *cfb*.

Genes de importância para *S. agalactiae* também foram encontrados fora das PAIs: *cyE* (codifica a  $\beta$ -hemolisina), *hylB* (hialuronidase), *sodA* (superóxido dismutase), *fbsA* (Proteína de ligação à fibronectina - Fbs), *scpB* (C5a peptidase) e *cas9* (proteína adaptada de um sistema de edição gênica).



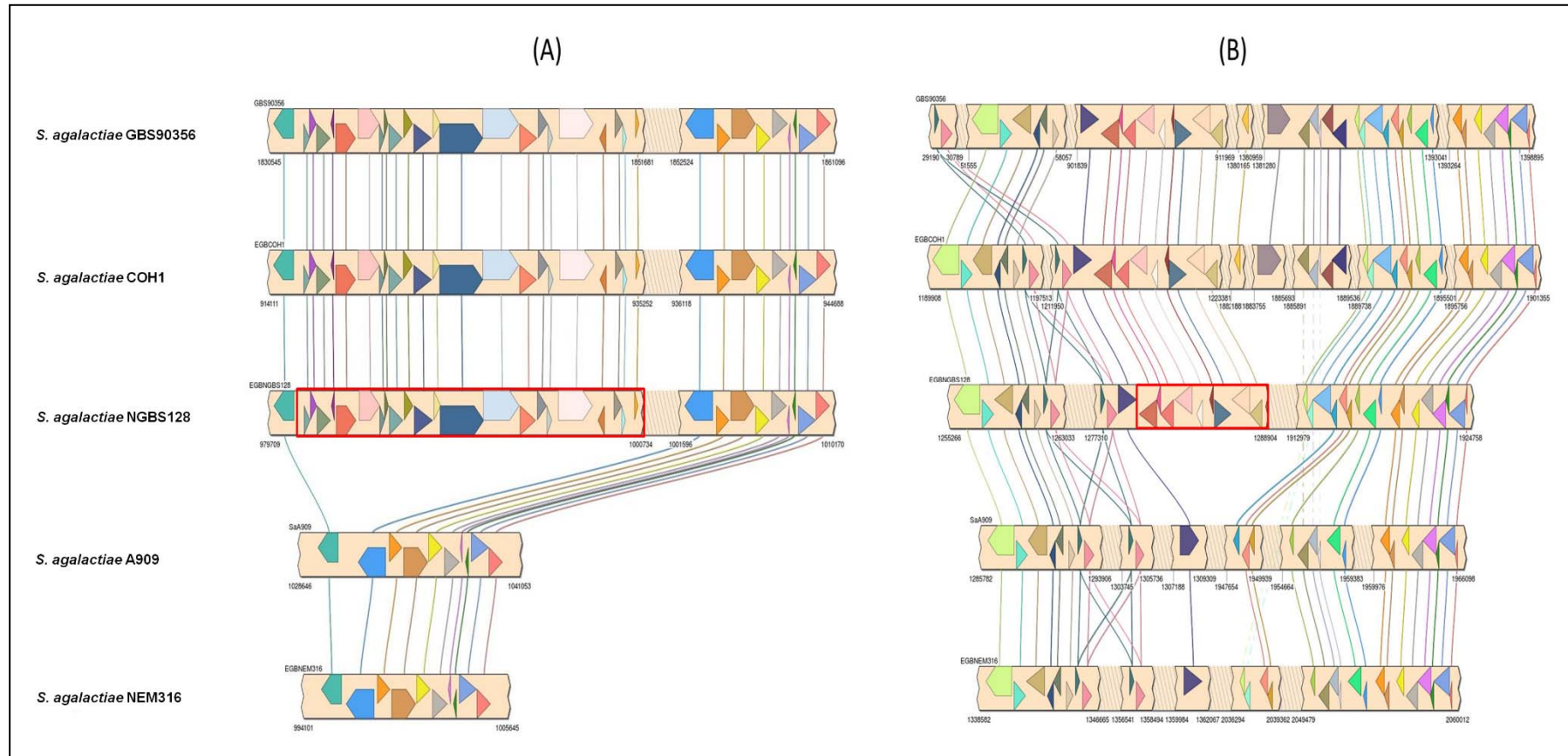


#### 4.3.4 Análise de sintenia das PAIs de *S. agalactiae* e predição do secretoma

Devido as SagPAI\_2 e 5 terem sido parcialmente detectadas nas amostras de *S. agalactiae* pertencentes a ST-17 (Figura 8) e relacionadas à integrases, genes dentro dessas PAIs (Tabela 6) foram comparados por blastn com a sequência de nucleotídeos das amostras NGBS128; COH1; NEM316; A909. A comparação foi realizada através da ferramenta web Simple Synteny (VELTRI *et al.*, 2016). As PAIs apresentaram alta conservação em todas as amostras de *S. agalactiae* analisadas em relação à GBS90356. Uma curadoria manual identificou um gene como Tn5252 presente apenas nas amostras ST-17 (COH1 e NGB128) e não identificado nas amostras NEM316 (TC III) e A909 (TC Ia) (Figura 9). Particularmente, na SagPAI\_2 identificamos um transposon Tn916 e um gene *tetM* que confere resistência à tetraciclina.

O SignalP detectou 20 proteínas com peptídeos sinal para as 352 proteínas descritas nas 10 PAIs identificadas, sugerindo que a maioria das proteínas envolvidas na patogenicidade de *S. agalactiae* não é secretada. Proteínas secretadas (60% - 12/20) foram localizadas na SagPAI\_5.

Figura 9- Análise de sintenia das PAIs 2 (A) e 5 (B) de *S. agalactiae*



Legenda: Os genomas das amostras GBS90356, COH1, NGBS128, A909 e NEM316 são apresentados de cima para baixo na figura. Os genes são representados por setas de cores diferentes. Os genes encontrados apenas nas amostras ST-17 foram destacados por retângulos vermelhos no genoma de *S. agalactiae* NGBS128. Esses genes estão descritos na Tabela 6 e APÊNDICE B.

Fonte: A autora, 2019.

Tabela 6- Genes das PAIs 2 e 5 encontrados em amostras de *S. agalactiae* ST-17

Tamanho do gene (pb)	Nome do gene	Nome do produto
<b><i>S. agalactiae</i> ilha de patogenicidade 2 (30,193 bp)</b>		
1194	-	Integrase, superantigen-encoding pathogenicity islands SaPI
201	-	Hypothetical protein
231	-	Hypothetical protein
423	-	Hypothetical protein
354	-	Hypothetical protein
1920	<i>tetM</i>	Tetracycline resistance protein TetM
276	-	Hypothetical protein
597	-	Hypothetical protein
1023	-	Lipoprotein, NLP/P60 family
2028	-	membrane protein, putative
2454	-	Hypothetical protein
396	-	Hypothetical protein
1026	-	Hypothetical protein
501	-	Antirestriction protein
762	-	Hypothetical protein
222	-	Conjugation related protein
273	-	Hypothetical protein
1185	-	Tn 916, transcriptional regulator, putative
1098	-	Hypothetical protein
159	-	Hypothetical protein
774	-	Hypothetical protein
378	-	Hypothetical protein
<b><i>S. agalactiae</i> ilha de patogenicidade 5 (124,906 bp)</b>		
1203	-	Integrase
258	-	Hypothetical protein
870	-	Hypothetical protein
	-	FtsK-like DNA segregation ATPase, YDCQ
1161	-	B.subtilis ortholog
426	-	Hypothetical protein
162	-	Hypothetical protein
1161	-	Hypothetical protein
1197	-	serine/threonine protein kinase
810	-	type II restriction enzyme <i>eco47ii</i>

Fonte: A autora, 2019.

#### 4.3.5 Análise funcional dos genes das PAIs

Os genes das PAIs de *S. agalactiae* foram classificados em grupos funcionais pelo Gene Orthology (GO). Após comparar a distribuição dos genes funcionais, foram encontrados 102 genes para processos biológicos, 105 genes para componentes celulares e 235 genes relacionados com funções moleculares (Tabela 7) compartilhados pelas amostras de *S. agalactiae*. Processos biológicos do GO foram relacionados principalmente com regulação da transcrição, açúcares dependentes de fosfoenolpirato, sistema de fosfotransferase (PTS), integração de DNA, recombinação de DNA, transporte transmembrana, tradução, transposição, processo metabólico de carboidratos, replicação de DNA, transporte de íon fosfato, transporte de íon metálico e homeostase redox celular.

Para componentes celulares no GO, dos 11 termos relacionados apenas a função de componente integral da membrana foi a mais presente. Além disso, 18 funções moleculares estavam principalmente relacionadas à ligação ao DNA, ligação de ATP, atividade de transferase, de quinase, de hidrolase, de ATPase, ligação de íons metálicos, atividade de metiltransferase, fator de transcrição de ligação a DNA, quinase do sensor de fósforo, peptidases, proteína fosfotransferase da glicose-N (PI)-fosfo-histidina, atividade da helicase, da endonuclease, do transportador transmembrana do fosfato inorgânico, da liase e da isomerase.

A análise GO demonstrou várias anotações funcionais, das quais 99 eram proteínas não caracterizadas. Assim, identificamos 36 proteínas não caracterizadas com os domínios Pfam e Interpro, das quais 16 apresentaram função desconhecida, 3 *helix-turn-helix*, 3 transposons, 3 proteínas de função desconhecida com motivo HXXEE, 1 proteína antirestrição, 1 reconhecimento de nucleosídeo, 1 PTS, 1 Transportador ABC, 1 ATPase, 1 organizador de replisoma de fago N-terminal, 1 família peptidase C39, 1 proteína de ligação efetora associada a integron, 1 lipoproteína exposta à superfície da célula hospedeira, 1 N-terminal fosfolipase\_D-nuclease e 1 nucleotidiltransferase .

Tabela 7- Análise funcional das PAIs de *S. agalactiae* (continua)

<b>Termo</b>	<b>Count<sup>a</sup></b>	<b>Id</b>
<b><i>Processo biológico</i></b>		
Pyruvatee-dependente sugar phosphotransferase system	11	GO:0009401
regulation of transcription, dna-template	11	GO:0006355
DNA integration	8	GO:0015074
DNA recombination	7	GO:0006310
transmenbrane transport	6	GO:0055085
phosphorelay signal transduction system	5	GO:0000160
transposition	3	GO:0032196
translation	3	GO:0006412
cell division	2	GO:0051301
cell redox homeostasis	2	GO:0045454
metal ion transport	2	GO:0030001
phosphate ion transport	2	GO:0006817
DNA replication	2	GO:0006260
carbohydrate metabolic process	2	GO:0005975
<b><i>Componente celular</i></b>		
integral component of membrane	73	GO:0016021
cytoplasm	9	GO:0005737
plasma membrane	6	GO:0005886
membrane	4	GO:0016020
cell wall	3	GO:0005618
ATP-biding cassette (ABC) transporter complex	2	GO:0043190
ribosome	2	GO:0005840
cell	2	GO:0005623
extracellular region	2	GO:0005576
exanuclease repair complex	1	GO:0009380
integral component of plasma membrane	1	GO:0005887
<b><i>Função molecular</i></b>		
DNA binding	32	GO:0003677
ATP binding	24	GO:0005524
sequence specific DNA binding	15	GO:0043565
transferase activity	11	GO:0016740
metal ion transport	7	GO:0046872
ATPase activity	7	GO:0016887
hydrolase activity	7	GO:0016787
kinase activity	7	GO:0016301

Tabela 7- Análise funcional das PAIs de *S. agalactiae* (conclusão)

methyltransferase activity	6	GO:0008168
DNA binding transcription factor activity	5	GO:0003700
protein-N(PI)-phosphohistidine-sugar phosphotransferase activity	4	GO:0008982
peptidase activity	4	GO:0008233
phosphorelay sensor kinase activity	4	GO:0000155
isomerase activity	3	GO:0016853
lyase activity	3	GO:0016829
Inorganic phosphate transmembrane transporter activity	3	GO:0005315
endonuclease activity	3	GO:0004519
helicase activity	3	GO:0004386

Legenda: <sup>a</sup>Count representa o número de genes envolvidos no “GO term”.

Fonte: A autora, 2019.

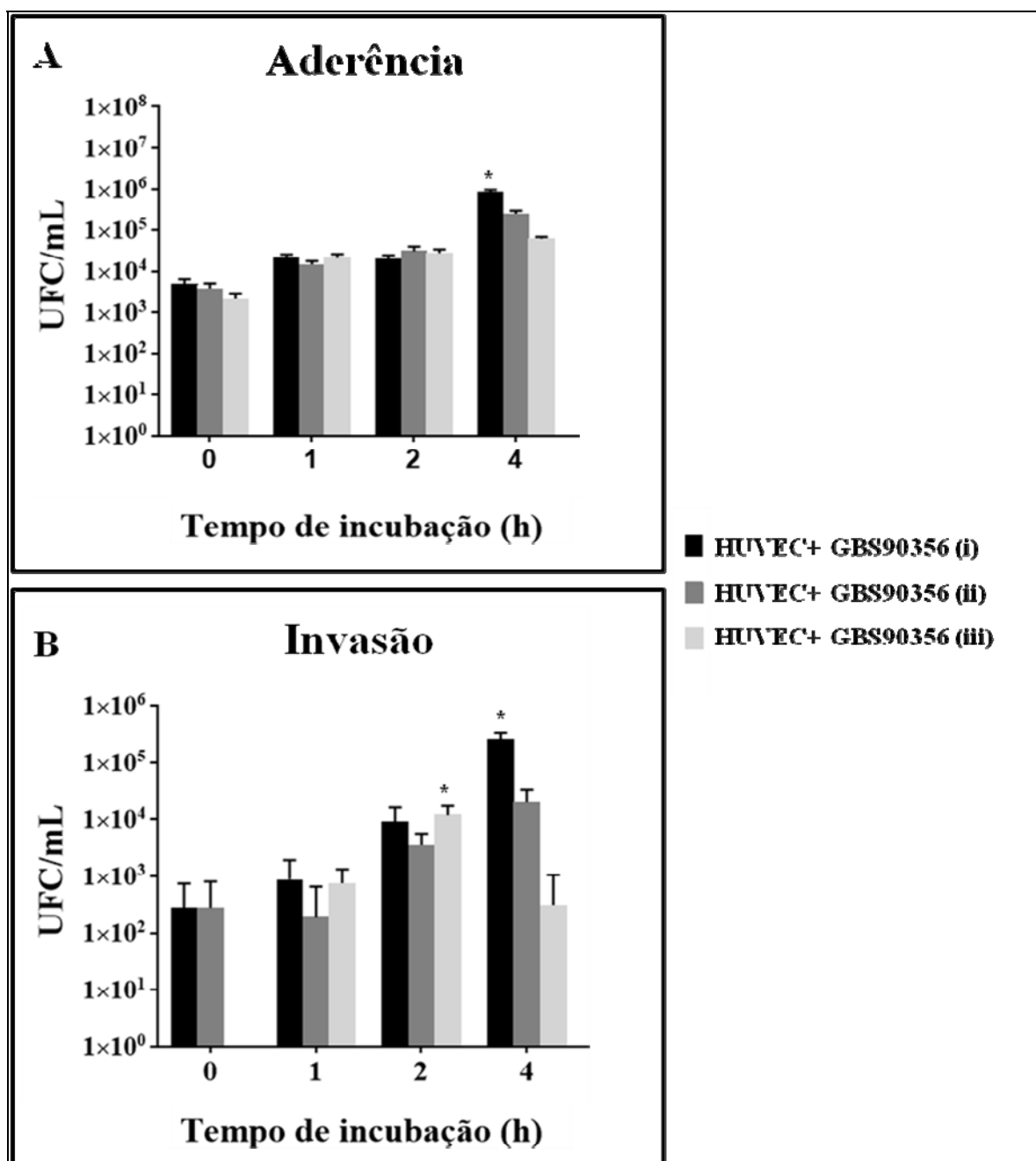
#### 4.4 Papel da hemolisina na interação de *S. agalactiae* GBS90356 com células endoteliais (HUVEC)

Análise do papel biológico da amostra GBS90356 foi realizada através da interação com células endoteliais humanas (HUVEC) na presença ou ausência de um dos principais fatores de virulência, a  $\beta$ -hemolisina. Os ensaios de aderência e invasão em HUVEC foram apenas realizados com a amostra GBS90356 (TC III / ST-17), cujos resultados da cinética de aderência e invasão estão representados no Gráfico 1. A amostra bacteriana demonstrou um perfil de aderência às células HUVEC similar para todas as condições nos tempos de 0h, 1h e 2h de incubação (Gráfico 1A). Após 4h de incubação, GBS90356 na condição (i) [não suplementado com glicose] apresentou diferença estatística quando comparada ao controle ( $8,3 \times 10^5$  UFC/mL;  $p < 0,05$ ).

Com relação à invasão, os dados demonstraram que a amostra mantida com suplemento de glicose [10 mg/mL], somente começou a invadir as HUVEC a partir de 1h de incubação quando comparadas ao controle (Gráfico 1B), com maior capacidade invasiva após 2h de interação ( $p < 0,05$ ). Porém, no tempo de 4h a invasão foi reduzida se comparada ao tempo de 2h (de  $1 \times 10^4$  para  $1 \times 10^2$  UFC/mL).

Quando a amostra GBS90356 foi crescida em meio não suplementado com glicose, ela foi capaz de invadir desde 0h de incubação e houve aumento significativo na invasão após 4 h de incubação ( $2,6 \times 10^5$ ;  $p < 0,05$ ).

Gráfico 1 - Interação da amostra GBS90356 com HUVEC



Legenda: Cinética de aderência (A) e invasão (B) da amostra GBS90356 com HUVEC. (i) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI; (ii) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI + glicose [10mg/mL]; (iii) HUVEC em meio 199 + glicose [10mg/mL] infectada com GBS90356 crescidas em BHI + glicose [10mg/mL]. Gráfico representativo da média  $\pm$  desvio padrão de três experimentos independentes. \*Diferenças significativas em relação ao controle  $P \leq 0,05$  através do teste  $t$  de Student.

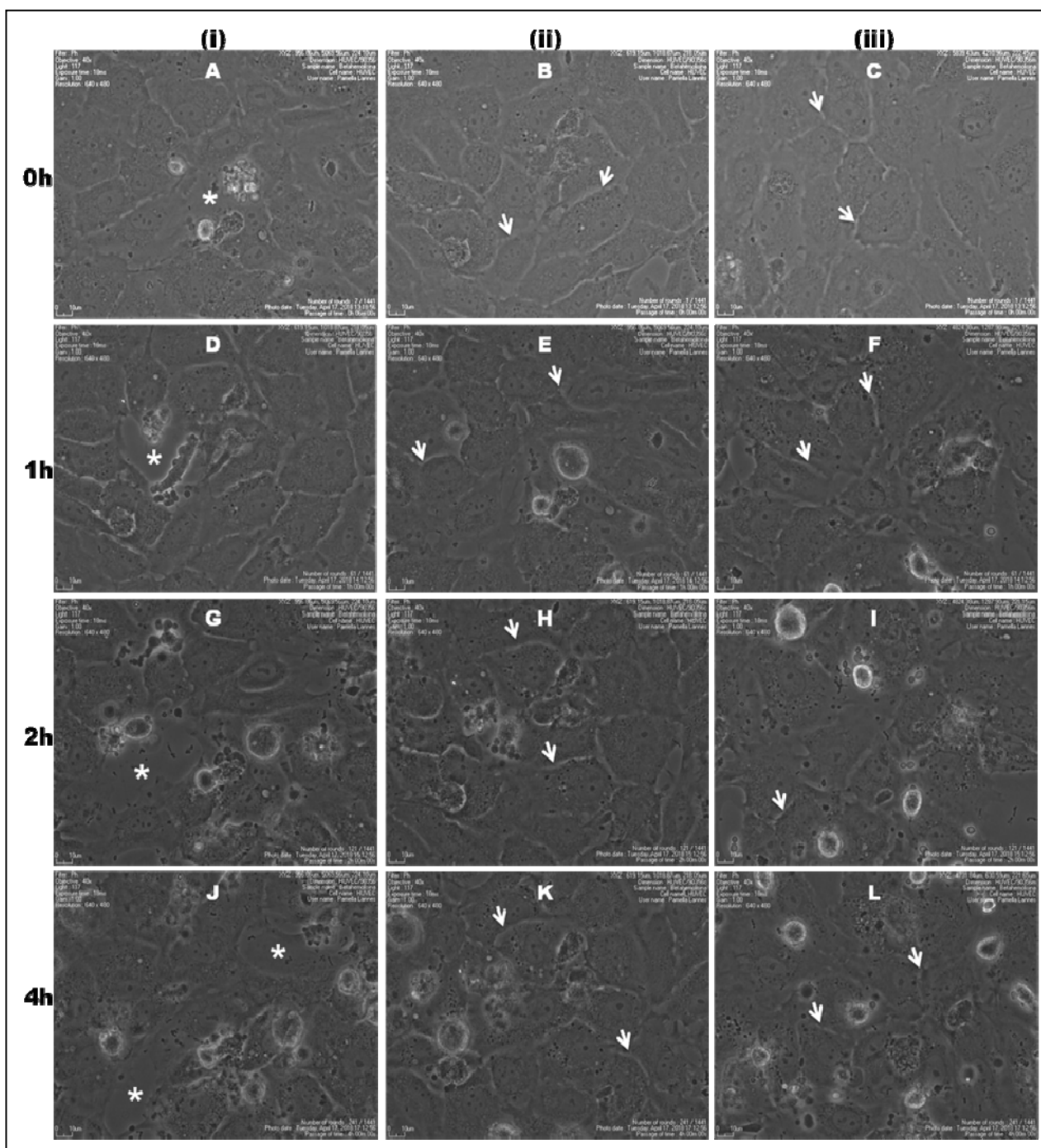
Fonte: A autora, 2019.



#### 4.5 Análise morfológica do perfil de interação celular (HUVEC X *S. agalactiae* GBS90356) por *time-lapse* na *Biostation*<sup>®</sup>

Micrografias da interação celular de *S. agalactiae* GBS90356 com HUVEC estão representadas na Figura 10 demonstrando as seguintes condições: (i) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI (A, D, G, J, M); (ii) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI + glicose [10mg/mL] (B, E, H, K, N); (iii) HUVEC em meio 199 + glicose [10mg/mL] infectada com GBS90356 crescidas em BHI + glicose [10mg/mL] (C, F, I, L, O). As análises das imagens foram realizadas através do programa ImageJ acoplado a *Biostation*<sup>®</sup>. Os resultados mostraram morfologia característica de células em apoptose (dados não mostrados) com contração celular e perda de contato célula-célula e célula-substrato (Figura 10D, G, J), desfazendo as junções intercelulares do endotélio, principalmente após 2 h de incubação na condição (i) (Figura 10G, J).

Figura 10 - Análise morfológica da interação da amostra GBS90356 com HUVEC



Legenda: Micrografias da interação celular de *S. agalactiae* com HUVEC nos tempos de 0h (A, B, C), 1h (D, E, F), 2h (G, H, I) e 4h (J, K, L) nas seguintes condições: (i) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI (A, D, G, J); (ii) HUVEC em meio 199 puro infectada com GBS90356 crescidas em BHI + glicose [10mg/mL] (B, E, H, K); (iii) HUVEC em meio 199 + glicose [10mg/mL] infectada com GBS90356 crescidas em BHI + glicose [10mg/mL] (C, F, I, L). \* Junções intercelulares desfeitas. Setas indicam as junções intercelulares. As análises das imagens foram realizadas através do programa ImageJ acoplado a Biostation®.

Fonte: A autora, 2019.

## 5 DISCUSSÃO

*S. agalactiae* tem sido considerado um importante patógeno responsável por infecções invasivas em humanos e animais (PÉRICHON *et al.*, 2017). As principais manifestações clínicas humanas incluem infecções como pneumonia, sepse e meningite que ocorrem imediatamente após o nascimento, relatadas como doença de início precoce (EOD; 0-6 dias) ou após a primeira semana de vida, descrita como doença de início tardio (LOD; 7-90 dias) (LIU *et al.*, 2018; SHABAYEK; SPELLERBERG, 2018). A ST-17 tem sido referida como um clone hipervirulento, distribuído mundialmente e associado ao aumento de infecções invasivas neonatais a partir da década de 60 (BELLAIS *et al.*, 2012; JONES *et al.*, 2006; LAMY *et al.*, 2006; SHABAYEK; SPELLERBERG, 2018). Desta forma, este trabalho relata o primeiro sequenciamento do genoma de uma amostra de *S. agalactiae* brasileira ST-17 (GBS90356) isolada na região metropolitana do Rio de Janeiro. A amostra GBS90356 apresentou tamanho de genoma e conteúdo G + C semelhante às amostras de *S. agalactiae* isoladas de diferentes hospedeiros (AGUIAR *et al.*, 2016; HE *et al.*, 2017).

A cápsula polissacarídica permite a classificação de amostras de *S. agalactiae* em diferentes tipos capsulares (Ia, Ib – IX), além de ser um dos principais fatores de virulência do microrganismo (NAKIB *et al.*, 2011). No entanto, alguns isolados são considerados não tipáveis (NT) porque não reagem com nenhum soro anticapsular. Amostras não tipáveis são responsáveis por 5-20% dos isolados que colonizam ou infectam humanos e por 30-77% daqueles obtidos a partir de mastite bovina (MARTINEZ *et al.*, 2000). Desta forma, métodos de tipagem molecular têm sido desenvolvidos, baseados na amplificação por PCR dos diferentes alelos da cápsula (operon *cps*) de *S. agalactiae* permitindo diminuir o número de amostras NT (ROSINI *et al.*, 2015).

O desenvolvimento de técnicas de tipagem baseadas no sistema convencional de PCR é simples e efetivo para identificar os diferentes TC em grande número de isolados (PAI *et al.*, 2006). De acordo com alguns autores (IMPERI *et al.*, 2010; KORIR *et al.*, 2016; POYART *et al.*, 2007; SLOTVED *et al.*, 2007) a PCR multiplex se caracteriza como uma das técnicas moleculares capazes de produzir resultado confiável, assim como o *Multi Locus Sequence Typing* (MLST). A tipagem

por PCR multiplex é uma ferramenta valiosa que parece ser suficientemente discriminatória para estudos epidemiológicos e que fornece de modo preciso e inequívoco a caracterização das amostras de *S. agalactiae*, permitindo promover uma investigação mais aprofundada sobre a biologia populacional deste microrganismo, além de possuir elevada reprodutibilidade (BRAGA *et al.*, 2013; DIEDRICK *et al.*, 2010). Sendo assim, os testes baseados em análise de DNA são menos propensos a erros que os métodos tradicionais de identificação bacteriana, além de serem mais rápidos e com maior potencial discriminatório (ALMEIDA *et al.*, 2013; BRAGA *et al.*, 2013).

Como o *S. agalactiae* é uma bactéria que apresenta características diversas variando de acordo com a região geográfica e hospedeiro, ferramentas moleculares epidemiológicas têm sido desenvolvidas com o objetivo de estudar a associação entre TC e doença, bem como a avaliação da variação genética dentro dos TC (NAKIB *et al.*, 2011). Nos últimos anos, a tipagem molecular de *S. agalactiae* com base na detecção do gene tipo-específico da região capsular tem sido objeto de estudo (KORIR *et al.*, 2014).

Diante de todas as informações descritas, o primeiro passo do nosso estudo foi implantar uma técnica molecular que conseguisse diferenciar os TC de amostras de *S. agalactiae* de hospedeiros distintos. A implantação da PCR multiplex para tipagem de amostras de *S. agalactiae* no Laboratório de Biologia Molecular e Fisiologia de Estreptococos (LBMFE) permitiu qualificá-lo como um dos poucos laboratórios no Brasil capacitado para identificação molecular, contribuindo para o melhor conhecimento sobre a epidemiologia do *S. agalactiae*. Além disso, a PCR multiplex implantada no LBMFE e em fase de patente, permitirá analisar também alguns fatores de virulência e genes de resistência aos antimicrobianos, agregando importantes colaboradores como a EMBRAPA, grupo coordenado pelo Dr. Hélio Langoni da UNESP e o grupo coordenado pelo Dr. Rinaldo Mota da UFRPE. A distribuição do TC varia geograficamente e o Brasil sendo um país com grande diversidade étnica, climática, social e econômica, necessita de estudos epidemiológicos que contribuam para análises da prevalência dessa bactéria nas diferentes regiões do país.

Nossos resultados demonstraram que dentre as 43 amostras de *S. agalactiae* de origem humana não oncológica avaliadas nesse estudo, 41,86% pertenceram ao TC III, seguido pelos TC V (20,93%), Ia e II (9,3% cada), Ib (4,65%) e VIII (2,33%),

estando de acordo com os dados encontrados no Canadá, entre 2003 e 2013, onde TC III (20%) foi o mais frequente, seguido pelos TC V (19%), Ia (19%), Ib (13%) e II (11 %) (RAABE; SHANE, 2018). Nossos resultados também corroboram com os resultados descritos na França e África do Sul, no qual o TC III foi o mais prevalente com 25,7% e 37,3%, respectivamente (MADZIVHANDILA, 2011; TAZI *et al.*, 2010). Na Dinamarca o TC III foi o mais isolado em adultos, representando 29,4%, seguido pelos TC Ia com 19,7% e TC V com 18,7% (LAMBERTSEN *et al.*, 2010). Pode-se destacar também o segundo TC mais prevalente no nosso estudo, o TC V, tem sido frequentemente isolado em infecções invasivas em adultos não grávidos em países como Espanha (BOLÑANOS *et al.*, 2005), Suíça (PERSSON *et al.*, 2004), Noruega (BERGSENG *et al.*, 2008), Estados Unidos (SKOFF *et al.*, 2009), Austrália e Nova Zelândia (ZHAO *et al.*, 2008), demonstrando a diversidade das amostras humanas. Cabe ressaltar que todas as publicações anteriormente descritas realizaram a tipagem por sorologia. FLORINDO (2014) evidenciou através da PCR convencional, entre os anos de 2005 a 2012 em Portugal, que os TC mais frequentes durante este período foram o III (23,3%), Ia (21,3%) e V (20,1%). Estudos realizados na América Latina reportaram que os TCs III (34,6%), V (28,7%), Ia (24,3%) e II (16,7%), foram os mais evidenciados em isolados de adultos não grávidos (COLLINS *et al.*, 1998; PALMEIRO *et al.*, 2010; PÉREZ *et al.*, 2004).

Até o momento, não existem dados na literatura que tenham estudado somente amostras de *S. agalactiae* isoladas de pacientes portadores de neoplasias. Neste contexto, as análises dos resultados obtidos neste trabalho foram realizadas, na sua maioria, com dados publicados na literatura de indivíduos não portadores de câncer. Na Colômbia, apesar dos autores terem descrito casos de infecções invasivas por *S. agalactiae* em pacientes com câncer e/ou malignidades hematológicas, esse pequeno percentual de amostras fazia parte de um grupo maior de indivíduos com diabetes, HIV, desordens hepáticas e pulmonares (CRESPOR-TIZ *et al.*, 2014). Em 2016, nosso grupo publicou o primeiro trabalho retrospectivo enfatizando infecções por *S. agalactiae* em pacientes com câncer no Brasil (PIMENTEL *et al.*, 2016).

Quando analisamos amostras isoladas exclusivamente de pacientes oncológicos nossos resultados demonstram a prevalência do TC Ia (44%), seguido do TC V (24%). O TC II apareceu como o terceiro mais encontrado entre as amostras de pacientes oncológicos com 15%, seguido do TC III com 11%. O TC IV

foi identificado em apenas 2% das nossas amostras de pacientes com câncer, assim como o TC VI (2%) e TC VII (2%). Não foram identificadas amostras oncológicas dos TC Ib e VIII que são frequentemente isoladas em neonatos na região Sul do Brasil (OTAGUIRI *et al.*, 2013) e Japão (KARUNAKARAN *et al.*, 2009), respectivamente. Desta forma, a capacidade de *S. agalactiae* em causar infecções em pacientes com diferentes comorbidades, incluindo o câncer reforçam a necessidade de estudos epidemiológicos mais aprofundados. Este é o primeiro trabalho, até o momento, que estudou o perfil capsular exclusivamente com amostras isoladas de pacientes com câncer.

No Brasil, informações sobre as características de *S. agalactiae* obtidas de origem bovina ainda são limitadas. Estudos prévios utilizando técnicas imunológicas convencionais em amostras de *S. agalactiae* de leite de animal com mastite bovina na região Sudeste do Brasil mostraram o TC III como predominante (77,6%), seguido dos TC II, Ia, Ib e VI (DUARTE *et al.*, 2004). Além disso, o TC V foi prevalente nas amostras de *S. agalactiae* isoladas de mastite nas regiões Sudeste e Sul, entre 1987 e 1989 e entre 2003 e 2006. Os TC Ia (23.5%), II (15%) e III (18.9%) bem como amostras NT (19.7%) também foram observados entre os isolados de *S. agalactiae* bovinos tipados por sorologia (PINTO *et al.*, 2013). No presente estudo, mostramos que das 87 amostras analisadas de *S. agalactiae* oriundas de mastite bovina, os TC prevalentes foram: II (48,28%), III (29,89%), V (10,34%) e VII (1,15%), demonstrando um novo perfil de incidência do TC II nas regiões Sudeste e Nordeste do Brasil, diferentemente do encontrado nos levantamentos realizados anteriormente nas regiões Sudeste e Sul do Brasil (DUARTE *et al.*, 2004; PINTO *et al.*, 2013).

Os isolados de *S. agalactiae* não tipáveis foram detectados em 7,5% das amostras. Neste trabalho, o tipo IX não foi investigado, pois se encontra em fase de testes, podendo explicar o número (n = 14) de amostras de *S. agalactiae* NT até o momento.

*S. agalactiae* compartilha algumas características com outros Estreptococos patogênicos. No entanto, o repertório preciso de atributos compartilhados e únicos que expliquem o surgimento de *S. agalactiae* como um importante patógeno para populações humanas permanece indefinido (TETTELIN *et al.*, 2002; SHABAYEK; SPELLERBERG, 2018). Para elucidar a base molecular da virulência de *S. agalactiae*, sequenciamos o genoma da primeira amostra GBS90356 isolada no Brasil pertencente ao TC III (ST-17) isolada de um neonato com meningite e

realizamos análises comparativas com outras amostras patogênicas de *S. agalactiae* recuperadas nas bases de dados públicas.

A sequência de um único genoma não reflete como a variabilidade genética impulsiona a patogênese dentro de uma espécie bacteriana (Tettelin *et al.*, 2005). Estudo prévio demonstrou através da análise de sequências do genoma de 24 amostras de *S. agalactiae*, sendo somente uma amostra ST-17 (COH1) que o pangenoma bacteriano é vasto, onde novos genes poderão ser adicionados ao *pool* genético da espécie (He *et al.*, 2017). Portanto, para fazer estimativas mais precisas do tamanho do pangenoma de *S. agalactiae*, incluímos 109 genomas de diversos hospedeiros disponíveis no GenBank, incluindo as amostras ST-17 (n = 6) para confirmar a alta plasticidade desta espécie.

Como o TC III tem sido reportado com elevada frequência em infecções invasivas em humanos, no qual a linhagem ST-17 hipervirulenta está envolvida na disseminação mundial e associada à elevada prevalência da síndrome de início tardio (SHABAYEK; SPELLERBERG, 2018), foram selecionadas 20 amostras humanas de *S. agalactiae* pertencentes ao TC III para análise do pangenoma. Os dados demonstraram que o pangenoma é aberto ( $\alpha = 0,83$ ), mostrando um conjunto complexo de material genético e um alto nível de adaptabilidade por transferência gênica horizontal (DONATI *et al.*, 2010). Em comparação com outras espécies de bactérias patogênicas, como *Pseudomonas syringae* com 64% do genoma central, *Streptococcus pneumoniae* (74%) e *Listeria monocytogenes* (80%), o *S. agalactiae* contém um dos menores genomas core (14%) descritos na literatura, sugerindo uma alta variabilidade genética dessa espécie (BALTRUS *et al.*, 2011; BARAÚNA *et al.*, 2017; HE *et al.*, 2017; PUYMÈGE *et al.*, 2015).

Embora a maior plasticidade possa estar associada aos múltiplos hospedeiros de *S. agalactiae* (HE *et al.*, 2017; LEFÉBURE; STANHOPE, 2007; TETTELIN *et al.*, 2005), nosso estudo usando apenas um único hospedeiro demonstrou que linhagem humana de ST-17 apresentou pangenoma aberto e uma alta plasticidade. Assim, podemos sugerir que outros fatores como elementos móveis genéticos e a evolução genética também possam influenciar na plasticidade genômica do microrganismo (HE *et al.*, 2017; TEATERO *et al.*, 2016). Além disso, a aquisição de elementos genéticos móveis parece correlacionar-se com a expansão clonal das amostras bacterianas (TEATERO *et al.*, 2016). Elementos genéticos móveis são características comuns de *S. agalactiae*, sendo que cada amostra possui seu próprio

repertório de transposons, fagos, integrases e plasmídeos integrados (DONATI *et al.*, 2010; HERBERT *et al.*, 2005) com potencial papel de virulência. Sendo assim, este estudo traz contribuições para a melhor compreensão da patogênese, especialmente da linhagem ST-17.

As PAIs são elementos genéticos distintos nos cromossomos de um grande número de patógenos bacterianos, codificando diversos fatores de virulência normalmente ausentes em amostras não patogênicas da mesma espécie ou espécies próximas evolutivamente (GAL-MOR; FINLAY, 2006). Diversas ferramentas têm sido desenvolvidas para detecção *in silico* de potenciais PAIs, incluindo programas que integram múltiplos recursos das PAIs, como composição atípica de sequências e genes associados à transferência gênica horizontal (GAL-MOR; FINLAY, 2006; HSIAO *et al.*, 2005). Segundo Silva-Filho (2018), o GIPSy apresentou os melhores resultados globais quando comparado ao IslandViewer4, alcançando 91% de cobertura. Neste estudo, o GIPSy predisse 10 PAIs (SagPAI\_1 a SagPAI\_10) e o IslandViewer4 identificou apenas 6 PAIs. Assim, os resultados da ferramenta GIPSy foram considerados no presente trabalho.

As PAIs de *S. agalactiae* GBS90356 contêm genes funcionais importantes, bem como genes desconhecidos. Curiosamente, as ilhas de patogenicidade das amostras ST-17 de *S. agalactiae* foram altamente conservadas. Em SagPAI\_1, *S. agalactiae* mostrou apenas os genes *cps* e *neu* que permitem que as bactérias evadam as defesas imunes do hospedeiro e conferem a capacidade de desenvolvimento de meningite (PAILHORIE *et al.*, 2013). Nossos resultados demonstraram que a SagPAI\_2 apresentou o elemento conjugativo integrativo (ECI) Tn916 carregando o gene *tetM* que confere resistência à tetraciclina (CUNHA *et al.*, 2014), corroborando a resistência à tetraciclina observada para GBS90356 pelo teste em disco, de acordo com CLSI, 2016 (dados não mostrados). Além disso, foram encontrados genes de resistência ao mercúrio (operon *mer*) (SOTERO-MARTINS *et al.*, 2008). Os genes de resistência são elementos capazes de transposição conjugativa para várias espécies bacterianas, contribuindo para o surgimento de resistência múltiplas drogas em patógenos Gram-positivos (CAMPISI *et al.*, 2016).

Os isolados de *S. agalactiae* possuem diferentes tipos de *lac* operon (RICHARDS *et al.*, 2013). No entanto, encontramos apenas o gene *lacR* localizado em SagPAI\_5, um regulador que se liga ao promotor na ausência de lactose



(RICHARDS *et al.*, 2013). Outro gene identificado na SagPAI\_5 foi o *lmb*, uma adesina ligada a laminina (Lmb) que apresenta um papel no tropismo bacteriano do sistema nervoso central, permitindo que *S. agalactiae* atravesse a barreira hematoencefálica, ocasionando a meningite (VAN WIJK *et al.*, 2017; SHABAYEK; SPELLERBERG, 2018). SagPAI\_2 e 5 apresentaram o gene transportador ABC, essencial para viabilidade celular, virulência, nutrição, aquisição e exportação de substâncias, incluindo polissacarídeos capsulares e hemolisina (DAVIDSON *et al.*, 2008; HE *et al.*, 2017). Além disso, ambas as PAIs apresentaram duas regiões de integrase relevantes, mostrando que os elementos genéticos móveis, assim como os genes que carregam podem contribuir para a hipervirulência e plasticidade da ST-17.

Os transposons conjugativos são elementos genéticos móveis capazes de transferência intercelular, permitindo sua manutenção, disseminação e regulação de ambos os processos (BROCHET *et al.*, 2009). Os mecanismos de disseminação e transferência de elementos móveis dentro de *S. agalactiae* permanecem pouco conhecidos. Na SagPAI\_8, o gene da integrase de um transposon conjugativo de estreptococos, Tn5252, foi identificado. O Tn5252 foi originalmente identificado como um elemento de origem cromossômica em *Streptococcus pneumoniae* (CHURCHWARD, 1995). A caracterização de elementos relacionados irá revelar a diversidade de combinações de módulos dentro do ECI, permitindo o estudo da interação entre diferentes módulos. Além disso, a hibridização de DNA de um conjunto de diversas amostras de origem humana e animal revelou que os transposons estão frequentemente presentes na população de *S. agalactiae*, compensando a ausência de plasmídeos (BROCHET *et al.*, 2008; MUNOZ-NAJAR, 1999).

O presente trabalho também mostrou genes de resistência à vancomicina (*vex3-vncR / S*) em SagPAI\_9, no entanto, como são escassos os dados reportados sobre a resistência de *S. agalactiae* à vancomicina, a base molecular é essencialmente desconhecida (GIZACHEW *et al.*, 2019; LIN *et al.*, 2011). A amostra GBS90356 mostrou-se susceptível à vancomicina (dados não mostrados). Além disso, a SagPAI\_10 apresentou o gene *cfb* codificando uma proteína extracelular formadora de poros, conhecida como fator CAMP, amplamente utilizada para a identificação fenotípica de isolados de *S. agalactiae* (CHRISTIE *et al.*, 1944; LANG; PALMER, 2003). Esta toxina formadora de poros é responsável pela disseminação sistêmica de *S. agalactiae* no hospedeiro (RAJAGOPAL, 2009). Apesar do exposto

acima, os mecanismos moleculares e a importância do fator CAMP na patogênese bacteriana ainda não são totalmente elucidados.

No presente estudo, foram preditas várias proteínas hipotéticas e genes desconhecidos que podem ser alvos no futuro em estudos para o estabelecimento de doenças por *S. agalactiae*. Portanto, os resultados indicam que uma grande parte do genoma codifica proteínas envolvidas na interação patógeno-hospedeiro. Nossos dados também relacionaram a presença de transposons com proteínas não caracterizadas, devido ao transposon ser integrado ao cromossomo. Tipicamente, um transposon conjugativo codifica apenas as proteínas necessárias para essa atividade e as proteínas que a regulam (BROCHET *et al.*, 2009).

Os peptídeos sinais foram analisados apenas nos genes dentro das PAIs e demonstramos que 20 genes têm peptídeos sinal, sugerindo que a maioria das proteínas envolvidas na patogenicidade de *S. agalactiae* dentro das PAIs analisadas não são secretadas. No entanto, evidências indicam que a linhagem ST-17 exibe uma combinação específica de proteínas secretadas e expostas à superfície como a SrtA e HvgA (BROCHET *et al.*, 2006; TAZI *et al.*, 2010).

Também analisamos outros genes de importância para *S. agalactiae* que foram preditos fora das PAIs como *cyfE* ( $\beta$ -hemolisina) que tem um papel central no equilíbrio das respostas pró e antiinflamatórias do hospedeiro infectado, necessárias para romper as barreiras epiteliais e endoteliais e a membrana do fagolisossomo (SIX *et al.*, 2015). Outros genes encontrados foram: (i) *hylB*, gene que codifica a hialuronidase, uma proteína exposta à superfície que cliva o ácido hialurônico e facilita a invasão e disseminação de bactérias através dos tecidos do hospedeiro (MELO *et al.*, 2016; OVIEDO *et al.*, 2013); (ii) *sodA*, superóxido dismutase, enzima essencial para a proteção de *S. agalactiae* contra o estresse oxidativo e para a sobrevivência em diferentes hospedeiros (POYART *et al.*, 2007); (iii) *fbsA* (Proteína de ligação à fibronectina - Fbs) relacionada com interação entre adesinas de superfície bacteriana e componentes das proteínas do plasma do hospedeiro e da matriz extracelular (MEC) (RAGUNATHAN; PONNURAJ, 2011). Estudos realizados com mutante *fbsA* demonstraram que a região repetida de FbsA contém o local de ligação para o ligante, sugerindo que *fbsA* seja um importante receptor de fibrinogênio do *S. agalactiae* (PIETROCOLA *et al.*, 2018; SHABAYEK; SPELLERBERG, 2018); (iv) *scpB*, peptidase C5a, uma serina protease localizada na superfície que cliva o C5a humano, reduzindo a quimiotaxia de neutrófilos e

diminuindo a morte opsonofagocítica. Adicionalmente, media a ligação à fibronectina e contribui para a invasão em células epiteliais (BECKMANN *et al.*, 2002; CHENG *et al.*, 2002) e (v) proteína cas9 adaptada de um sistema de edição de genoma natural em bactérias. As bactérias capturam fragmentos de DNA de vírus invasores e os usam para criar segmentos de DNA conhecidos como matrizes CRISPR (LIER *et al.*, 2015). Estes genes preditos fora das PAIs revelam que o leque de fatores de virulência de *S. agalactiae* é vasto e complexo.

Proteínas não caracterizadas identificadas nos genomas de *S. agalactiae* publicados até a presente data, permanecem pouco conhecidas. Desta forma, neste trabalho incluímos análises de proteínas não caracterizadas, avaliando os domínios Pfam e Interpro que incluem: o sistema de fosfotransferase de açúcar dependente de fosfoenolpiruvato (PTS), um importante sistema de transporte ativo de carboidratos em *Streptococcus* que, através da translocação na membrana celular, catalisa a fosforilação de substratos de açúcar (DEUTSCHER *et al.*, 2006); o domínio estrutural *helix-turn-helix* capaz de se ligar ao DNA, auxiliando na regulação da expressão gênica de *S. agalactiae* (ARAVIND *et al.*, 2005); o transportador de ligação a ATP (ABC) responsável pelo transporte de bacteriocinas peptídicas através da membrana citoplasmática. O transportador ABC é a protease de maturação e seu domínio proteolítico reside na parte N-terminal da proteína (HAVARSTEIN; DIEP; NES, 1995), demonstrando sua importância para as bactérias patogênicas; além da atividade transferase; atividade hidrolase; componente integral da membrana; processo metabólico de carboidratos; e da ligação ao DNA sequência específica.

Todos esses achados sugerem que a virulência de *S. agalactiae* está diretamente relacionada à sua alta plasticidade e à presença de elementos genéticos móveis, conseqüentemente, o sequenciamento de novos genomas ajudará a determinar um conjunto mais robusto de fatores de virulência. A compreensão do conceito de PAIs pode afetar profundamente a maneira como percebemos a virulência bacteriana e a evolução microbiana. Em conjunto, estas observações mostraram a importância da vigilância genômica de amostras epidemiologicamente relevantes e mostram o poder da análise *in silico* do genoma de *S. agalactiae* (ST-17) para entender a epidemiologia e plasticidade do genoma da população de *S. agalactiae*.

Nossos resultados mostraram ter implicações para a patogênese, evolução e o conceito de espécie e sugerem que as estratégias de pesquisa para genomas microbianos podem precisar de comprovações através de métodos biológicos. Deste modo, incluímos neste estudo como modelo biológico a interação da amostra GBS90356 com endotélio, visando avaliar o papel da  $\beta$ -hemolisina, um importante fator de virulência bem caracterizado em *S. agalactiae* localizado fora das PAIs identificadas pelo GIPSy e IslandViewer4. Os mecanismos que desencadeiam o processo de infecção e morte celular induzido pelo *S. agalactiae* permanecem ainda pouco elucidados, principalmente quando falamos de amostras ST-17. A  $\beta$ -hemolisina é um fator de virulência que desempenha papel chave na toxicidade de *S. agalactiae*. Em fibroblastos e trofoblastos a  $\beta$ -hemolisina auxiliou na morte dessas células quando em contato com o *S. agalactiae* (KAPLAN *et al.*, 2008).

Nossos resultados com células endoteliais (HUVEC) demonstraram que a inatividade da  $\beta$ -hemolisina utilizando glicose [10mg/mL], diminuiu o perfil de invasão do patógeno em HUVEC a partir de 2h de incubação quando comparado ao controle. Embora sejam escassos na literatura dados que relacionem o papel da hemolisina de *S. agalactiae* com HUVEC, nossos dados são equivalentes aos relatos de Fettucciari *et al.* (2011), onde a inibição da atividade  $\beta$ -hemolítica do *S. agalactiae* foi correlacionada com a diminuição da aderência do patógeno em macrófagos. Este resultado foi consistente também com o trabalho descrito anteriormente, onde o perfil de invasão em macrófagos reduziu significativamente após a inatividade da hemolisina (MARCHLEWICZ; DUNCAN, 1980; 1981). Ullet *et al.* (2004) demonstraram que a  $\beta$ -hemolisina de *S. agalactiae* do TCIII desempenhou papel na fagocitose, sobrevivência intracelular e indução da morte celular por apoptose em macrófagos. Neste mesmo ano, Liu; Nizet, ao inibirem a  $\beta$ -hemolisina também verificaram que houve bloqueio da citólise e a apoptose de macrófagos, restaurando a morte fagocítica.

A morfologia das HUVEC infectadas com GBS90356 demonstrou que as junções intercelulares foram desfeitas quando não houve a suplementação com glicose. Dados de Stoner *et al.* (2015) também demonstraram que a  $\beta$ -hemolisina desfez as junções de hBMEC auxiliando na pinocitose. Relatos com células epiteliais ME180 (linhagem celular epitelial cervical humana) e Caco2 (linhagem celular epitelial de carcinoma do cólon humano) verificaram que o *S. agalactiae* foi capaz de atravessar a monocamada epitelial, associando-se a junções intercelulares,

especificamente nas proteínas juncionais (SORIANO *et al.*, 2006). Além disso, a indução da apoptose pela amostra GBS90356 em HUVEC foi demonstrada previamente pelo nosso grupo (OLIVEIRA *et al.*, 2018). Contudo, experimentos de indução da apoptose após suplementação com glicose encontram-se em andamento por citometria de fluxo em nosso laboratório.

De acordo com os dados obtidos neste estudo, podemos concluir que a implantação da tipagem capsular por PCR multiplex foi bem sucedida, caracterizando os diferentes TC de diversos hospedeiros. Além disso, os resultados mostraram que o sequenciamento da amostra GBS90356 pertencente ao TC III / ST-17 tornou possível análises comparativas que contribuem para o melhor entendimento a nível molecular da patogênese ocasionada por esse patógeno. A inibição da  $\beta$ -hemolisina reduziu o perfil de aderência e invasão da amostra GBS90356 e desfez as junções intercelulares do endotélio, sugerindo o envolvimento no processo de disseminação bacteriana. Nosso trabalho evidenciou que a virulência do *S. agalactiae* e a capacidade em causar diferentes infecções dependem da expressão de importantes fatores de virulência, reforçando a necessidade de estudos mais aprofundados, inclusive das proteínas não caracterizadas.

**CONCLUSÕES:**

- a) dentre as amostras humanas não oncológicas de *S. agalactiae*, o TC III foi o mais prevalente, enquanto nas amostras de pacientes com câncer o TC Ia foi o mais frequentemente isolado. Já nas amostras bovinas o TC mais frequente foi o TC II;
- b) a implantação da tipagem capsular por PCR multiplex foi bem sucedida;
- c) o sequenciamento genômico da amostra de *S. agalactiae* TCIII / ST-17 resultou em um conteúdo G+C de 35,2% e 2.03Mpb;
- d) a análise comparativa demonstrou a heterogeneidade da amostra sequenciada, GBS90356, frente às demais amostras de diferentes tipos capsulares com genoma completo depositados no NCBI;
- e) amostras bacterianas da mesma sequência tipo mostraram relações evolutivas próximas e as ramificações de amostras ST-17 humanas foram agrupadas juntas, indicando um gargalo evolutivo;
- f) o pangenoma de *S. agalactiae* é aberto apresentando uma alta plasticidade;
- g) foram descritas, através do software GIPSY, 10 ilhas de patogenicidade com um total de 352 genes que corresponderam a 17% do tamanho do genoma;
- h) as PAIs 2 e 5 foram parcialmente detectadas nas amostras do ST-17 e relacionadas à integrases;
- i) as PAIs mostraram importantes genes de fatores de virulência para *S. agalactiae* (*neu*, *cps*, *dlt*, *fb*s, *cfb*, *lmb*);
- j) 20 proteínas com peptídeos sinal foram detectadas pela ferramenta SignalP, sugerindo que a maioria das proteínas envolvidas na patogenicidade de *S. agalactiae* não é secretada;
- k) Gene Orthology encontrou 102 genes para processos biológicos, 105 genes para componentes celulares e 235 genes relacionados com funções moleculares para os genes das PAIs;
- l) a inibição da atividade  $\beta$ -hemolítica da GBS90356 reduz o perfil de aderência e invasão da amostra GBS90356, sugerindo que esta é

requerida no processo de disseminação de *S. agalactiae* em HUVEC;

m) mudanças na morfologia, sinais de apoptose (dados não mostrados) e junções intercelulares do endotélio desfeitas foram verificados na interação HUVEC X *S. agalactiae* (GBS90356).

## REFERÊNCIAS

- AGUIAR, E.L. *et al.* Complete genome sequence of *Streptococcus agalactiae* strain GBS85147 serotype of type Ia isolated from human oropharynx. **Stand Genomic Sci.**, v. 3, p. 1–8, 2016.
- AILA EI, N.A. *et al.* Comparison of different sampling techniques and of different culture methods for detection of group B streptococcus carriage in pregnant women. **BMC Infect Dis.**, v. 29, p. 285, 2010.
- AIRD, W.C. Endothelium as an organ system. **Crit. Care. Med.**, v. 32, p. S271–S279, 2004.
- AKHTAR, M.S.; KRISHNAN, M.Y.; BHAKUNI, V. Insights into the mechanism of action of hyaluronate lyase: role of C-terminal domain and Ca<sup>2+</sup> in the functional regulation of enzyme. **The Journal of Biological Chemistry**, United States, v. 281, n.38, p. 28336–28344, 2006.
- ALIKHAN, N-F. *et al.* BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. **BMC Genomics**, v. 12, p. 402. 2011
- ALBERTS, B. *et al.* **Biologia molecular da célula**. 5. ed. Porto Alegre: Artmed, 2004.
- ALMEIDA, A. *et al.* Detection and discrimination of common bovine mastitis causing streptococci. **Veterinary Microbiology**, Amsterdam, v. 164, p. 370–377, 2013.
- AMIN, A.; ABDULRAZZAQ, Y.M.; UDUMAN, S. Group B streptococcal serotype distribution of isolates from colonized pregnant women at the time of delivery in United Arab Emirates. **J. Infect.**, v. 45, p. 42–46, 2002.
- AMUNDSON, S.A. *et al.* Stress- specific signatures: expression profiling of p53 wild-type and –null human cells. **Oncogenes**, v. 24, p. 4572-4579, 2005.
- AMBORSKI, R. *et al.* A non-hemolytic, group B streptococcus infection in cultured bullfrogs, *Rana catesbeiana*, in Brazil. **J Wildlife Dis.**, 1983.
- ANDREWS, J.L. *et al.* Group B streptococci causing neonatal bloodstream infection: antimicrobial susceptibility and serotyping results from SENTRY centers in the Western Hemisphere. **Am J Obstet Gynecol.**, v. 183, p. 859–862, 2000
- AREAL, A. *et al.* Infecção perinatal por *Streptococcus agalactiae* pode ser evitada: Prevalência da colonização em parturientes no Hospital São Marcos, factores de risco e a sua relação com a infecção perinatal. **Acta Pediátrica Portuguesa**, n. 41, p. 16-21, 2010.
- ARAVIND, L. *et al.* The many faces of the helix-turn-helix domain: transcription regulation and beyond. **FEMS Microbiology Reviews**, v. 29, p. 231–262, 2005.



ARNDT, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. **Nucleic Acids Research**, v. 44, n. W1, p. 16–21, 2016.

AZEVEDO, J. Mastite cresce no verão. **Prod. Parmalat.**, v. 1, n. 10, p. 50-53, 1997.

AZIZ, R.K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. **BMC Genomics**, 9, 75, 2008.

BAKER, M. De novo genome assembly: what every biologist should know. **Nature methods**, v. 9, n. 4, p. 333-337, 2012.

BAKER, C.J.; EDWARDS, M.S. Group B streptococcal conjugate vaccines. **Arch Dis Child**, v. 88, p. 375–378, 2003.

BALLARD, M.S. *et al.* International Bacteremia Surveillance Collaborative The changing epidemiology of group B Streptococcus bloodstream infection: a multi-national population-based assessment. **Infectious Diseases**, v. 48, 386–391; 2016.

BALDAUF, S.L. Phylogeny for the faint of heart: a tutorial. **Trends Genet.**, v. 19, n. 6, p. 345–351, 2003.

BALTRUS, D.A. *et al.* Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. **PLoS Pathog.**, v. 7, p. e1002132, 2011.

BANKEVICH, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455-477, 2012.

BARAÚNA, R.A. *et al.* Assessing the Genotypic Differences between Strains of *Corynebacterium pseudotuberculosis* biovar equi through Comparative Genomics. **PLOS ONE**, v. 17, n. 6, p.76-80, 2017.

BARON, M.J. *et al.*  $\alpha$ C protein of Group B Streptococcus binds host cell surface glycosaminoglycan and enters cells by an actin-dependent mechanism. **J Biol Chem.**, v. 279, p.24714–24723, 2005.

BARROS, R.R.; SOUZA, A.F.; BAPTISTA, F.O.L. Polyclonal spread of *Streptococcus agalactiae* resistant to clindamycin among pregnant women in Brazil. **J Antimicrob Chemother**, v. 71, n. 7, p. 2054-2056, 2016.

BASSET, C. *et al.* Innate immunity and pathogen-host interaction. **Vaccine**, v. 21, p. 12-23, 2003.

BECKMANN, C. *et al.* Identification of novel adhesions from group B Streptococci by use of phage display reveals C5a peptidase mediates fibronectin binding. **Infect. Immun.**, v. 70, n. 6, p. 2869-2876, 2002.

BEITUNE, E.L.P. *et al.* Colonization by *Streptococcus agalactiae* during pregnancy: Maternal and Perinatal prognosis. **Braz J Infec Disease**, v. 9, n. 3, p. 276-282, 2005.

BELLAIS, S. A. *et al.* Capsular Switching in Group B *Streptococcus* CC17 Hypervirulent Clone: A Future Challenge for Polysaccharide Vaccine Development. **The Journal of Infectious Diseases**, v. 206, p. 1745-1752, 2012.

BENDTSEN, J.D. *et al.* Improved prediction of signal peptides: SignalP 3.0. **J Mol Biol.**, v.16, n. 340, p. 783-795, 2004.

BERTELLI, C. *et al.* IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. **Nuc. Acids Research**, v. 1, p. 1-7, 2017.

BERTI, F. *et al.* Structure of the type IX group B *Streptococcus* capsular polysaccharide and its evolutionary relationship with types V and VII. **J Biol Chem.**, v. 289, n. 34, p. 23437-23448, 2014.

BISHOP, E.J. *et al.* Necrotizing fasciitis in captive juvenile *Crocodylus porosus* caused by *Streptococcus agalactiae*: an outbreak and review of the animal and human literature. **Epidemiol Infect.**, v. 135, p. 1248-1255, 2007.

BISCHARAT, N.D.W. *et al.* Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor. **J. Clin. Microbiol.**, v. 42, p. 2161–2167, 2004.

BODASZEWSKA-LUBAS, M. *et al.* Adherence of Group B *Streptococci* to Human Rectal and Vaginal Epithelial Cell Lines in relation to Capsular Polysaccharides as well as Alpha-like Protein Genes – Pilot Study. **Polish Journal of Microbiology.**, v. 55, p. 96-112, 2013.

BOGATCHEVA, N.V.; GARCIA, J.G.N.; VERIN, A.D. Molecular mechanisms of thrombin-induced cell permeability. **Biochem.**, v. 67, p. 75-84, 2002.

BOLAÑOS, M. *et al.* Distribution of *Streptococcus agalactiae* serotypes in samples from non-pregnant adults. **Clin. Microbiol. Newsl.**, v. 27, p. 151–153, 2005.

BOLGER, A.M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.

BORGER, I.L. *Streptococcus agalactiae* em gestantes: prevalência de colonização e avaliação da suscetibilidade aos antimicrobianos. **Rev Bras Ginecol Obstet.**, v. 25, n. 10, p. 575-579, 2005.

BOTELHO, A.C.N. *et al.* *Streptococcus agalactiae* carriage among pregnant women living in Rio de Janeiro, Brazil, over a period of eight years. **PLoS ONE.**, v.13, n. 5, p. e0196925. 2018.

BRAGA, P.A.C. *et al.* Bacterial identification: from the agar plate to the mass spectrometer. **RSC Advance: an international journal to further the chemical sciences**, Cambridge, v. 3, p. 994–1008, 2013.

BROCHET, M. *et al.* Genomic diversity and evolution within the species *Streptococcus agalactiae*. **Microbes Infect.**, v. 8, p. 1227–1243, 2006.

BROCHET, M. *et al.* Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. **J Bacteriol.**, v. 190, p. 6913–6917, 2008.

BROCHET, M. *et al.* Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. **Mol. Microbiol.**, v. 71, p. 948–959, 2009.

BROWN, D. Structure and function of membrane rafts. **Int J Med Microbiol.**, v. 291, n. 6–7, p. 433–437, 2002.

BROWN, C.K. *et al.* Structure of the Streptococcal cell wall C5a peptidase. **Proc. Natl. Acad. Sci.**, v. 102, p. 18391–18396, 2005.

BRUEGGEMANN, A.B. *et al.* Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. **J Infect Dis.**, v. 187, p. 1424–1432, 2003.

BUSCETTA, M. *et al.* FbsC, a novel fibrinogen-binding protein, promotes *Streptococcus agalactiae*- host cell interactions. **J Biol Chem.**, v. 289, n. 30, p. 21003–21015, 2014.

CAMPISI, E. *et al.* Genomic Analysis Reveals Multi-Drug Resistance Clusters in Group B *Streptococcus* CC17 Hypervirulent Isolates Causing Neonatal Invasive Disease in Southern Mainland China. **Frontiers in Microbiology.**, v. 7, p. 1265, 2016.

CAPANNA, F. *et al.* Antibiotic resistance patterns among group B *Streptococcus* isolates: implications for antibiotic prophylaxis for early-onset neonatal sepsis. **Swiss medical weekly.**, v. 143, p. 13778, 2013

CAPANEMA, X.E.R. *et al.* Brazilian genome sequencing projects: state of the art. Recent patents on DNA end gene sequences. **Recent Pat DNA Gene Seq.**, v. 2, n. 2, p. 111–132, 2008.

CAPOSIO, P. *et al.* A novel role of the interferon-inducible protein IFI16 as inducer of proinflammatory molecules in endothelial cells. **J. Bio. Chem.**, v. 282, p. 33515–33529, 2007.

CARLIN, A.F. *et al.* Group B *Streptococcus* suppression of phagocyte functions by protein-mediated engagement of human Siglec-5. **J Exp Med.**, v. 206, n.8, p. 1691–9, 2009.

CARVALHO-CASTRO, G.A. *et al.* Molecular epidemiology of *Streptococcus agalactiae* isolated from mastitis in Brazilian dairy herds. **Braz J Microbiol.**, v. 48, p. 551–559, 2017.

CARVER, T. *et al.* DNAPlotter: circular and linear interactive genome visualization. **Bioinformatics**, v. 25, p. 119-120, 2012.

CHATTOPADHYAY, D. *et al.* Phylogenetic Lineage and Pilus Protein Spb1/SAN1518 Affect Opsonin- Independent Phagocytosis and Intracellular Survival of Group B Streptococcus. **Microbes and infection / Institut Pasteur**, v. 13, n.4, p. 369-382, 2011.

CHENG, Q. *et al.* The Group B Streptococcal C5a peptidase is both a specific protease and an invasin. **Infect Immun.**, v. 70, p. 2408–2413, 2002.

CHOU, L. *et al.* Phenotypic and genotypic heterogeneity among Streptococcus iniae isolates recovered from cultured and wild fish in North America, Central America and Caribbean islands. **Journal of Aquatic Animal Health**, v. 26, p. 263–271, 2018.

CHRISTIE, R.; ATKINS, N.E.; MUNCH-PETERSEN, E.A. Note on a lytic phenomenon shown by group B streptococci. **Aust J Exp Biol Med Sci.**, v. 22, p. 197–200, 1944.

CHURCHWARD, J.R.S.G.G. Conjugative transposition. **Annual Review in Microbiology**, v. 49, p. 367-397, 1995.

CIESLEWICZ, M.J. *et al.* Structural and genetic diversity of group B Streptococcus capsular polysaccharides. **Infect. Immun**, v. 73, n. 5, p. 3096-3103, 2005

COLLINS, T.S. *et al.* Group B streptococcal colonization in a developing country: its association with sexually transmitted disease and socioeconomic factors. **Am. J. Trop. Med. Hyg.**, v. 59, p. 633-636, 1998.

CONESA, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674-3676, 2005.

COSTA, A.L. *et al.* Prevalence of colonization by group B Streptococcus in pregnant women from a public maternity of Northwest region of Brazil. **Rev Bras Ginecol Obstet.**, v. 30, p. 274-280, 2008.

COX, F. *et al.* Prevention of group B streptococcal colonization and bacteremia in neonatal mice with topical vaginal inhibitors. **J. Infect. Dis.**, v. 167, p. 1118-1122, 1993.

CRESPO-ORTIZ, M del P. *et al.* Emerging trends in invasive and noninvasive isolates of Streptococcus agalactiae in a Latin American hospital: a 17-year study. **BMC Infectious Diseases**, v. 14, p. 428, 2014.

CUNHA, V. *et al.* Streptococcus agalactiae clones infecting humans were selected and fixed through the extensive use of tetracycline. **Nature communications**, v. 5, p. 4544, 2014.

- DANGWETNGAM, M. *et al.* Serotype distribution and antimicrobial susceptibilities of *Streptococcus agalactiae* isolated from infected cultured tilapia (*Oreochromis niloticus*) in Thailand: Nine-year perspective. **J Med Microbiol.**, v.65, n.3, p.247-254, 2016.
- DAY, W.H.E. Computational complexity of inferring phylogenies from dissimilarity matrices. **Bull Math Biol.**, v. 49, n. 4, p. 461–467, 1987.
- DARLING, A.E.; MAU, B.; PERNA, N.T. Progressive Mauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. **PLoS ONE**, v. 5, n. 6, p. e11147, 2010. DOI: 10.1371/journal.pone.0011147.
- DAVIDSON, S.M. Endothelial mitochondria and heart disease. **Cardiovasc Res.**, v. 88 p. 58–66, 2010.
- DEUTSCHER, J.; FRANCKE, C.; POSTMA, P.W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. **Microbiology and molecular biology reviews**, v. 70, p. 939–1031, 2006.
- DIEDRICK, M.J. *et al.* Clonal Analysis of Colonizing Group B *Streptococcus*, Serotype IV, an Emerging Pathogen in the United States. **Journal of Clinical Microbiology**, v. 48, p. 3100-3104, 2010.
- DOARE, K.L.; HEATH, P. T. An overview of global GBS epidemiology. **Vaccine**, v.31, p.D7-D12, 2013.
- DONATI, C. *et al.* Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. **Genome Biol.**, v. 11, p.107, 2010.
- DONDERS, G.G. *et al.* Maternal immunization with an investigational Trivalent Group B *Streptococcal* vaccine: a randomized controlled trial. **Obstet. Gynecol.** v. 127, p. 213–221, 2016.
- DORAN, K.S.; NIZET, V. Molecular pathogenesis of neonatal group B *Streptococcal* infection: no longer in its infancy. **Mol. Microbiol.**, v. 54, n. 1, p. 23-31, 2004.
- DRAMSI, S. *et al.* Assembly and role of *pili* in Group B *Streptococci*. **Mol Microbiol.**, v. 60, p. 1401–1413, 2006.
- DUARTE, R.S. *et al.* Phenotypic and molecular characteristics of *Streptococcus agalactiae* isolates recovered from milk of dairy cows in Brazil. **Journal of Clinical Microbiology**, v. 42, p. 4214-4222, 2004.
- DUTRA, V. G. *et al.* *Streptococcus agalactiae* in Brazil: serotype distribution, virulence determinants and antimicrobial susceptibility. **BMC Infectious Diseases**, v. 14, p. 323, 2014.
- D'URZO, N. *et al.* Acidic pH Strongly Enhances In Vitro Biofilm Formation by a Subset of Hypervirulent ST-17 *Streptococcus agalactiae* Strains. **Applied and Environmental Microbiology**, v. 8, n. 7, p. 2176-2185, 2014.

- EDMOND, K.M. *et al.* Group B streptococcal disease in infants aged younger than 3 months: systematic review and metaanalysis. **Lancet**, v. 379, p. 547–556, 2012.
- EDWARDS, M.S.; BAKER, C. J. Streptococcus agalactiae (group B Streptococcus). In: MENDELL, G. L.; BENNET, J. E.; DOLIN, R. **Principles and practice of infectious diseases**. 5 ed. Philadelphia: Churchill Livingstone, 2001. p. 2156-2167.
- ELLING, R. *et al.* Synchronous recurrence of group B streptococcal late-onset sepsis in twins. **Pediatrics**, v. 133, p. e1388-91, 2014.
- EI-METWALLY, S. *et al.* Next-Generation Sequence Assembly: Four Stages of Data Processing and Computational Challenges. **PLoS Computational Biology**, 9, n. 12, p. e1003345, 2013.
- ESCOLANO, S.S. *et al.* Infección tardía por estreptococo grupo B en gemelos nacidos mediante cesárea. **An Pediatr.**, v. 13, p. 33-39, 2014.
- EVANS, J. *et al.* Phylogenetic relationships among Streptococcus agalactiae isolated from piscine, dolphin, bovine and human sources: a dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. **J Med Microbiol.**, v. 37, p. 54–55, 2008.
- FACKLAM, R. What Happened to the Streptococci: Overview of Taxonomic and Nomenclature Changes. **Clinical Microbiology Reviews**, Washington, v. 15, n. 4, p. 613-630, 2002.
- FALLAVENA, P.R. *et al.* The influences of CD14 -260C>T polymorphism on survival in ICU critically ill patients. **Immunology Investigation**, v. 38, p. 797-811, 2009.
- FARLEY, M.M. Group B Streptococcal disease in nonpregnant adults. **Clin. Infect. Dis.**, v. 33, p. 556-561, 2001.
- FEIL, E.J. *et al.* eBurst: inferring patterns of evolutionary descent among cluster of related bacterial genotypes from multilocus sequence typing data. **Journal of Bacteriology**, Washington, v. 186, n. 5, p. 1518-1530, 2004.
- FELSENSTEIN, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. **Evolution**, v. 39, n. 4, p. 783, 1985.
- FETTUCCIARI, K. *et al.* Group B streptococcus induces apoptosis in macrophages. **J Immunol.**, v.165, p. 3923–3933, 2011
- FIOLO, K. *et al.* Infection rate and Streptococcus agalactiae serotypes in samples of infected neonates in the city of Campinas (São Paulo), Brazil. **Revista Brasileira de Ginecologia e Obstetria**, v. 34, n. 9, p. 544-549, 2012.
- FITCH, W.M.; MARGOLIASH, E. Construction of Phylogenetic Trees. **Science**, v. 155, n. 3760, p. 279–284, 1967.

- FLORINDO, C. *et al.* Epidemiological surveillance of colonising 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. **Euro Surveillance**, v. 19, n. 23, p. 20825, 2014.
- FREY, M.N. *et al.* Streptococcus agalactiae involved in the etiology of Sexually Transmitted Diseases. **An Bras Dermatol.**, v. 86, p.1205–7, 2011.
- FRY, R.M. Fatal infections by hemolytic streptococcus group B. **Lancet**, v. 231, p. 199–201. 1938.
- FURFARO, L. *et al.* Group B streptococcus prevalence, serotype distribution and colonization dynamics in Western Australian pregnant women. **J. Med. Microbiol.**, v. 68, n. 5, p. 728-740, 2019.
- FUNAMOTO, S. *et al.* Biomaterials preparation and characterization of decellularized cornea using high-hydrostatic pressurization for corneal tissue engineering. **Biomaterials**, v. 31, n. 14, p. 3941–3948, 2017.
- GARCIA, M.V. *et al.* Mechanism of infection and colonization of *Rhipicephalus sanguineus* eggs by *Metharhizium anisopliae* as revealed by scanning electron microscopy and histopathology. **Brazilian Journal of Microbiology**, v.36, p. 368-372, 2008.
- GAL-MOR, O.; FINLAY, B.B. Pathogenicity islands: a molecular toolbox for bacterial virulence. **Cellular Microbiology**, Oxford, v. 8, n. 11, p. 1707-1719, 2006.
- GAO, J. *et al.* Antibiotic resistance of Streptococcus agalactiae from cows with mastitis. **The Veterinary Journal**, n. 194, p. 423-424, 2014.
- GENDRIN, C. *et al.* Mast cell chymase decreases the severity of group B Streptococcus infections. **J Allergy Clin Immunol.**, v. 142, n. 1, p. 120-129, 2018.
- GIACHETTO, P.F.; HIGA, R.H. Bioinformática Aplicada na Agricultura. In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (Ed.). **Tecnologias da informação e comunicação e suas relações com a agricultura**. 4. ed. Brasília, DF: Embrapa, 2015.
- GIBBS, R.S.; SCHRAG, S.; SCHUCHAT, A. Perinatal infections due to group B streptococci. **Obstet Gynecol.**, v. 104, p. 1062–1076, 2004.
- GIZACHEW, M. *et al.* Streptococcus agalactiae maternal colonization, antibiotic resistance and serotype profiles in Africa: a meta-analysis. **Annals of clinical microbiology and antimicrobials**, v. 18, n. 1, p. 14, 2019.
- GLASER, P. *et al.* Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease. **Molecular Microbiology**, v. 45, p. 1499-1513, 2002.

GLENN, T.C. Field guide to next-generation DNA sequencers. **Mol. Ecol. Resour.**, v. 11, n. 5, p. 759-769, 2011.

GLÜCK, T. *et al.* Parameters influencing membrane CD14 expression and soluble CD14 levels in sepsis. **European Journal of Medical Research**, v. 6, p. 351–358, 2001.

GUTEKUNST, H.; EIKMANN, B.J.; REINSCHIED, D.J. The novel fibrinogen binding protein FbsB promotes *Streptococcus agalactiae* invasion into epithelial cells. **Infect. Immun.**, v. 72, p.3495–3504, 2004.

HAMILTON, A. *et al.* Penicillin-binding protein 1a promotes resistance of group B *Streptococcus* to antimicrobial peptides. **Infection and Immunity**, Washington, v. 74, n. 11, p. 6179-6187, 2006.

HARRIS, T.O. *et al.* A novel streptococcal surface protease promotes virulence, resistance to opsonophagocytosis, and cleavage of human fibrinogen. **J Clin Invest.**, v. 111, p. 61–70, 2003.

HAVARSTEIN, L.S.; DIEP, D.B.; NES, I.F. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. **Mol Microbiol.**, v. 16, p. 229-240, 1995.

HE, J. *et al.* RNA-Seq reveals enhanced sugar metabolism in *Streptococcus mutans* Co-cultured with *Candida albicans* within mixed-species biofilms. **Front. Microbiol.**, v. 8, p.1036, 2017.

HELLERQVIST, C.G.; SUNDELL, H.; GETTINS, P. Molecular basis for group B beta-hemolytic streptococcal disease. **Proc Natl Acad Sci U S A.**, v. 84, p. 51-55, 1987.

HENNEKE, P. *et al.* Role of lipoteichoic acid in the phagocyte response to group B streptococcus. **J. Immunol.**, v. 174, p. 6449-6455, 2005.

HEATH, P.T. Status of vaccine research and development of vaccines for GBS. **Vaccine**, v. 34, p. 2876–2879, 2016. DOI: 10.1016/j.vaccine.2016.12.072.

HENSON, J.; TISCHLER, G.; NING, Z. Next-generation sequencing and large genome assemblies. **Pharmacogenomics**, v. 13, n. 8, p. 901-915, 2012.

HERBERT, M.A. *et al.* Genetic islands of *Streptococcus agalactiae* strains NEM316 and 2603VR and their presence in other Group B Streptococcal strains. **BMC Microbiology**, v. 5, p. 31, 2005.

HICKMAN, M.E. *et al.* Changing epidemiology of group B Streptococcal colonization. **Pediat.**, v. 10, n. 2, p. 203-209, 1999.

HO, Y.R. *et al.* Variation in the number of tandem repeats and profile of surface protein genes among invasive group B Streptococci correlates with patient age. **J Clin Microbiol.**, v. 45, p. 1634-1636, 2007.



HO, Y.R. *et al.* The enhancement of biofilme formation in Group B streptococcal isolates at vaginal pH. **Med: Microbiol. Immunol.**, v. 202, p. 105-115, 2013.

HONIG, E. *et al.* The epidemiology of vaginal colonization with group B Streptococci in a sexually transmitted disease clinic. **Obstet. Gynecol.**, p. 177-180, 2002.

HSIAO, W.W.L. *et al.* Evidence of a Large Novel Gene Pool Associated with Prokaryotic Genomic Islands. **PLoS Genet.**, v. 1, n. 5, p. e62, 2005.

HYNES, W.L.; WALTON, S.L. Hyaluronidases of Gram-positive bacteria. **FEMS Microbiol. Lett.**, v.183, p. 201–207, 2000.

HUTCHISON III, C.A. DNA sequencing: bench to bedside and beyond. **Nucleic Acids Res.**, v. 35, p. 6227-6237, 2007.

IMPERI, M. *et al.* A multiplex PCR assay for the direct identification of the capsular type (Ia to IX) of *Streptococcus agalactiae*. **J Microbiol Methods.**, v. 80, p. 212-214, 2010.

JAFFE, E.A. *et al.* Antibiotic administration to treat possible occult bacteremia in febrile children. **N Engl J Med.** v. 317, n. 19, p. 1175-1180, 1987 *apud* JIETANG, M. *et al.* An evolving new paradigm: endothelial cells--conditional innate immune cells. **Journal of hematology & oncology**, v. 6, p. 61, 2013.

JAFFE, E.A. *et al.* Culture of human endothelial cells derived from umbilical veins. Identification of morphologic and immunologic criteria. **J. Clin. Invest.**, v. 221, p. 281-283, 1973.

JANNECHEVITZ, F.F. *et al.* Caracterização genética e genômica de *Helicobacter pylori* / Genetic and genomic characterization of *helicobacter pylori*. **MIOC**, Rio de Janeiro. v.73, p. 188, 2015.

JERLSTRÖM, P.G.; CHHATWAL, G.S.; TIMMIS, K.N. The IgA-binding beta antigen of the c protein complex of group B streptococci: sequence determination of its gene and detection of two binding regions. **Mol Microbiol.**, v. 5, n. 4, p. 843–849, 1996.

JIANG, S.M. *et al.* Regulation of virulence by a twocomponent system in Group B Streptococcus. **J Bacteriol.**, v.187,p.1105–1113, 2005.

JIETANG, M. *et al.* An evolving new paradigm: endothelial cells--conditional innate immune cells. **Journal of hematology & oncology**, v. 6, p. 61, 2013.

JONSSON, I. *et al.* Role of Fibrinogen- Binding Adhesin Expression in Septic Arthritis and Septicemia Caused by *Streptococcus agalactiae*. **Journal of Infectious Disease**, United States, v. 192, p. 1456-1464, 2005.

JOHRI, A.K. *et al.* Epidemiology of Group B Streptococcus in developing countries. **Vaccine**, v. 31, p. D43–D45, 2013.

- JOHRI, A.K.L. *et al.* Group B Streptococcus: global incidence and vaccine development. **Vaccine**, v. 4, p. 932-942, 2006.
- JONES, N. *et al.* Multilocus sequence typing system for group B Streptococcus. **J Clin Microbiol.**, v. 41, p. 2530–2536, 2003.
- JONES, N.K.A. *et al.* Enhanced Invasiveness of Bovine-Derived Neonatal Sequence Type 17 Group B Streptococcus Is Independent of Capsular Serotype. **Enhanced Invasiveness of Neonatal GBS**, v. 42, p. 915-924, 2006.
- JUHAS, M. *et al.* Genomic islands: tools of bacterial horizontal gene transfer and evolution. **FEMS microbiology reviews**, V. 33, n. 2, p. 376-393, 2009.
- KAAMBWA, B.; BRYAN, B.; GRAY, J. Cost-effectiveness of rapid tests and other existing strategies for screening and management of early-onset group B Streptococcus during labour. **BJOG**, v. 117, p. 1616-1627, 2010.
- KANG, H.M. *et al.* Genotype Characterization of Group B Streptococcus Isolated From Infants With Invasive Diseases in South Korea. **Pediatr Infect Dis J.**, v. 36, n. 10, p. e242-e247, 2017.
- KANNIKA, K. *et al.* Molecular serotyping, virulence gene profiling and pathogenicity of Streptococcus agalactiae isolated from tilapia farms in Thailand by multiplex PCR. **J Appl Microbiol.**, v.122, n. 6, p. 1497-1507, 2017.
- KAPLAN, A. *et al.* Group B streptococcus induces trophoblast death. **Microb Pathog.**, v. 45, p. 231–235, 2008.
- KARATAN, E.; WATNICK, P. Signals, Regulatory Networks and Materials That Build and Break Bacterial Biofilms. **Microb Pathog.**, v. 73, p. 310-347, 2009.
- KARUNAKARAN, R. *et al.* Group B Streptococcus infection: epidemiology, serotypes, and antimicrobial susceptibility of selected isolates in the population beyond infancy (excluding females with genital tract- and pregnancy-related isolates) at the University Malaya Medical Centre. **J Infect Dis.**, v. 62, p. 192-194, 2009.
- KAYANSAMRUJ, P. *et al.* Molecular characterization and virulence gene profiling of pathogenic *Streptococcus agalactiae* populations from tilapia (*Oreochromis* sp.) farms in Thailand. **Journal of veterinary Diagnostic Investigation**, v. 26, n. 4, p. 488-495, 2014.
- KEEFE, GP. Streptococcus agalactiae mastitis: a review. **Can Vet J.**, v. 38, p. 429–437, 1997.
- KEEFE, G. Update on Control of Staphylococcus aureus and Streptococcus agalactiae for management of mastitis. **Veterinary Clinics Food Animal**, Philadelphia, v. 28, p. 203–216, 2012.
- KELLER, T.T., *et al.* Infections and endothelial cells. **Card Res.**, v. 60, p. 40-48, 2003.

KONEMAN, E.W. *et al.* The gram-positive cocci: Part II: Streptococci, Enterococci, and the “Streptococci-like” Bacteria. In: KONEMAN, E. W. *et al.* **Color Atlas and Textbook of Diagnostic Microbiology**. 5th ed. USA: ED Lippincott, 1997. p. 577-603

KONG, F. *et al.* Use of phenotypic and molecular serotype identification methods to characterize previously nonserotypeable group B streptococci. **J Clin Microbiol.**, v. 46, p. 2745-2750, 2008.

KONG, F. *et al.* Molecular Profiles of Group B Streptococcal Surface Protein Antigen Genes: Relationship to Molecular Serotypes. **Journal of Clinical Microbiology**. v. 40, p. 620-626, 2002.

KONG, F. *et al.* Serotype Identification of Group B Streptococci by PCR and Sequencing. **Journal of Clinical Microbiology**, v. 40, p. 216-226, 2002.

KORIR, M.L. *et al.* Association and virulence gene expression vary among serotype III group B streptococcus isolates following exposure to decidual and lung epithelial cells. **Infect. Immun.**, v. 82, n. 11, p. 4587-4595; 2016.

KONTO-GHIORGHI, Y. *et al.* Papel duplo para *Pilus* na adesão às células epiteliais e formação de biofilme em *Streptococcus agalactiae*. **PLoS Pathogens**, v. 5, n. 5, p. e1000422, 2009.

LAMAGNI, T.L. *et al.* Emerging Trends in the Epidemiology of Invasive Group B Streptococcal Disease in England and Wales, 1991–2010. **Clinical Infectious Diseases**, v. 57, p. 682–688, 2013.

LAMBERTSEN, L. *et al.* Characterization of invasive group B streptococci from adults in Denmark 1999 a 2004. **Eur J Clin Microbiol. Infect Dis.**, v. 26, n. 9, p. 1071-1077, 2010.

LAMY, M-C. S. *et al.* Rapid detection of the “highly virulent” group B Streptococcus ST-17 clone. **Microbes and Infection.**, v. 8, p. 1714-1722, 2006.

LANCEFIELD, R. C. A serological differentiation of specific types of bovine hemolytic streptococci (group B). **J. Exp. Med.**, v. 59, p. 441-458, 1934.

LANG, S.; PALMER, M. Characterization of *Streptococcus agalactiae* CAMP Factor as a Pore-forming Toxin. **J Biol Chem.**, v. 278, p. 38167–38173, 2003.

LANGONI, H. *et al.* Tratamento de Mastite Bovina com Amoxicilina e Enrofeoxarina Bem como sua associação. São Paulo. **Arquivos do Instituto Biológico**, v. 67, p. 177-180, 2000.

LANOTTE, P. *et al.* Proteomic Biomarkers Associated with *Streptococcus agalactiae* Invasive Genogroups. **PLOS ONE**, v.8, n.1, p.1-11, 2013.

LAUER, P. *et al.* Genome analysis reveals *pili* in Group B Streptococcus. **Science**, v. 309, p. 105, 2005.

LAW, M.R. *et al.* The prevention of neonatal group B streptococcal disease: a report by a working group of the Medical Screening Society. **J Med Screen.**, v. 12, p. 60-8, 2005.

LEFÉBURE, T.; STANHOPE, M.J. Evolution of the core and pan-genome of Streptococcus: positive selection, recombination, and genome composition. **Genome Biology**, v. 8, p. 71, 2007.

LE DOARE, K.; HEATH, P.T. An overview of global GBS epidemiology. **Vaccine Review**, v. 31, p. D7–D12, 2013.

LEMIEUX, J. *et al.* Roles of the sortases of Streptococcus pneumoniae in assembly of the RlrA pilus. **Journal of Bacteriology**, v. 190, n. 17, p. 6002-6013, 2008.

LIBSTER, R. *et al.* Long-term outcomes of group B streptococcal meningitis. **Pediatrics**, v. 130, p. e8–e15, 2012.

LIER, C. *et al.* Analysis of the type II-A CRISPR-Cas system of Streptococcus agalactiae reveals distinctive features according to genetic lineages. **Front. Genet.**, v. 6, p. 214, 2015.

LIU, G.Y.; NIZET, V. Extracellular virulence factors of group B Streptococci. **Front biosci.**, v. 5, n. 9, p. 1794-1802, 2004.

LIN, FP-Y. *et al.* Computational Bacterial Genome-Wide Analysis of Phylogenetic Profiles Reveals Potential Virulence Genes of Streptococcus agalactiae. **PLoS ONE**, v. 6, n. 4, p. e17964, 2011.

LIN, H-C. *et al.* Clonal dissemination of invasive and colonizing clonal complex 1 of serotype VI group B Streptococcus in central Taiwan, **Journal of Microbiology, Immunology and Infection**, v. 8, p. 25-29, 2014.

LIU, L. *et al.* Comparison of next-generation sequencing systems. **BioMed Research International**, v. 12, p. 80-88, 2012.

LIU, G.Y. *et al.* Sword and shield: linked group B streptococcal beta-haemolysin/cytolysin and carotenoid pigment function to subvert host phagocytedefense. **Proc Natl Acad Sci USA**, v. 5, p. 14491–14496, 2018.

MADZIVHANDILA, M. *et al.* Serotype distribution and invasive potential of group B Streptococcus isolates causing disease in infants and colonizing maternal-newborn dyads. **Plos One**, v. 6, p. 17861, 2011.

MAIDEN, M.C.J., *et al.* Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. **Proc Natl Acad Sci USA.**, v. 95, p. 3140-3145, 2006.

MAISEY, H.C., *et al.* Recent advances in understanding the molecular basis of Group B Streptococcus virulence. **Expert Rev Mol Med.**, v. 10, p. e. 27, 2008.

MAISEY, H.C., *et al.* Group B streptococcal *pilus* proteins contribute to adherence to and invasion of brain microvascular endothelial cells. **J Bacteriol.**, v. 189, p. 1464–1467, 2007.

MANNING, S.D., *et al.* Multilocus sequence types associated with neonatal group B streptococcal sepsis and meningitis in Canada. **J Clin Microbiol.** v. 47, p. 1143–1148, 2009.

MARCHLEWICZ, D.A.; DUNCAN, J.L. Properties of a haemolysin produced by group B streptococci. **Infect Immun.**, v. 30, p. 805–813, 1980.

MARCHLEWICZ, D.A.; DUNCAN, J.L. Lysis of erythrocytes by a haemolysin produced by group B streptococcus sp. **Infect Immun.**, v. 34, p. 787–794, 1981.

MARTINEZ, G.J. *et al.* Characterization of Streptococcus agalactiae isolates of bovine and human origin by randomly amplified polymorphic DNA analysis. **J. Clin. Microbiol.**, v. 38, p. 71–78, 2000.

MARTINS, A.F. Identificação de cocos gram-positivos. In: MARTINS, A. F. (Org.). **Bacteriologia clínica: manual de aulas práticas.** Porto Alegre: Sulina, 2010. p. 58.

MARTINS, R.P. *et al.* Prevalência e etiologia infecciosa de mastite bovina na microregião de Cuiabá, MT. **Ci. Anim. Bras.**, v. 11/n. 1, p. 181-187, 2010.

MARESSO, A.W.; SCHNEEWIND, O. Sortase as a target of anti-infective therapy. **Pharmacological Reviews**, v. 60, n. 1, p. 128-141, 2008.

MATSUBARA, K. *et al.* Seroepidemiologic studies of serotype VIII group B Streptococcus in Japan. **J. Infect. Dis.**, v 186, n. 6, p. 855–858, 2012.

MÉDIGUE, C.; MOSZER, I. Annotation, comparison and databases for hundreds of bacterial genomes. **Research in microbiology**, v. 158, n. 10, p. 724–36, 2007.

MELO, S.C., *et al.* Antimicrobial susceptibility of Streptococcus agalactiae isolated from pregnant women. **Rev Inst Med Trop.**, v. 58, p. 83–86, 2016.

MELIN, P.; EFSTRATIOU, A. Group B streptococcal epidemiology and vaccine needs in developed countries. **Vaccine**, v. 31, p. D31–D42, 2013.

MENGIST, A.; KANNAN, H.; ABDISSA, A. Prevalence and antimicrobial susceptibility pattern of anorectal and vaginal group B Streptococci isolates among pregnant women in Jimma, Ethiopia. **BMC Res Notes**, v. 19, p. 351, 2016.

MIAN, G.F. *et al.* Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. **Vet microbiol.**, v. 136, p. 180-183, 2009.

MILLER, S.I., MEKALANOS, J.J. Constitutive expression of the *phoP* regulon attenuates *Salmonella* virulence and survival within macrophages. **J Bacteriol.**, v. 172, p. 2485–2490, 1990.

MILLER, R.J.; KOREN, S.; SUTTON, G. Assembly algorithms for next-generation sequencing data. **Genomics.**, v. 95, p. 315-327, 2010.

MIRA, A., *et al.* The bacterial pan-genome: a new paradigm in microbiology. **Genomics**, v. 3, p. 35-57, 2010.

MIRANDA, P.S.D. *et al.* Biofilm formation on different pH conditions by *Streptococcus agalactiae* isolated from bovine mastitic milk. **Letters in applied microbiology**, v. 235, p. 235-243, 2018.

MITHAL, L. B. *et al.* Cord Blood Acute Phase Reactants Predict Early Onset Neonatal Sepsis in Preterm Infants. **PLoS One**, v. 12, n. 1, p. e0168677, 2017.

MIYATA, A. *et al.* Early-onset group B streptococcal disease following culture-based screening in Japan: a single center study. **J Obstet Gynaecol Res.**, v. 38, n. 8, p. 1052-1056. 2012.

MOHAMED, S.; SYED, B. A. Commercial prospects for genomic sequencing technologies. **Nature reviews Drug discovery**, v. 12, n. 5, p. 341-342, 2013.

MOLINARI, G.; CHHATWAL, G. S. Streptococcal invasion. **Curr. Opin. Microbiol.**, v. 2, p. 6-61, 1999.

MONTEIRO, G.C.T.S. *et al.* Surface carbohydrates as recognition determinants in non-opsonic interactions and intracellular viability of group B *Streptococcus* strains in murine macrophages. **Int J Mol Med**, v. 13, p. 175-180, 2004.

MORAES, A.M. *et al.* **Tecnologia do cultivo de células animais: de biofármacos e terapia gênica.** São Paulo: Rocca, 2007. 503p.

MOUSLIM, C. *et al.* Activation of the RcsC/YojN/RcsB phosphorelay system attenuates *Salmonella* virulence. **Mol Microbiol.**, v. 54, p. 386–395, 2004.

MUÑOZ, P. *et al.* Group B streptococcus bacteremia in nonpregnant adults. **Arch Intern Méd.**, v.157, p. 213-216, 1997.

MUNOZ-NAJAR, U.V.M. An operon that confers UV resistance by evoking the SOS mutagenic response in streptococcal conjugative transposon Tn5252. **Journal of Bacteriology.**, v. 181, n. 9, p. 2782-2788, 1999.

MUZZI, A. *et al.* The pan-genome: towards a knowledgebased discovery of novel targets for vaccines and antibacterials. **Drug Discov Today.**, v. 12, p. 429–439, 2007.

NAMAVAR, J.B. *et al.* The prevalence and adverse effects of group B streptococcal colonization during pregnancy. **Arch Iran Med.** v. 11, n. 6, p. 654-657, 2008.

NAGARAJAN, N; POP, M. Sequence assembly demystified. *Nature reviews. Genetics, Nature Publishing Group.*, v.14, n. 3, p. 157–167, 2013.

NAKIB, M. A. *et al.* Comparison of the Diversilab® system with multi-locus sequence typing and pulsed-field gel electrophoresis for the characterization of *Streptococcus agalactiae* invasive strains. **Journal of Microbiological Methods**, v. 85, p. 137–142, 2011.

NIELSEN, H. Predicting Secretory Proteins with SignalP. **Methods Mol Biol.**, v.16, n. 11, p. 59-73, 2017.

NIZET, V. *et al.* Group B streptococcal beta-hemolysin expression is associated with injury of lung epithelial cells. **Infect Immun.**, v. 64, p. 3818–3826, 1996.

NIZET, V.; RUBENS, C. Pathogenic Mechanisms and virulence factors for group B streptococci. **Asociación Argentina de Microbiología.** v. 14, p. 23-29, 2002.

NOCARD, M.; MOLLEREAU, R. Sur une mammite contagieuse des vaches laitières. **Ann. Inst. Pasteur.**, v. 1, p. 109-126, 1887 *apud* BISCHARAT, N., D. W. *et al.* Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor. **J. Clin. Microbiol.**, v. 42, p. 2161–2167, 2004.

OLIVEIRA, J. *et al.* Reactive oxygen species generation mediated by NADPH oxidase and PI3K/Akt pathways contribute to invasion of *Streptococcus agalactiae* in human endothelial cells. **Memorias do Instituto Oswaldo Cruz**, v. 113, n. 6, p. e140421, 2018

OSTER, G. *et al.* Prevention of group B streptococcal disease in the first 3 months of life: would routine maternal immunization during pregnancy be cost-effective? **Vaccine**, v. 32 p. 4778–4785, 2014.

OTAGUIRI, E.S. *et al.* Commensal *Streptococcus agalactiae* isolated from patients seen at University Hospital of Londrina, Paraná, Brazil: capsular types, genotyping, antimicrobial susceptibility and virulence determinants. **BMC Microbiol.** v. 13, p. 297–305, 2013.

OVIEDO, P. *et al.* Phenotypic and genotypic characterization of *Streptococcus agalactiae* in pregnant women. First study in a province of Argentina. **Brazilian Journal of Microbiology.**, v. 44, n. 1, p. 253-258, 2013.

PAI, R. *et al.* Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. **J Clin Microbiology.**, v. 44, p. 124-131, 2006.

PAILHORIES, H. *et al.* The transcription of the *neuD* gene is stronger in serotype III group B streptococci strains isolated from cerebrospinal fluid than in strains isolated from vagina. **FEMS Microbiology Letters**, v. 349, p. 71–75, 2013.

PALMEIRO, J.K. *et al.* Phenotypic and genotypic characterization of group B streptococcal isolates in southern Brazil. **J Clin Microbiol.**, v. 12, p. 4397-4403, 2010.

PANTOJA, Y. *et al.* PanWeb: A web interface for pan-genomic analysis. **PLoS ONE.** v. 12, n. 5, p. e0178154, 2017.

PATIL, K.P. *et al.* Group B streptococci colonization in pregnant women is screening necessary? **J South Asian Feder Obstet Gynecol**, v.5, p.64–67, 2013.

PEOTTA, V.A. *et al.* Human monocytes and monocyte-derived macrophages phagocytosis of serotype III group B streptococci strains. **Curr Microbiol.**, v. 43, p. 64-68, 2001.

PETERS, K. *et al.* Molecular basis of endothelial dysfunction in sepsis. **Cardiovasc. Res.**, v. 60, p. 49-57, 2003.

PÉREZ, J. *et al.* Distribution of capsular types and antimicrobial susceptibility of *Streptococcus agalactiae* causing infections in Argentina. **Rev Argent Microbiol** .v. 36, p. 63-67, 2004.

PÉRICHON, B. *et al.* Regulation of PI-2b Pilus Expression in Hypervirulent *Streptococcus agalactiae* ST-17 BM110. **PLoS ONE.**, v. 12, n. 1, p. e0169840, 2017.

PETREY, D.; HONIG, B. Structural Bioinformatics of the Interactome. **Annu Rev Biophys.**,v. 43, p. 193–210. 2014.

PEZZICOLI, A. *et al.* Pilus backbone contributes to Group B *Streptococcus* paracellular translocation through epithelial cells. **J Infect Dis.**, v. 198, p. 890–898, 2008.

PIETROCOLA, G. *et al.* FbsA, a fibrinogen-binding protein from *Streptococcus agalactiae*, mediates platelet aggregation. **Blood.**, v. 105, n. 3, p. 1052–1059, 2005.

PIETROCOLA, G. *et al.* *Streptococcus agalactiae* Non-Pilus, Cell Wall-Anchored Proteins: Involvement in Colonization and Pathogenesis and Potential as Vaccine Candidates. **Front. Immunol.**, v. 9, p. 602, 2018.

PIMENTEL, B.A. *et al.* *Streptococcus agalactiae* infection in cancer patients: a five-year study. **Eur. J. Clin. Microbiol. Infect. Dis.**, v.35, p.927-933, 2016.

PINTO, T.C.A. *et al.* Distribution of serotypes and evaluation of antimicrobial susceptibility among human and bovine *Streptococcus agalactiae* strains isolated in Brazil between 1980 and 2006. **Braz J Infect Dis.**, v. 17, p. 131–136, 2013.

PIZARRO-CERDÁ, J.; COSSART, P. Bacterial adhesion and entry into host cells. **Cell.**, v.124, p. 715–727, 2006.



POP, M.; SALZBERG, S.L. Bioinformatics challenges of new sequencing technology. **Trends in genetics.**, v. 24, n. 3, p. 142–149, 2008.

POHLHAUS, J.R.; COOK-DEEGAN, R.M. Genomics Research: World Survey of Public Funding. **BMC Genomics.**, v. 20, p. 472-489, 2008.

POYART, C. *et al.* Multiplex PCR Assay for Rapid and Accurate Capsular Typing of Group B Streptococci. **J Clin Microbiol.**, p. 1985–1988, 2007.

POYART, C. *et al.* Invasive Group B Streptococcal Infections in Infants, France. **Emerging Infectious Diseases.**, v. 14, n. 10, p. 1647-1649, 2008.

PRESENTATION, C. *et al.* Perinatal/neonatal late-onset group B streptococcal infection in identical twins: Insight to disease pathogenesis. **J Perinatol.**, v. 11 p. 16-19, 2002.

PRITZLAFF, C.A. *et al.* Genetic basis for the beta-haemolytic/cytolytic activity of group B streptococcus. **Mol Microbiol.**, v. 39, p. 236–247, 2001.

PUYMÈGE, A. *et al.* Analysis of Streptococcus agalactiae pan-genome for prevalence, diversity and functionality of integrative and conjugative or mobilizable elements integrated in the tRNA<sup>Lys</sup> CTT gene. **Mol. Genet. Genomics.**, v. 38, n. 15, p. 1031-1039, 2015.

QUAIL, M.A. *et al.* A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. **BMC Genomics**, v. 13, n. 341. 2012.

RAABE, V.N.; SHANE, A.L. Group B Streptococcus (*Streptococcus agalactiae*). **Microbiol Spectrum.**, v. 7, n. 2, p. 3-9, 2018.

RAGUNATHAN, P.; PONNURAJ, K. Expression, purification and structural analysis of a fibrinogen receptor FbsA from Streptococcus agalactiae. **Protein J.**, v. 3, p. 159-166, 2011.

RAGUNATHAN, P. *et al.* Metal Binding Is Critical for the Folding and Function of Laminin Binding Protein, Lmb of Streptococcus agalactiae. **PLoS ONE**, v. 8, n. 6, p. e67517, 2013.

RAJAGOPAL, L. Understanding the regulation of Group B Streptococcal virulence factors. **Future Microbiol.**, v. 4, p. 201-221, 2009.

RANNALA, B.; YANG, Z. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. **J Mol Evol**, v. 43, n. 3, p. 304–311, 1996.

RESCHE-RIGON, M. *et al.* Evaluating mortality in intensive care units: contribution of competing risks analyses. **Critical Care**, v. 10, p.103, 2006.

RICHARDS, V.P. *et al.* Comparative genomics and the role of lateral gene transfer in the evolution of bovine adapted *Streptococcus agalactiae*. **Infect Gen Evol.**, v. 11, n. 6, p. 1263-1275, 2011.

RICHARDS, V.P. *et al.* Transcriptomic and genomic evidence for *Streptococcus agalactiae* adaptation to the bovine environment. **BMC Genomics.**, v. 14, p. 920, 2013.

RINAUDO, C.D. *et al.* Specific Involvement of Pilus Type 2a in Biofilm Formation in Group B *Streptococcus*. **PLoS ONE**, v. 5, n.2, p. 9216, 2010.

RIVERA, L. *et al.* Incidence and serotype distribution of invasive group B streptococcal disease in young infants: a multi-country observational study. **BMC Pediatr**, v.15, p.143, 2015.

ROCCHETTI, T.T. *et al.* Group B streptococci colonization in pregnant women: risk factors and evaluation of the vaginal. **Arch. Gynecol. Obstet.**, v. 283, p. 717–721, 2011.

ROH, K.K. *et al.* Effects of conventional or lower doses of hormone replacement therapy in postmenopausal women. **Arterioscler Tromb Vasc Biol.** v. 24, n. 8, p.1516-1521, 2004.

ROSINI, R. *et al.* Identification of novel genomic islands coding for antigenic *pilus*-like structures in *Streptococcus agalactiae*. **Mol Microbiol.**, v. 61, p. 126-141, 2006.

ROY, C.R. *et al.* Eukaryotic cells and microbial pathogens: a familiar couple take centre stage. **Nat. Cell. Biol.**, v.5, p.16-19, 2003.

ROSA-FRAILE, M. *et al.* Group B streptococcal haemolysin and pigment, a tale of twins. **FEMS microbiology reviews**, v. 38, n. 5, p. 932–946, 2014.

ROSINI, R. *et al.* Genomic analysis reveals the molecular basis for capsule loss in the group B *Streptococcus* population. **PLoS ONE**, v. 2, p. 25-39, 2015.

SAFADI, AL R. *et al.* Enhanced Expression of *lmb* Gene Encoding Laminin-Binding Protein in *Streptococcus agalactiae* Strains Harboring IS1548 in *scpBlmb* Intergenic Region. DeLeo FR, ed. **PLoS ONE**, v. 5, n. 5, p. e10794, 2010.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Mol Biol Evol.**, v. 4, n. 4, p. 406–425, 1987.

SANTOS, G.S. *et al.* The effects of interferon-gamma and transforming growth factor-beta on adherence and survival of group B *Streptococcus* type III strains in ECV304 cells. **Int J Mol Med.**, v. 11, p. 401-406, 2003.

SANTOS, G. S. *et al.* Signal transduction in human endothelial cells induced by their interaction with group B *Streptococci*. **Int J Mol Med.**, v. 15, p. 859-863, 2005.

SALZBERG, S.L. *et al.* GAGE: A critical evaluation of genome assemblies and assembly algorithms. **Genome Research**, v. 22, n. 3, p. 557-567, 2012.

SCHAEFER, R. *et al.* Characterization of Aujeszky's disease virus isolated from South Brazil in the last twenty years by restriction enzyme analysis. **Brazilian Journal of Microbiology**, v. 37, n. 3, p. 390-394, 2006.

SCHUCHAT, A. Epidemiology of group B Streptococcal disease in the United States: shifting paradigms. **Clin. Microbiol. Rev.**, v. 11, p. 497-513, 1998.

SCHUCHAT, A. Group B Streptococcus. **Lancet**, v. 353, p. 51-56, 1999.

SCHUBERT, A. *et al.* The fibrinogen receptor FbsA promotes adherence of Streptococcus agalactiae to human epithelial cells. **Infect. Immun.**, v. 72, p. 6197–6205, 2004.

SHABAYEK, S. *et al.* Serotype and surface protein gene distribution of colonizing group B streptococcus in women in Egypt. **Epidemiol. Infect.** v. 142, p. 208–210, 2014.

SHABAYEK, S.; SPELLERBERG, B. Group B Streptococcal Colonization, Molecular Characteristics, and Epidemiology. **Frontiers in Microbiology**, v.9, p. 437, 2018.

SHARMA, P. *et al.* Role of pilus proteins in adherence and invasion of Streptococcus agalactiae to the lung and cervical epithelial cells. **J Biol Chem.**, v. 288, p. 4023-4032, 2013.

SIGGE, A. *et al.* Heterogeneity of hemolysin expression during neonatal Streptococcus agalactiae sepsis. **J Clin Microbiol.**, v. 46, p. 807–809, 2008.

SILVA, T. T. **Mastite bovina e sua relação com a produção e composição do leite**. 37f. 2014. Trabalho de conclusão de curso (Bacharel em Zootecnia) – Universidade Federal de Goiás, Goiânia, 2014. Disponível em: [https://evz.ufg.br/up/66/o/15\\_MASTITE\\_BOVINA\\_E\\_SUA\\_RELACAO\\_COM\\_A\\_PRODUCAO\\_E\\_COMPOSICAO\\_DO\\_LEITE.pdf](https://evz.ufg.br/up/66/o/15_MASTITE_BOVINA_E_SUA_RELACAO_COM_A_PRODUCAO_E_COMPOSICAO_DO_LEITE.pdf). Acesso em: 20 maio 2019.

SILVA-FILHO, A.C. *et al.* Comparative Analysis of Genomic Island Prediction Tools. **Frontiers in Genetics.**, v. 9, p. 22-30, 2018.

SIMPSON, A.J. *et al.* The genome sequence of the plant pathogen Xylella fastidiosa. **Nature**, v. 406, p. 151-157, 2000.

SIX, A. *et al.* Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent group B streptococcus involved in binding to both fibrinogen and plasminogen. **Mol Microbiol.**, v. 97, p. 1209–1222, 2015.

SLOTVED, H.C. *et al.* Serotype IX, a proposed new Streptococcus agalactiae serotype. **J. Clin. Microbiol.**, v. 45, n. 9, p. 2929-2936, 2007.

SLOTVED, H. C.; HOFFMANN, S. Evaluation of procedures for typing of group B Streptococcus: a retrospective study. **PeerJ.**, v. 5, p. e3105, 2017.

SKOFF, T.H. *et al.* Increasing burden of invasive group B streptococcal disease in non-pregnant adults, 1990–2007. **Clin Infect Dis.**, v. 49, n. 1, p. 85–92, 2009.

SNEATH, P.H.A.; SOKAL, R.R. **Numerical taxonomy**. San Francisco: W.H. Freeman, 1973.

SOARES, S.C. *et al.* PIPS: pathogenicity island prediction software. **Plos one**, v. 7, p. e30848, 2012.

SOARES, T.C.G. *et al.* Prevalence of Group B Streptococcus serotypes III and V in pregnant women of Rio de Janeiro. **Brazil Brazilian Journal of Microbiology**, v. 44, n. 3, p. 869-872, 2013.

SOARES, S.C. *et al.* GIPSY: Genomic island prediction software. **Journal of Biotechnology**, v. 232, p. 2-11, 2015.

SORENSEN, U.B.S. *et al.* Emergence and global dissemination of host-specific Streptococcus agalactiae clones. **MBio**, v. 1, n. 3, p. e00178, 2010.

SOTERO-MARTINS, A. *et al.* A conservative region of the mercuric reductase gene (*merA*) as a molecular marker of bacterial mercury resistance. **Brazilian Journal of Microbiology**, v. 39, n. 2, p. 307–310, 2008.

SORIANO, M. *et al.* Streptococcus do Grupo B crosses Human Epithelial Cells by a Paracellular Route. **The Journal of Infectious Diseases**, v. 193, p. 241–250, 2006.

SPELLERBERG, B. *et al.* Lmb, a protein with similarities to the Lral adhesion family, mediates attachment of Streptococcus agalactiae to human laminin. **Infect Immun.**, v. 67, p. 871-878, 1999.

SPELLERBERG, B. *et al.* The *cyl* genes of Streptococcus agalactiae are involved in the production of pigment. **FEMS Microbiol Lett.**, v. 188, p. 125–128, 2000.

STEIN, L. From sequence to biology. **Genome annotation**. v.2, p. 23-27, 2001.

STONER, T.D. *et al.* Group B Streptococcal Infection and Activation of Human Astrocytes. **PLoS ONE**, v. 10, n. 6, p. e0128431, 2015.

SUHR, F. *et al.* Regulation of extracellular matrix compounds involved in angiogenic processes in short and long-track elite runners. **Journal of Medicine and Science in Sports**, v. 20, n. 3, p. 441- 448, 2010.

SUMPIO, B.E. *et al.* Cells in focus: endothelial cell. **The International Journal of Biochemistry & Cell Biology**, v. 32, n. 12, p. 1508–1512, 2002.

TAMURA, G.S. *et al.* Adherence of group B streptococci to cultured epithelial cells: roles of environmental factors and bacterial surface components. **Infect. Immun.**, v. 62, p. 2450-2458, 1994.

TAZI, A. *et al.* The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. **The Journal of Experimental Medicine**, v. 207, n. 11, p. 2313-2322, 2010.

TEATERO, S. *et al.* Clonal Complex 17 Group B Streptococcus strains causing invasive disease in neonates and adults originate from the same genetic pool. **Scientific Reports**, v. 6, p. 200-247, 2016.

TENENBAUM, T. *et al.* Adherence to and invasion of human brain microvascular endothelial cells are promoted by fibrinogen-binding protein FbsA of Streptococcus agalactiae. **Infect Immun.**, v. 73, p. 4404-4409, 2007.

TETTELIN, H. *et al.* Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae. **Proc. Natl. Acad. Sci.**, v. 99, p. 12391–12396, 2002.

TETTELIN, H. *et al.* Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial 'pan-genome'. **Proc Natl Acad Sci.**, v. 102, p. 13950–13955, 2005.

THUDI, M. *et al.* Current state-of-art of sequencing technologies for plant genomics research. **Briefings in Functional Genomics**, v. 11, n. 1, p. 3-11, 2012.

TIZARD, I.R. Imunidade Inata: Inflamação. In: TIZARD, I. R. **Imunologia Veterinária: uma introdução**. 6.ed. São Paulo: ROCCA, 2002. p. 39-45.

TONIOLO, C. *et al.* Streptococcus agalactiae Capsule Polymer Length and Attachment Is Determined by the Proteins CpsABCD. **Journal of Biological Chemistry**, v. 290, n. 15, p. 9521–9532, 2015.

TORRES-ESTAY, V. *et al.* Androgens modulate male-derived endothelial cell homeostasis using androgen receptor-dependent and receptor-independent mechanisms. **Angiogenesis**, v. 20, n. 1, p. 25–38, 2016.

TRACEY, K.J. The inflammatory reflex. **Nature**, n. 420, p. 853-859, 2002.

ULETT, G.C. *et al.*  $\beta$ -Hemolysin–Independent Induction of Apoptosis of Macrophages Infected with Serotype III Group B Streptococcus. **The Journal of Infectious Diseases**, v. 188, n. 7, p. 1049–1053, 2004.

VALBUENA, G.; WALKER, D.H. The endothelium as a target for infections. **Annu. Rev. Pathol. Mech. Dis**, v.1, p.171-198, 2006.

VAN WIJK, X.M. *et al.* Whole-Genome Sequencing of Invasion-Resistant Cells Identifies Laminin  $\alpha$ 2 as a Host Factor for Bacterial Invasion. **Mol bio.** v. 8, p. e02128-16, 2017.

VANDAMME, P. *et al.* Streptococcus *difficilis* is a nonhemolytic group B, type Ib Streptococcus. **Int J Sust Bacteriol.** v. 1, p. 81-5, 1997.

VASCONCELOS, A.T.R. *et al.* Swine and poultry pathogens: The complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. **J Bacteriol.**, v. 187, p. 5568-5577, 2005.

VECINI, J.F. *et al.* Importância do fibrinogênio plasmático na identificação de processos inflamatórios de cães. **Ciência Veterinária nos Trópicos**, v. 9, n. 1, p. 31-35, 2006.

VELTRI, D. *et al.* SimpleSynteny: a web-based tool for visualization of microsynteny across multiple species. **Nucleic acids research.**, v. 44, 41–45, 2016.

VELASCO, E. *et al.* Epidemiology of bloodstream infection at a cancer center. **São Paulo Méd J/Rev Paul Méd.**, v. 118, p. 131-8, 2000.

VERANI, J. R. *et al.* Prevention of perinatal group B streptococcal disease—revised guidelines from CDC 2010. **MMWR Recomm Rep.**, v. 59, n. 10, p. 1-36, 2010.

WANG, Z. *et al.* De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). **BMC Genomics**, v. 11, p. 726, 2010.

WANG, Z., *et al.* Pharmacologic targeting of sphingosine-1-phosphate receptor 1 improves the renal microcirculation during sepsis in the mouse. **J Pharmacol Exp Ther**, v.352, n. 1, p.61-66, 2015.

WANG, J. *et al.* Pathological analysis, detection of antigens, FasL expression analysis and leucocytes survival analysis in tilapia (*Oreochromis niloticus*) after infection with green fluorescent protein labeled *Streptococcus agalactiae*. **Fish Shellfish Immunol**, v.12, n. 1, p.e0168677, 2017.

WEISNER, A.M. *et al.* Characterization of group B streptococci recovered from infants with invasive disease in England and Wales. **Clin Infect Dis.**, v. 38, p. 1203-1208, 2004.

WEISS V.A. *et al.* Draft Genome Sequence of *Herbaspirillum lusitanum* P6-12, an Endophyte Isolated from Root Nodules of *Phaseolus vulgaris*. **J Bacteriol**, v. 194, n. 15, p. 4136–4137, 2012.

YAKES, F.M. *et al.* CM101 treatment overrides tumor-induced immunoprivilege leading to apoptosis. **Cancer Res.**, v. 60, p. 5740-5746, 2000.

YI, M. *et al.* Performance comparison of SNP detection tools with illumina exome sequencing data—an assessment using both family pedigree information and sample-matched SNP array data. **Nucl. Acids Res.**, v. 42, n. 12. p. 1-14. 2014.

YANCEY, M. K. *et al.* Perípartum infection associated with vaginal group B Streptococcal colonization. **Obstet. & Gynecol.**, v. 84, p. 816-819, 1994.

YAO, K. *et al.* Capsular gene typing of *Streptococcus agalactiae* compared to serotyping by latex agglutination. **J Clin Microbiol.** v., 51, n. 2, p. 503-507, 2013.

ZHAO, Z. *et al.* Genome-Wide Analysis of Tandem Repeats in Plants and Green Algae. **Genes, Genomes and Genetics**, v. 4, n. 1, p. 67–78, 2012.

ZHANG, W. *et al.*, A Practical Comparison of *De Novo* Genome Assembly Software Tools for Next-Generation Sequencing Technologies. **PLoS ONE**, v. 6, n. 3, p. e17915, 2011.

ZAFALON, L.F. *et al.* Staphylococcus aureus portadores de genes de toxinas isolados em amostras de diferentes fontes de transmissão durante a ordenha. São Paulo. **Revista do Instituto Adolfo Letz**, v. 68, n. 2, p. 269-277, 2009.

ZALEZNIK, D.F. *et al.* Invasive disease due to group B streptococcus in pregnant women and neonatos from diverse population groups. **Clin Infect Dis.** v. 30, p. 276-281, 2000.

ZUBAIR, S. *et al.* Genome Sequence of *Streptococcus agalactiae* Strain 09mas018883, Isolated from a Swedish Cow. **Genome Announcements**, v.1, n. 4, p. e00456-13, 2013.

**APÊNDICE A – Informações dos 109 genomas de *S. agalactiae* utilizadas nesse estudo**

Strain	Capsular type	Source	ST <sup>a</sup>	Host	Geographical origin	GenBank accession	Size (Mb)	G+C (%)	Genes	References
<b>A909</b>	Ia	neonatal sepsis	7	Human	USA <sup>b</sup>	CP000114.1	2.12	35.60	2207	Tettelin et al., 2005
<b>GBS85147</b>	Ia	oropharynx	103	Human	Brazil	CP010319.1	1.99	35.50	2040	De Aguiar et al., 2016
<b>GD201008-001</b>	Ia	meningoencephalitis	7	Tilapia	China	CP003810.1	2.06	35.60	2122	Liu et al., 2012
<b>HN016</b>	Ia	brain	7	Tilapia	China	CP011325.1	2.06	35.70	2120	Wang et al., 2015
<b>WC1535</b>	Ia	-	7	Tilapia	China	CP016501.1	2.21	35.80	2262	Zhang, 2016
<b>TFJ0901</b>	Ia	-	-	Tilapia	China	CP034315.1	2.08	35.70	2158	Mo et al., 2018
<b>138P</b>	Ib	kidney	261	Tilapia	USA	CP007482.1	1.84	35.50	1831	Pridgeon et al., 2014
<b>138spar</b>	Ib	kidney	261	Tilapia	USA	CP007565.1	1.84	35.50	1825	Pridgeon et al., 2014
<b>ND2-22</b>	Ib	brain	261	Tilapia	Israel	FO393392.1	1.84	35.50	1854	Rosinski-Chupin et al., 2013
<b>GX026</b>	Ib	brain	261	Tilapia	China	CP011328.1	1.84	35.50	1875	Chen, 2015
<b>SA1</b>	Ib	-	-	Tilapia	Brazil	CP019804.1	1.84	35.50	1911	Barony et al., 2017
<b>SA5</b>	Ib	-	-	Tilapia	Brazil	CP019805.1	1.84	35.50	1907	Barony et al., 2017
<b>SA9</b>	Ib	-	-	Tilapia	Brazil	CP019806.1	1.84	35.50	1910	Barony et al., 2017
<b>S13</b>	Ib	eye	552	Tilapia	Brazil	CP018623.1	1.84	35.40	1904	Pereira et al., 2016
<b>SA16</b>	Ib	brain	-	Tilapia	Brazil	CP019807.1	1.84	35.50	1903	Barony et al., 2017
<b>SA20</b>	Ib	-	552	Tilapia	Brazil	CP003919.2	1.84	35.50	1912	Pereira et al., 2016
<b>S25</b>	Ib	peritoneal cavity	552	Tilapia	Brazil	CP015976.1	1.84	35.50	1880	Pereira et al., 2016
<b>SA30</b>	Ib	brain	-	Tilapia	Brazil	CP019800.1	1.84	35.50	1909	Barony et al., 2017
<b>SA33</b>	Ib	brain	-	Tilapia	Brazil	CP019801.1	1.84	35.50	1906	Barony et al., 2017
<b>SA53</b>	Ib	brain	260	Tilapia	Brazil	CP019802.1	1.85	35.50	1903	Barony et al., 2017
<b>SA73</b>	Ib	brain	260	Tilapia	Brazil	CP019803.1	1.85	35.50	1914	Barony et al., 2017
<b>SA75</b>	Ib	brain	260	Tilapia	Brazil	CP019808.1	1.85	35.50	1910	Barony et al., 2017



SA79	Ib	kidney	-	Tilapia	Brazil	CP019809.1	1.84	35.50	1909	Barony et al., 2017
SA81	Ib	-	-	Tilapia	Brazil	CP019810.1	1.84	35.50	1899	Barony et al., 2017
SA85	Ib	brain	927	Tilapia	Brazil	CP019811.1	1.85	35.50	1915	Barony et al., 2017
SA95	Ib	brain	927	Tilapia	Brazil	CP019812.1	1.86	35.50	1910	Barony et al., 2017
SA97	Ib	brain	927	Tilapia	Brazil	CP019813.1	1.86	35.50	1923	Barony et al., 2017
SA102	Ib	brain	927	Tilapia	Brazil	CP019814.1	1.85	35.50	1913	Barony et al., 2017
SA132	Ib	brain	260	Tilapia	Brazil	CP019815.1	1.85	35.50	1911	Barony et al., 2017
SA136	Ib	brain	260	Tilapia	Brazil	CP019816.1	1.85	35.50	1907	Barony et al., 2017
SA159	Ib	brain	-	Tilapia	Brazil	CP019817.1	1.84	35.50	1905	Barony et al., 2017
SA184	Ib	brain	-	Tilapia	Brazil	CP019818.1	1.84	35.50	1910	Barony et al., 2017
SA191	Ib	brain	260	Tilapia	Brazil	CP019819.1	1.85	35.50	1915	Barony et al., 2017
SA195	Ib	brain	-	Tilapia	Brazil	CP019820.1	1.84	35.50	1912	Barony et al., 2017
SA201	Ib	brain	-	Tilapia	Brazil	CP019821.1	1.84	35.50	1909	Barony et al., 2017
SA209	Ib	brain	-	Tilapia	Brazil	CP019822.1	1.84	35.50	1907	Barony et al., 2017
SA212	Ib	brain	-	Tilapia	Brazil	CP019823.1	1.84	35.50	1906	Barony et al., 2017
SA218	Ib	brain	927	Tilapia	Brazil	CP019824.1	1.85	35.50	1912	Barony et al., 2017
SA220	Ib	brain	-	Tilapia	Brazil	CP019825.1	1.84	35.50	1905	Barony et al., 2017
SA245	Ib	brain	260	Tilapia	Brazil	CP019826.1	1.85	35.50	1907	Barony et al., 2017
SA256	Ib	brain	260	Tilapia	Brazil	CP019827.1	1.85	35.50	1909	Barony et al., 2017
SA289	Ib	brain	260	Tilapia	Brazil	CP019828.1	1.85	35.50	1908	Barony et al., 2017
SA330	Ib	brain	-	Tilapia	Brazil	CP019829.1	1.84	35.50	1910	Barony et al., 2017
SA333	Ib	brain	-	Tilapia	Brazil	CP019830.1	1.84	35.50	1908	Barony et al., 2017
SA341	Ib	brain	-	Tilapia	Brazil	CP019831.1	1.84	35.50	1915	Barony et al., 2017
SA343	Ib	brain	-	Tilapia	Brazil	CP019832.1	1.84	35.50	1904	Barony et al., 2017
SA346	Ib	brain	-	Tilapia	Brazil	CP019833.1	1.84	35.50	1903	Barony et al., 2017
SA374	Ib	brain	-	Tilapia	Brazil	CP019834.1	1.84	35.50	1911	Barony et al., 2017
SA375	Ib	brain	-	Tilapia	Brazil	CP019835.1	1.84	35.50	1914	Barony et al., 2017
SA623	Ib	brain	-	Tilapia	Brazil	CP019836.1	1.84	35.50	1912	Barony et al., 2017

SA627	Ib	brain	-	Tilapia	Brazil	CP019837.1	1.84	35.50	1913	Barony et al., 2017
Sag37	Ib	-	12	Human	China	CP019978.1	2.20	35.80	2292	Zhou et al., 2017
QMA0271	Ib	-	261	Fish	Australia	CP029632.1	1.80	35.30	1870	Ben Zakour et al., 2018
GBS-NY	II	blood	22	Human	USA	CP007570.1	2.24	35.90	2300	Srinivasan et al., 2014
GBS-NM	II	blood	22	Human	USA	CP007571.1	2.21	35.90	2275	Srinivasan et al., 2014
GBS6	II	-	22	Human	USA	CP007572.1	2.23	35.80	2306	Srinivasan et al., 2014
FDAARGOS_254	II	blood	22	Human	USA	CP020449.2	2.22	35.70	2299	Goldberg et al., 2018
SA111	II	milk	61	Bovine	Portugal	LT545678.1	2.28	35.60	2320	Institut Pasteur, 2016
GBS90356	III	neonatal CSF <sup>c</sup>	17	Human	Brazil	This study	2.03	35.20	2023	This study
COH-1	III	neonatal sepsis	17	Human	USA	HG939456.1	2.06	35.40	2123	Da Cunha et al., 2015
NGBS128	III	neonatal	17	Human	Canada	CP012480.1	2.07	35.70	2124	Teatero et al., 2016
CU_GBS_08	III	blood	283	Human	China	CP010874.1	2.08	35.40	2184	Zhou et al., 2015
CU_GBS_98	III	CSF	283	Human	China	CP010875.1	2.02	35.40	2129	Zhou et al., 2015
SG-M1	III	meningitis	283	Human	Singapore	CP012419.2	2.11	35.50	2219	Mehershahi et al., 2005
NEM316	III	neonatal sepsis	23	Human	France	AL732656.1	2.21	35.60	2273	Glaser et al, 2002
H002	III	pregnant vagina	736	Human	China	CP011329.1	2.14	35.70	2224	He et al., 2017
BM110	III	clinical isolate	17	Human	-	LT714196.1	2.17	35.50	2283	Institut Pasteur, 2017
FWL1402	III	-	739	Frog	China	CP016391.1	2.09	35.40	2177	Zhang, 2016
Sag158	III	-	19	Human	China	CP019979.1	2.10	35.70	2190	Zhou et al., 2017
CUGBS591	III	joint aspirate	12	Human	Hong Kong	CP021862.1	2.23	35.80	2330	Kalimuddin et al., 2017
SG-M50	III	blood	283	Human	Singapore	CP021865.1	2.12	35.50	2221	Kalimuddin et al., 2017
SG-M158	III	blood	283	Human	Singapore	CP021864.1	2.12	35.50	2221	Kalimuddin et al., 2017
SG-M29	III	blood	283	Human	Singapore	CP021866.1	2.12	35.50	2221	Kalimuddin et al., 2017
SG-M25	III	blood	19	Human	Singapore	CP021867.1	2.21	35.70	2319	Kalimuddin et al., 2017
SG-M6	III	blood	17	Human	Singapore	CP021869.1	2.11	35.60	2188	Kalimuddin et al., 2017
SG-M4	III	blood	23	Human	Singapore	CP021870.1	2.07	35.50	2133	Kalimuddin et al., 2017
SG-M163	III	blood	283	Human	Singapore	CP021863.1	2.11	35.50	2221	Kalimuddin et al., 2017
C001	III	milk	-	Bovine	China	CP008813.1	2.12	35.60	2189	Chan, 2014

<b>874391</b>	III	vagina	17	Human	Japan	CP022537.1	2.15	35.50	2224	Sullivan et al., 2017
<b>SGEHI2015-107</b>	III	muscles	283	Fish	Singapore	CP025027.1	2.02	35.40	2118	Kalimuddin et al., 2017
<b>SGEHI2015-95</b>	III	muscles	283	Fish	Singapore	CP025028.1	2.12	35.50	2245	Kalimuddin et al., 2017
<b>SGEHI2015-25</b>	III	muscles	283	Fish	Singapore	CP025029.1	2.05	35.40	2171	Kalimuddin et al., 2017
<b>SGEHI2015-113</b>	III	muscles	283	Fish	Singapore	CP025026.1	2.04	35.40	2154	Kalimuddin et al., 2017
<b>HU-GS5823</b>	III	blood	335	Human	Japan	AP018935.1	2.23	35.60	2317	Nagoaka et al., 2018
<b>NGBS572</b>	IV	-	452	Human	Canada	CP007632.1	2.06	35.50	2127	Teatero et al., 2015
<b>NGBS061</b>	IV	-	459	Human	Canada	CP007631.2	2.22	35.50	2314	Teatero et al., 2015
<b>SG-M8</b>	IV	blood	1	Human	Singapore	CP021868.1	2.17	35.60	2237	Kalimuddin et al., 2017
<b>NGBS357</b>	V	blood	1	Human	Canada	CP012503.1	2.17	35.60	2248	Teatero et al., 2016
<b>SS1</b>	V	blood	1	Human	USA	CP010867.1	2.09	35.50	2173	Flores et al., 2015
<b>2603 V/R</b>	V	clinical isolate	110	Human	USA	AE009948.1	2.16	35.60	2279	Tettelin et al., 2002
<b>CNCTC 10/84</b>	V	clinical isolate	26	Human	USA	CP006910.1	2.01	35.40	2094	Hooven et al., 2014
<b>09mas018883</b>	V	mastitis	1	Bovine	Sweden	HF952104.1	2.14	35.50	2225	Zubair et al., 2013
<b>ILRI005</b>	V	abscess	609	Camel	Kenya	HF952105.1	2.11	35.40	2214	Zubair et al., 2013
<b>GBS ST-1</b>	V	oral cavity	1	Canine	USA	CP013202.1	2.17	35.40	2261	Harden et al., 2018
<b>GBS-M002</b>	VI	vagina	1	Human	Taiwan	CP013908.1	2.09	35.60	2169	He et al., 2017
<b>ILRI112</b>	VI	abscess	617	Camel	Kenya	HF952106.1	2.03	35.30	2173	Zubair et al., 2013
<b>B509</b>	-	vagina	-	Human	China	CP021769.1	2.06	35.50	2158	Wu et al., 2017
<b>B508</b>	-	vagina	-	Human	China	CP021770.1	2.20	35.60	2282	Wu et al., 2017
<b>B105</b>	-	neonatal blood	-	Human	China	CP021773.1	2.27	35.70	2363	Wu et al., 2017
<b>B507</b>	-	vagina	-	Human	China	CP021771.1	2.08	35.40	2199	Wu et al., 2017
<b>B111</b>	-	neonatal blood	-	Human	China	CP021772.1	2.15	35.40	2252	Wu et al., 2017
<b>S73</b>	-	brain	-	Tilapia	Brazil	CP030845.1	2.06	35.40	2177	Pereira et al., 2016
<b>Sag27</b>	-	perianal region	-	Human	China	CP031556.1	2.21	35.70	2333	Zhou et al., 2018
<b>FDAARGOS_512</b>	-	endotracheal aspiration	-	Human	USA	CP033822.1	2.13	35.60	2233	Goldberg et al., 2018
<b>NCTC11930</b>	-	-	-	-	Denmark	LS483342.1	2.06	35.40	2164	Doyle, 2018
<b>NCTC8187</b>	-	-	-	-	United Kingdom	LS483387.1	2.05	35.50	2147	Doyle, 2018

<b>NCTC8184</b>	-	-	-	-	United Kingdom	LR134265.1	2.27	35.80	2366	Doyle, 2018
<b>NCTC13949</b>	-	-	-	-	United Kingdom	LR134512.1	2.03	35.60	2111	Welcome Trust Sanger Institute, 2018
<b>NCTC13947</b>	-	-	-	-	United Kingdom	LR134520.1	2.10	35.60	2197	Welcome Trust Sanger Institute, 2018

<sup>a</sup> sequence type; <sup>b</sup> United States of America; <sup>c</sup> cerebrospinal fluid.

**APÊNDICE B** – Lista das CDSs detectadas dentro das PAIs do genoma da *S. agalactiae* GBS90356. Ilhas de patogenicidade são abreviadas com SagPAI

Pathogenicity island GIPSY	Size (bases)	Gene name	Product name	Number Secreted proteins	Location (pb)	Pathogenicity island IslandViewer4
SagPAI_1	12023			0	7632..19654	
	450	<i>cpsF</i>	Polysaccharide biosynthesis protein CpsF		7656..8105	
	474	<i>cpsG</i>	Glycosyl transferase CpsG		8105..8578	
	1146	<i>cpsH</i>	Capsular polysaccharide repeating-unit polymerase CpsH		8575..9720	
	969	<i>cpsI</i>	Capsular polysaccharide biosynthesis protein CpsI		9717..10685	
	948	<i>cpsJ</i>	Polysaccharide biosynthesis glycosyl transferase CpsJ		10719..11666	
	957	<i>cpsK</i>	Capsular polysaccharide biosynthesis protein CpsK(V)		11750..12706	
	1401	<i>cpsL</i>	Capsular polysaccharide repeat unit transporter CpsL		12703..14103	
	1026	<i>neuB</i>	N-acetylneuraminase synthase		14103..15128	
	1155	<i>neuC</i>	UDP-N-acetylglucosamine 2-epimerase		15205..16359	
	630	<i>neuD</i>	Sialic acid biosynthesis protein NeuD, O- acetyltransferase		16356..16985	

	1242	<i>neuA</i>	N-Acetylneuraminat cytidyltransferase		16996..18237	
	486	–	Hypothetical protein		18351..18836	
	654	<i>ung</i>	Uracil-DNA glycosylase, family 1		18935..19588	
SagPAI_2	30193			3	193219..223411	
	753	–	ABC transporter ATP-binding protein		193221..193973 c	
	963	–	nisin-resistance protein, putative		193999..194961 c	
	180	–	Hypothetical protein		195127..195306	
	144	–	Hypothetical protein		195471..195614	
	849	–	Methionine ABC transporter substrate-binding protein		195646..196494 c	
	744	–	Protein export cytoplasm protein SecA ATPase RNA helicase		196627..197370 c	
	699	–	Transcriptional regulator in cluster with beta- lactamase, GntR family		198896..199594 c	
	1563	<i>guaA</i>	GMP synthase [glutamine- hydrolyzing]		199807..201369	
	1194	–	Integrase, superantigen-encoding pathogenicity islands SaPI		201437..202630 c	
	201	–	Hypothetical protein		202657..202857 c	
	231	–	Hypothetical protein		203353..203583 c	
	423	–	Hypothetical protein		203580..204002 c	
	354	–	Hypothetical protein		204534..204887	
	1920	<i>tetM</i>	Tetracycline resistance protein TetM		205231..207150 c	
	276	–	Hypothetical protein		207527..207802 c	
	597	–	Hypothetical protein		207832..208428 c	
	1023	–	Lipoprotein, NLP/P60 family		208445..209467 c	

Islandviewer4  
SagPAI\_2

	2028	-	membrane protein, putative		209464..211491 c	
	2454	-	Hypothetical protein		211488..213941 c	
	396	-	Hypothetical protein		213925..214320 c	
	1026	-	Hypothetical protein		214391..215416 c	
	501	-	Antirestriction protein		215488..215988 c	
	762	-	Hypothetical protein		216050..216811 c	
	222	-	Conjugation related protein		216871..217092 c	
	273	-	Hypothetical protein		217089..217361 c	
	1185	-	Tn916, transcriptional regulator, putative		217358..218542 c	
	1098	-	Hypothetical protein		218724..219821 c	
	159	-	Hypothetical protein		219938..220096	
	774	-	Hypothetical protein		220149..220922 c	
	378	-	Hypothetical protein		220932..221309 c	
	315	-	Hypothetical protein		221330..221644 c	
	1128	-	Na <sup>+</sup> /H <sup>+</sup> antiporter		222221..223348	
SagPAI_3	6179			0	282273..288451	
	270	-	Retron-type RNA-directed DNA polymerase		282422..282691 c	
	894	-	Retron-type RNA-directed DNA polymerase		282805..283698 c	
	1356	-	Predicted arginine uptake transporter		284415..285770	
	135	<i>rpmH</i>	LSU ribosomal protein L34p		285954..286088	
	675	<i>dltR</i>	Regulation of D-alanyl- lipoteichoic acid biosynthesis, DltR		286444..287118	

	1188	<i>dltS</i>	Regulation of D-alanyl-lipoteichoic acid biosynthesis, sensor histidine kinase		287118..288305	
SagPAI_4	56006			1	681956..737961	
	615	–	Nitroreductase family protein		682089..682703	
	981	–	Hypothetical protein		682875..683855 c	
	198	–	Hypothetical protein		684038..684235 c	
	639	–	4-hydroxy-2-oxoglutarate aldolase		685433..686071 c	
	1008	–	2-dehydro-3-deoxygluconate kinase		686083..687090 c	
	639	–	predicted 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase		687116..687754 c	
	813	–	2-deoxy-D-gluconate 3-dehydrogenase		687771..688583 c	
	435	–	PTS system, hyaluronate-oligosaccharide-specific IIA component		688861..689295	
	1197	–	Unsaturated glucuronyl hydrolase		689298..690494	
	495	–	PTS system, hyaluronate-oligosaccharide-specific IIB component		690549..691043	
	867	–	PTS system, hyaluronate-oligosaccharide-specific IIC component		691079..691945	
	816	–	PTS system, N-acetylgalactosamine-specific IID component		691932..692747	
	1905	–	Heparinase II/III-like		692827..694731	



1002	<i>regR</i>	Transcriptional regulator RegR, repressor of hyaluronate and KDG utilization	694811..695812
615	<i>def</i>	Peptide deformylase	695884..696498 c
645	–	Predicted N-ribosylNicotinamide CRP-like regulator	696564..697208 c
1164	–	membrane protein, putative	697320..698483
984	–	oxidoreductase, Gfo/Idh/MocA family	698529..699512 c
1896	–	Neutral endopeptidase O	699660..701555
954	–	Muramoyltetrapeptide carboxypeptidase	701605..702558 c
237	–	lysine rich protein	702725..702961
2070	–	Na <sup>+</sup> /H <sup>+</sup> antiporter	702963..705032
1068	–	Integrase	705176..706243 c
405	–	Phage terminase	706367..706771 c
747	–	Hypothetical protein	706793..707539 c
381	–	Phage transcriptional regulator, Cro/CI family	707548..707928 c
165	–	Phage protein	708216..708380
471	–	Hypothetical protein	708363..708833 c
213	–	Phage protein	708894..709106
342	–	Hypothetical protein	709106..709447
531	–	HNH homing endonuclease	709494..710024
285	–	Phage protein	710024..710308
219	–	Phage protein	710332..710550
1332	<i>dnaC- 1</i>	Replicative DNA helicase	710559..711890


Islandviewer4  
SagPAI\_4

603	-	putative protein	711892..712494
747	-	Phage replication initiation protein	712472..713218
822	-	Helicase loader DnaI	713219..714040
249	-	Hypothetical protein	714040..714288
495	-	Phage DNA binding protein	714281..714775
330	-	Phage protein	714772..715101
420	-	Hypothetical protein	715101..715520
207	-	Hypothetical protein	715517..715723
417	<i>ssb4</i>	Single-stranded DNA-binding protein	715713..716129
261	-	Phage protein	716139..716399
348	-	prophage LambdaSa2, transcriptional regulator, Cro/CI family	716396..716743
465	-	Hypothetical protein	716740..717204
543	-	Mobile element protein	717314..717856
288	-	Hypothetical protein	718146..718433
480	-	Hypothetical protein	718886..719365
1713	-	Phage terminase, large subunit	719358..721070
927	-	Hypothetical protein	721295..722221
543	-	prophage LambdaSa2, protease, putative	722271..722813
1263	-	Phage capsid and scaffold	722824..724086
336	-	Glycerate kinase	724107..724442
306	-	Hypothetical protein	724439..724744
348	-	Hypothetical protein	724744..725091
345	-	Hypothetical protein	725078..725422
684	-	Hypothetical protein	725437..726120

	477	–	Hypothetical protein		726120..726596	
	129	–	Hypothetical protein		726635..726763	
	2736	–	Phage tail length tape-measure protein		726782..729517	
	723	–	Hypothetical protein		729514..730236	
	3675	–	Phage hyaluronidase		730237..733911	
	231	–	Hypothetical protein		733928..734158	
	339	–	Hypothetical protein		734161..734499	
	411	–	Hypothetical protein		734534..734944	
	330	–	Phage holin		734947..735276	
	1407	–	Phage lysin, N-acetylmuramoyl-L-alanine amidase		735280..736686	
	186	–	Hypothetical protein		736893..737078	
	405	–	Phage-related protein		737139..737543	
SagPAI_5	124906			12	907232..1032137	
	516	<i>lacB</i>	Galactose-6-phosphate isomerase, LacB subunit		912816..913331 c	
	348	<i>lacA</i>	Galactose-6-phosphate isomerase, LacA subunit		913429..913776 c	
	2451	–	Sialidase		914025..916475 c	
	378	–	PTS system, galactose-specific IIC component		916826..917203 c	
	306	–	PTS system, galactose-specific IIB component		917243..917548 c	
	465	–	PTS system, galactose-specific IIA component		917550..918014 c	
	777	<i>lacR</i>	Lactose phosphotransferase system repressor		918232..919008	

2583	–	Streptococcal histidine triad protein	919289..921871 c
924	<i>lmb</i>	Laminin-binding surface protein	921898..922821 c
444	<i>dtd</i>	D-tyrosyl-tRNA(Tyr) deacylase	923076..923519 c
2217	–	GTP pyrophosphokinase	923529..925745 c
2403	<i>cpdB</i>	2',3'-cyclic-nucleotide 2'-phosphodiesterase	926005..928407
456	<i>nrdI-2</i>	Ribonucleotide reduction protein NrdI	928564..929019
1038	–	Deblocking aminopeptidase	929030..930067
498	–	Hypothetical protein	930067..930564
120	–	Hypothetical protein	930621..930740
1038	–	Ferric iron ABC transporter, iron-binding protein	930786..931823 c
774	–	Two-component response regulator, associated with ferric iron transporter	931835..932608 c
1650	–	Two-component sensor kinase, associated with ferric iron transporter	932612..934261 c
828	–	PTS system, mannose-specific IID component	934399..935226 c
810	–	PTS system, mannose-specific IIC component	935223..936032 c
504	–	PTS system, mannose/fructose family IIB component	936036..936539 c
426	–	PTS system, IIA component, putative	936557..936982 c
1035	–	membrane protein, putative	937229..938263 c

183	-	Hypothetical protein	938315..938497
1155	-	Putative transmembrane protein	938492..939646 c
849	-	ABC transporter ATP-binding protein	939652..940500 c
159	-	Hypothetical protein	940569..940727 c
183	-	competence-stimulating peptide ABC transporter ATP-binding protein ComA	941031..941213
120	-	Hypothetical protein	941170..941289
1335	<i>rgfC</i>	Histidine kinase of the competence regulon ComD	941284..942618 c
753	-	Response regulator of the competence regulon ComE	942615..943367 c
816	<i>rgfB</i>	Exodeoxyribonuclease III	943583..944398 c
2184	-	PTS system, maltose and glucose-specific IIC component / IIB component / IIA component	944452..946635 c
1656	-	Phosphate regulon sensor protein PhoR (SphS)	946790..948445 c
678	-	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	948438..949115 c
657	-	Phosphate transport system regulatory protein PhoU	949115..949771 c
750	<i>pstB3</i>	Phosphate transport ATP-binding protein PstB	949768..950517 c
879	-	Phosphate transport system permease protein PstA	950510..951388 c

846	-	Phosphate transport system permease protein PstC	951390..952235 c	
882	<i>pstS2</i>	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS	952250..953131 c	
285	-	Thymidylate kinase	953332..953616 c	
306	-	Thymidylate kinase	953613..953918 c	
741	-	Ribosomal RNA small subunit methyltransferase E	953915..954655 c	
954	-	Ribosomal protein L11 methyltransferase	954655..955608 c	
309	-	Hypothetical protein	955605..955913 c	
717	-	Transcriptional regulator, MerR family	956187..956903	
471	-	acetyltransferase, GNAT family	956942..957412 c	
459	-	MutT/nudix family protein; 7,8-dihydro-8-oxoguanine-triphosphatase	957384..957842 c	
144	-	Hypothetical protein	957829..957972 c	
471	-	Hypothetical protein	957978..958448 c	
477	-	Protein export cytoplasm protein SecA ATPase RNA helicase	958790..959266	
1269	-	Chromosome segregation helicase	959259..960527	
1122	-	Integrase	960982..962103 c	
327	-	Hypothetical protein	962153..962479 c	
633	-	Hypothetical protein	962672..963304 c	
792	-	cell wall surface anchor family protein, putative	963385..964176 c	

Islandviewer4  
SagPAI\_5

549	–	Hypothetical protein	964193..964741 c
1374	–	Glycerate kinase	964741..966114 c
2001	–	membrane protein, putative	966610..968610 c
2502	–	conjugal transfer protein, putative	968626..971127 c
411	–	Hypothetical protein	971148..971558 c
267	–	Hypothetical protein	971569..971835 c
954	–	Hypothetical protein	971856..972809 c
255	–	YoeB toxin protein	972862..973116 c
267	–	YefM protein	973109..973375 c
474	–	Hypothetical protein	973520..973993 c
483	–	Hypothetical protein	974001..974483 c
273	–	Hypothetical protein	974536..974808 c
570	–	Hypothetical protein	974943..975512 c
1350	–	Transporter	975515..976864 c
300	–	Transcriptional regulator, XRE family	977061..977360 c
378	–	Hypothetical protein	977344..977721 c
1233	–	Transcriptional regulator, Cro/CI family	977737..978969 c
1719	–	FtsK/SpoIIIE family protein	979257..980975 c
462	–	putative transfer protein	980995..981456 c
297	–	putative transfer protein	981478..981774 c
2592	–	Hypothetical protein	981834..984425 c
2262	–	N-acetylmuramoyl-L-alanine amidase	984514..986775 c
264	–	Hypothetical protein	986791..987054 c
462	–	Hypothetical protein	987074..987535 c
312	–	Hypothetical protein	987520..987831 c

432	-	Transcriptional regulator, Cro/CI family	987824..988255 c
315	-	Transcriptional regulator, XRE family	989386..989700
153	-	Hypothetical protein	989713..989865 c
339	-	Transcriptional regulator, XRE family	989900..990238
423	-	Hypothetical protein	990245..990667
369	-	<i>unknown</i>	990742..991110 c
339	-	Hypothetical protein	991114..991452 c
294	-	DNA-damage-inducible protein J, putative	991449..991742 c
759	-	CAMP factor	992276..993034 c
1308	-	Mobile element protein	993331..994638 c
1641	<i>merA-2</i>	Mercuric ion reductase	994798..996438 c
393	<i>merR-2</i>	Mercuric resistance operon regulatory protein	996467..996859 c
1569	-	Manganese transport protein MntH	997565..999133
723	-	ABC-2 family transporter protein YydJ	999701..1000423 c
618	-	ABC transporter	1000427..1001044 c
159	-	Hypothetical protein	1001102..1001260 c
750	-	Membrane-embedded metalloprotease YydH	1001320..1002069 c
948	-	Hypothetical protein	1002029..1002976 c
306	-	<i>Unknown</i>	1003548..1003853 c
402	-	Hypothetical protein	1003837..1004238 c



1218	–	Hypothetical protein	1004226..1005443 c
387	–	<i>unknown</i>	1005480..1005866 c
510	–	Hypothetical protein	1006037..1006546 c
333	–	<i>unknown</i>	1006548..1006880 c
297	–	<i>unknown</i>	1007140..1007436 c
855	–	Aminoglycoside 6-adenylyltransferase	1007699..1008553 c
210	–	Hypothetical protein	1008732..1008941 c
609	–	<i>unknown</i>	1009566..1010174 c
336	–	Transcriptional regulator, PadR family	1010161..1010496 c
489	–	acetyltransferase, GNAT family	1011018..1011506 c
762	–	membrane protein, putative	1011675..1012436 c
903	–	ABC transporter, ATP-binding protein	1012433..1013335 c
207	–	Hypothetical protein	1013332..1013538 c
1077	–	transcriptional regulator, Cro/CI family	1013540..1014616 c
156	–	Hypothetical protein	1014640..1014795 c
1782	–	Excinuclease ABC subunit C	1015019..1016800 c
768	–	Diacylglycerol hydrolase like	1016959..1017726
1341	–	Multi antimicrobial extrusion protein (Na <sup>+</sup> )/drug antiporter), MATE family of MDR efflux pumps	1017785..1019125 c
411	–	PhnB protein; putative DNA binding 3- demethylubiquinone-9 3-methyltransferase domain protein	1019225..1019635 c

	498	-	Hypothetical protein		1019645..1020142 c	
	597	-	putative superinfection immunity protein		1020364..1020960	
	291	-	Mobile element protein		1021093..1021383	
	636	-	Mobile element protein		1021563..1022198	
	1236	-	Hypothetical protein		1022408..1023643	
	1203	-	Integrase		1023823..1025025 c	
	258	-	Hypothetical protein		1025028..1025285 c	
	870	-	Hypothetical protein		1025305..1026174c	
	1161	-	FtsK-like DNA segregation ATPase, YDCQ B.subtilis ortholog		1026276..1027436 c	
	426	-	Hypothetical protein		1027757..1028182 c	
	162	-	Hypothetical protein		1028654..1028815 c	
	1161	-	Hypothetical protein		1028859..1030019	
	1197	-	serine/threonine protein kinase		1030073..1031269 c	
	810	-	type II restriction enzyme eco47ii		1031327..1032136 c	
SagPAI_6	7103			0	1034362..1041464	
	2109	<i>clpL</i>	putative ATP-dependent Clp proteinase (ATP- binding subunit)		1034795..1036903	
	945	-	Homocysteine S-methyltransferase		1036996..1037940 c	
	1377	-	S-methylmethionine permease		1037945..1039321 c	
	651	-	Hypothetical protein		1039451..1040101 c	
	411	-	<i>unknown</i>		1040259..1040669 c	
	120	-	Hypothetical protein		1040679..1040798 c	
	549	-	Transcriptional regulator, TetR family		1040854..1041402	
SagPAI_7	13935			0	13.184.281.332.362	
	708	-	Hypothetical protein		1318706..1319413 c	

	312	-	Transcriptional regulator, PadR family		1319604..1319915	
	684	-	Hypothetical protein		1319936..1320619	
	717	-	Methionine ABC transporter ATP-binding protein		1320883..1321599	
	1098	-	ABC transporter permease protein		1321602..1322699	
	609	-	Transcriptional regulator, TetR family		1322723..1323331	
	162	-	Phosphonate ABC transporter phosphate-binding periplasmic component		1323466..1323627 c	
	492	-	RNA polymerase sigma factor, ECF subfamily		1323911..1324402	
	948	-	Hypothetical protein		1324399..1325346	
	708	-	Protein export cytoplasm protein SecA ATPase RNA helicase		1325606..1326313 c	
	231	-	acetyltransferase, GNAT family		1326454..1326684 c	
	561	-	acetyltransferase, GNAT family		1327012..1327572 c	
	351	-	Hypothetical protein		1328020..1328370 c	
	273	-	Hypothetical protein		1328380..1328652 c	
	1635	-	Hypothetical protein		1328671..1330305 c	
	393	-	Phosphonate ABC transporter phosphate-binding periplasmic component		1330369..1330761 c	
SagPAI_8	35230			2	1414023..1449252	
	312	-	Hypothetical protein		1414291..1414602 c	
	255	-	Aldose 1-epimerase		1414711..1414965 c	

216	-	Ubiquinone/menaquinone biosynthesis methyltransferase UbiE	1415613..1415828	
249	-	Uncharacterized conserved protein, contains double-stranded beta-helix domain	1415862..1416110	
207	-	Integrase	1416387..1416593 c	
252	-	Hypothetical protein	1416782..1417033	
372	-	<i>unknown</i>	1417338..1417709 c	
189	-	Hypothetical protein	1417862..1418050 c	
294	-	DNA-damage-inducible protein J, putative	1418047..1418340 c	
207	-	Surface protein Rib	1419052..1419258 c	
1077	-	Surface protein Rib	1419236..1420312 c	
1194	-	Transcriptional regulator, AraC family	1420560..1421753	Islandviewer4
525	-	Transcriptional regulator SczA, TetR family	1421948..1422472 c	SagPAI_8
864	-	Cobalt-zinc-cadmium resistance protein CzcD	1422604..1423467	
855	-	2,5-diketo-D-gluconic acid reductase	1423552..1424406	
456	<i>adh2</i>	Alcohol dehydrogenase	1424516..1424971	
573	<i>adh</i>	Alcohol dehydrogenase	1424971..1425543	
399	<i>adhR</i>	transcriptional regulator, MerR family	1425550..1425948	
396	-	Transcriptional regulator	1425958..1426353	
162	-	4-carboxymuconolactone decarboxylase	1426381..1426542	

285	–	Hypothetical protein	1426785..1427069
492	–	Transcriptional regulator, AraC family	1427308..1427799
120	–	Hypothetical protein	1427882..1428001
1146	–	Osmotically activated L-carnitine/choline ABC transporter, ATP-binding protein OpuCA	1428100..1429245
636	–	smotically activated L-carnitine/choline ABC transporter, permease protein OpuCB	1429245..1429880
927	–	Osmotically activated L-carnitine/choline ABC transporter, substrate-binding protein OpuCC	1429883..1430809
651	–	Osmotically activated L-carnitine/choline ABC transporter, permease protein OpuCD	1430810..1431460
1182	–	MFS general substrate transporter	1431522..1432703 c
633	–	Positive transcriptional regulator, MutR family	1432776..1433408 c
222	–	Positive transcriptional regulator, MutR family	1433408..1433629 c
126	–	Hypothetical protein	1433629..1433754 c
1443	–	ATP-dependent DNA helicase recG	1434492..1435934 c
3756	–	DNA-binding protein	1435960..1439715 c
4371	–	Type II restriction endonuclease	1439875..1444245 c
690	–	Methylase	1444250..1444939 c
195	–	Hypothetical protein	1445251..1445445 c
333	–	Tn5252, Orf 9 protein	1445426..1445758 c

	300	-	Tn5252, Orf 9 protein		1445766..1446065 c	
	840	-	Phage replication initiation		1446118..1446957 c	
	282	-	Hypothetical protein		1446966..1447247 c	
	519	-	Hypothetical protein		1447395..1447913	
	1155	-	Integrase		1447965..1449119	
	114	-	Hypothetical protein		1449124..1449237	
SagPAI_9	11994			0	1856445..1868438	
	405	-	Transposase		1857473..1857877	
	201	-	Mobile element protein		1858039..1858239	
	258	-	Mobile element protein		1858264..1858521	
	162	-	<i>unknown</i>		1859029..1859190 c	
	1278	<i>vex1</i>	ABC transporter membrane-spanning permease, Pep export, Vex1		1859932..1861209	
	657	<i>vex2</i>	ABC transporter, ATP-binding protein Vex2		1861219..1861875	
	1377	<i>vex3</i>	ABC transporter membrane-spanning permease, Pep export, Vex3		1861875..1863251	
	654	<i>vncR</i>	Two-component response regulator VncR		1863348..1864001	
	1320	<i>vncS</i>	Sensor histidine kinase VncS		1863998..1865317	
	612	-	Mobile element protein		1865434..1866045 c	
	201	-	Hypothetical protein		1866259..1866459	
	189	-	Paratox		1866501..1866689	
	1206	-	Cell division protein FtsW		1867114..1868319	
SagPAI_10	12495			2	1907736..1920230	
	420	-	Hypothetical protein		1907936..1908355 c	

Islandviewer4  
SagPAI\_9

	552	-	Hypothetical protein	1908575..1909126 c	
	345	-	Hypothetical protein	1909916..1910260 c	
	375	-	Phosphonate ABC transporter phosphate-binding periplasmic component	1910253..1910627 c	
	204	-	Hypothetical protein	1910844..1911047 c	
	432	-	Hypothetical protein	1911117..1911548 c	
	615	-	PAP2 family protein	1912147..1912761 c	
	318	-	MazG related NTP pyrophosphohydrolase	1912975..1913292	
	561	-	Lysine decarboxylase family	1913289..1913849	
	864	-	protease, putative	1914154..1915017	
	303	-	Rhodanese-like domain protein	1915174..1915476 c	
	768	<i>cfb</i>	CAMP factor	1915649..1916416 c	
	432	-	DNA topology modulation protein	1917186..1917725	
	1086	-	Glycerol dehydrogenase	1917926..1919011	
	708	-	Acid-resistant locus arl7 (Fragment)	1919176..1919883 c	
	3453	<i>scpB</i>	c5a peptidase precursor	1937763..1942839 c	

ANEXO – Formato do 1º artigo submetido

## Comparative genomic analysis and identification of pathogenicity islands of hypervirulent ST-17 *Streptococcus agalactiae* Brazilian strain

Pamella Silva Lannes-Costa<sup>1</sup>, Rafael Azevedo Baraúna<sup>2</sup>, Juliana Nunes Ramos<sup>3</sup>, João Flavio Carneiro Veras<sup>3</sup>, Marcos Vinícius Reis Conceição<sup>2</sup>, Verônica Viana Vieira<sup>4</sup>, Ana Luiza de Mattos-Guaraldi<sup>5</sup>, Rommel Thiago Juca Ramos<sup>2</sup>, Kelly S. Doran<sup>6</sup>, Artur Silva<sup>2</sup>, Prescilla Emy Nagao<sup>1\*</sup>.

<sup>1</sup>Laboratory of Molecular Biology and Physiology of Streptococci, Institute of Biology Roberto Alcantara Gomes, State University of Rio de Janeiro (UERJ), Rio de Janeiro, RJ, Brazil.

<sup>2</sup>Center of Genomics and Systems Biology, Institute of Biological Sciences, Federal University of Pará (UFPA), Belém, PA, Brazil.

<sup>3</sup>National Institute for Quality Control in Health (INCQS), Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, RJ, Brazil.

<sup>4</sup>Interdisciplinary Laboratory of Medical Research, Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, RJ, Brazil.

<sup>5</sup>Faculty of Medical Sciences, State University of Rio de Janeiro (UERJ), Rio de Janeiro, RJ, Brazil.

<sup>6</sup>Department of Biology, San Diego State University, San Diego, CA, Department of Immunology and Microbiology, University of Colorado School of Medicine, Aurora, CO, United States of America.

### \* Correspondence:

Prescilla Emy Nagao

[pnagao@uerj.br](mailto:pnagao@uerj.br) / [pnagao@uol.com.br](mailto:pnagao@uol.com.br)

### Abstract

**Background:** *Streptococcus agalactiae* are important pathogenic bacteria that cause severe infections in humans, especially neonates. The mechanism by which ST-17 causes invasive infections than other STs is not well understood. In this study, we sequenced the first genome of a *S. agalactiae* ST-17 strain isolated in Brazil using the Illumina HiSeq 2500 technology. *S. agalactiae* GBS90356 ST-17 belongs to the capsular type III and was isolated from a neonatal with a fatal case of meningitis.

**Results:** The genome presented a size of 2.03 Mbp and a G+C content of 35.2%. *S. agalactiae* has 706 genes in its core genome and an open pan-genome with a size of 5.020 genes, suggesting a high genomic plasticity of *S. agalactiae*. GIPSY software was used to identify 10 Pathogenicity islands (PAIs) which corresponded to 15% of the genome size. IslandViewer4 corroborated the prediction of five of these PAIs. These pathogenicity islands



showed important virulence factors genes for *S. agalactiae* e.g. *neu*, *cps*, *dlt*, *fbs*, *cfb*, *lmb*. SignalP detected 20 proteins with signal peptides among the 352 proteins found in PAIs, which 60% were located in the SagPAI\_5. SagPAI\_2 and 5 were partially detected in ST-17 strains studied. Moreover, we identified unique genes that potentially belong to ST-17 strains, which can be related with virulence during host-pathogen interactions.

**Conclusions:** We provide evidence of high genomic plasticity in *S. agalactiae* although this is not usual in pathogenic bacteria with small genomes. The detected PAIs comprise a large portion of the genome, indicating that the majority of the genes are involved in *S. agalactiae* pathogenicity process. All these findings suggest that the virulence of *S. agalactiae* could be related to its high plasticity and the presence of the mobile genetic elements.

**Keywords:** *Streptococcus agalactiae*, ST-17, hypervirulent, comparative genomic, pathogenicity islands, pan-genome.

## Background

*Streptococcus agalactiae* (also known as Group B Streptococcus; GBS) colonizes asymptotically the gastrointestinal and genitourinary tracts of up to 30% of healthy individuals (Lancefield, 1934; Edwards et al., 2005; Baker, 2013). This specie is responsible for invasive infections in neonates, elderly and immunocompromised adults (Gibbs et al., 2004; Edwards et al., 2005; Johri et al., 2006; Saad et al., 2017).

*S. agalactiae* has been classified in ten different capsular types according to their capsular polysaccharide composition (Ia, Ib, and II up to IX) (Slotved et al., 2007; Liu et al., 2018). Several epidemiological studies have pinpointed a remarkable association of *S. agalactiae* capsular type III strains with meningitis, particularly during late-onset disease (LOD) (Phares et al., 2008; Périchon et al., 2017). Capsular type III strains belonging to the clonal complex 17 specifically sequence type 17 (ST-17) have historically been designated as “hypervirulent” (Périchon et al., 2017). The *S. agalactiae* GBS90356 described in this study belongs to the capsular type III and was isolated from a 3-day-old neonate with fatal meningitis.

*S. agalactiae* GBS90356 is able to survive within human peripheral blood monocytes, macrophages and endothelial cells (Peotta et al., 2001; Santos et al., 2005), and to interact with human ICAM-1 and  $\beta$ 1-integrin during endothelial cells infection (Lione et al., 2005). Activation of PKC and the F-actin organization into stress fibers was also described previously (Santos et al., 2005). GBS90356 culture supernatants, rich in metallopeptidases, were able to cleave fibronectin, laminin, type IV collagen, fibrinogen and albumin (Soares et al., 2008). Cleavage of the host extracellular matrix by *S. agalactiae* may be a relevant factor in the process of bacterial dissemination and/or invasion.

*S. agalactiae* genomes comparative analysis demonstrated that the specie contains a pan genome consisting of a core genome composed of genes shared by all strains and a accessory genome consisting of partially shared and strain-specific genes (Tettelin et al., 2005). *S. agalactiae* shares certain features with other pathogenic streptococci. However, the precise repertoire of shared and unique attributes that account for the emergence of *S. agalactiae* as a major pathogen for human populations remains undefined (Tettelin et al., 2002). *S. agalactiae* can also survive in different hosts as fish, bovine, camel and canine (Shabayek and Spellerberg, 2018).

To elucidate the molecular basis for *S. agalactiae* virulence, we sequenced the genome of the first GBS90356 strain isolated in Brazil belonging to the capsular type III (ST-17) isolated from a neonatal with meningitis and performed comparative analyses with other *S. agalactiae* pathogenic strains recovered from public databases.

## Methods

### Growth bacterial and genomic DNA preparation

*S. agalactiae* GBS90356 capsular type III isolated from the cerebrospinal fluid of a 3-day-old neonate with fatal meningitis is the first hypervirulent-ST-17 strain isolated in Brazil. The isolate was identified as group B streptococci and capsular type according to the method described by Poyart and colleagues (2007) with modifications. GBS90356 was cultured on

blood agar base (BAB; Oxoid) containing 5% defibrinated sheep's blood for 24 h at 37°C and subsequently grown in Brain Heart Infusion broth (BHI-Difco) at 37°C up to an optical density of 0.4 at  $\lambda = 540$  nm ( $OD_{540}$ ) ( $\sim 10^8$  C.F.U. ml<sup>-1</sup>). Genomic DNA was extracted using the Pure Link™ Genomic DNA® Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions (Akter et al., 2015).

### Genome Sequencing, assembly and annotation

Genome was sequenced in Illumina HiSeq 2500 platform of the Institute Oswaldo Cruz. Genomic DNA was randomly fragmented at approximately 500 bp for library construction using the Nextera XT kit and sequenced as paired-end reads in Illumina instruments and applied in the steps of quantification, amplification and sequencing according to manufacturer's protocol. The quality of the reads was analyzed in FastQC v. 0.11.2 (Andrews, 2010). Reads were filtered and trimmed by Trimmomatic 0.36 (Bolger et al., 2014). Genome was assembled using SPAdes 3.0.0 (Bankevich et al., 2012). The contigs generated were aligned and ordered by Mauve v. 2.0 (Darling et al., 2010), using the reference genome of *Streptococcus agalactiae* 2603V / R (GenBank Accession AE009948.1). Automatic annotation was performed using Rapid Annotation Subsystem Technology (RAST) server (Aziz et al, 2008). The genome was manually curated with Artemis (Rutherford et al., 2000) and UniProt database (Consortium, 2013).

### Pan-genome of *Streptococcus agalactiae*

The value of  $\alpha$  of the Heap's Law was calculated based on the results of the pan genome obtained in Pan-Genomes Analysis Pipeline (PGAP v.1.11) (Zhao et al., 2012), through the web tool PanWeb (Pantoja et al., 2017). The sequence coverage and identity parameters were 50%, and E-value 0.001. Based on the complete genomes available in Genbank, a total of 109 *S. agalactiae* strains were selected to represent the known diversity of *S. agalactiae* population (included the study strain, GBS90356), being 42 human strains, 3 bovine strains, 55 fish strains and 9 others strains (camel, frog, canine and unknown) used to determine the specie pan genome (Table S1). Subsequently, we calculated the pan genome for 20 *S. agalactiae* genomes belonging capsular type III (included 6 strains ST-17) isolated from human hosts: COH-1 (HG939456.1); NGBS128 (CP012480.1); CU\_GBS\_08 (CP010874.1); CU\_GBS\_98 (CP010875.1); SG-M1 (CP012419.2); NEM316 (AL732656.1); H002 (CP011329.1); BM110 (LT714196.1); Sag158 (CP019979.1); CUGBS591 (CP021862.1); SG-M50 (CP021865.1); SG-M158 (CP021864.1); SG-M29 (CP021866.1); SG-M25 (CP021867.1); SG-M6 (CP021869.1); SG-M4 (CP021870.1); SG-M163 (CP021863.1); 874391 (CP022537.1); HU-GS5823 (AP018935.1) and GBS90356 strain sequenced in this study (Table S1).

### Prediction and analysis of pathogenicity islands and prophages

Pathogenicity islands (PAIs) were identified with GIPSy v1.1.2 software (Soares et al., 2015). The non-pathogenic bacterial strain *Streptococcus thermophilus* JIM8232 isolated from cheese was used as reference. The E-value used was 0.000001 and coverage above 90% to ensure high identity between the sequences. All PAIs and genes identified were annotated manually in the Genbank file using Artemis. The prediction of PAIs was confirmed in the web tool IslandViewer4 using all four methods IslandPick, SIGI-HMM, IslandPath-DIMOB and Islander (Bertelli et al, 2017). The synteny of genes was analyzed using blastn with a minimum coverage cutoff of 50% in the web tool Simple Synteny (Veltri et al., 2016). The

genome of study strain GBS90356 and reference strains, COH1, NGBS128, NEM316 and A909 were used for the synteny analysis. The identification of prophages in the genome was performed with the web tool PHASTER (Phage Search Tool) (Zhou et al., 2011).

### **Nucleotide comparison and secretome characterization**

BRIG software (Blast Ring Image Generator) (Alikhan et al., 2011) was used to design a circular map of *S. agalactiae* genomes including the pathogenicity islands of GBS90356 strain. In this work we used 13 complete genomes of distinct capsular types available in Genbank. Capsular type III was the most represented (7 genomes), particularly 6 isolates belonging to the hypervirulent clonal complex 17 (III/ST-17), followed by types Ia (n= 1), Ib (n= 1), II (n= 1), IV (n= 1), V (n= 1) and VI (n=1). The image was displayed as a series of concentric rings with the center ring being the reference GBS90356. SignalP v.4.1 (Bendtsen et al, 2004) is the currently most used program for prediction of signal peptides from amino acid sequences. Only genes within PAIs were analyzed.

### **Phylogenomic and functional analysis**

A phylogenomic tree using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was calculated in PGAP v.1.11 (Zhao et al., 2012) using the *S. agalactiae* core genome. The tree was designed using SplitsTree v.4.14.5 (Huson et al., 1998) from the .tree file generated by PGAP. Genes within PAIs were functionally analyzed using the web tool GO FEAT (Araújo et al., 2018), which assigns Gene Ontology terms to each sequence and integrates results from Uniprot, Interpro, Pfam and EMBL databases.

## **Results**

### **Genomic features of *Streptococcus agalactiae* GBS90356**

We obtained a genome of *S. agalactiae* GBS90356 with 40 contigs and a size of 2,029,093 bp, N50 of 124,554, 2,023 CDSs, 52 tRNA genes and a G+C content of 35.2%. (Table S1). These features are in accordance with other genomes of *S. agalactiae* sequenced and deposited in public databases. No complete prophage was identified in the genome. A total of 1588 (78.5%) CDSs were classified in subsystems of the RAST server. The most represented categories were Carbohydrates; Protein Metabolism; Amino Acids and Derivatives; Cell Wall and Capsule; DNA Metabolism; Cofactors, Vitamins, Prosthetic Groups, Pigments; Virulence, Disease and Defense. Categories of subsystems important for bacterial virulence such as Cell Wall and Capsule; Virulence, Disease and Defense; Phages, Prophages, Transposable Elements and Plasmids; and Stress Response, comprises 14.2% of the CDSs predicted.

### **Core and Pan-genome evaluation of *Streptococcus agalactiae***

The pan-genome of the 109 strains of *S. agalactiae* presented a total of 5,020 genes of which 706 (14%) belong to the core genome. A total of 51 genes are specific to GBS90356 strain. Based on the Heap's law, the pan-genome of the specie is open ( $\alpha = 0.7525$ ).

The pan-genome of 20 strains of *S. agalactiae* belonging to capsular type III was calculated, including the ST-17 strain sequenced in this study. The pan-genome of *S. agalactiae* capsular type III presented a total of 3,304 genes, which 1,610 (48.73%) belong to the core genome, 1,050 (31.78%) to the accessory genome and 644 (19.49%) genes unique, being 22 single

genes to GBS90356 strain. The core genome of capsular type III (1,610 genes / 48.73% of the pan-genome) is higher than the core genome calculated including the other capsular types (706 genes / 14% of the pan-genome), suggesting a higher similarity among genomes of type III. It should be noted that the pan-genome of *S. agalactiae* capsular type III is open ( $\alpha = 0.8365$ ), although they present the same capsular type, this species demonstrates a great plasticity, in which the insertion of new strains in the analysis other genes may arise.

### Identification and Comparison of Genomes and PAIs

The detection and analysis of Pathogenicity Islands (PAIs) in *S. agalactiae* GBS90356 genome was performed in GIPSY and IslandViewer4 (Table S2). PAIs have a large number of virulence genes grouped in blocks that play important roles in the bacterial pathogenicity. GIPSY predicted 10 PAIs (SagPAI\_1 up to SagPAI\_10) which corresponded to 17% of the genome size with a total of 352 genes (Table S2). Conversely, IslandViewer4 detected only 6 PAIs (SagPAI\_2, 4, 5, 8, 9 and 11). Therefore, results predicted by the GIPSY were considered in this work. PAIs 2 and 5 were partially detected in ST-17 strains (Figure 1), and related to integrase (Figure 2) (Table 1).

The SagPAI\_1 has 13 genes, including the *neu* and *cps* genes which encode sialic acid and the polysaccharide capsule, respectively. In contrast, 33 coding sequences in SagPAI\_2 were identified as uncharacterized proteins and conserved in ST-17 strains (Figure 2). Only 6 genes were described in SagPAI\_3 as the *dltS* (D-alanyl-lipoteichoic acid biosynthesis regulators) and *dltR* (histidine kinase sensors) genes responsible for characterization of *S. agalactiae* specie. Moreover, 4 genes were identified as uncharacterized proteins.

In-depth analysis of the genes presented in SagPAI\_4 was verified the existence of genes with products classified as uncharacterized proteins. It is important to perform studies in order to better characterize the biologic functions. It is noteworthy that in SagPAI\_5, 135 genes were identified, including laminin binding protein (*lmb* gene) and the *lacR*, a putative negative regulator of the operon that binding to the promoter in the absence of lactose.

The genes related to mercury resistance were observed in SagPAI\_6, *merA* (encoding a mercuric reductase) and *merR* (encoding a mercuric resistance operon regulatory protein). Ethanol-Induced Alcohol Dehydrogenase (*adh*) was localized in SagPAI\_8, a housekeeping gene for sequence type characterization by MLST. Genes *vex1-3* and *vncR/S* were located within the SagPAI\_9; although the exact functions of the *vex-vnc* clusters remain to be elucidated. It has been demonstrated that *vex3* genes are associated with resistance to vancomycin, but GBS90356 strain was vancomycin sensitive (data not shown). Finally, in SagPAI\_10 was identified the CAMP (acronym for Christie, Atkins and Munch-Peterson) factor encode *cfb* gene.

As shown in the circular map of Figure 1, the PAIs 2, 4, 5 and 8 are poorly conserved among the *S. agalactiae* strains analyzed in this study. The ST-17 lineages seem to present a higher conservation of these islands compared to GBS90356 strain.

SignalP detected 20 signal peptide proteins for the 352 proteins described within the 10 PAIs identified, suggesting that the majority of proteins involved in the *S. agalactiae* pathogenicity are not secreted. Secreted proteins (60% - 12/20) were located in the SagPAI\_5.

### Synteny analysis of *S. agalactiae* pathogenicity islands

Genes within PAIs were compared by *blastn* with the nucleotide sequence of strains NGBS128; COH1; NEM316; A909. The comparison was performed using the web tool Simple Synteny (Veltri et al., 2016). The PAIs showed high conservation in all *S. agalactiae* strains analyzed compared to the GBS90356 strain. A manual curation identified a gene as Tn5252 present only in ST-17 strains (COH1 and NGB128) and not identified in the NEM316 (CTIII) and A909 (CTIa) strains (Figure 2). Particularly, in SagPAI\_2 we identified a transposon Tn916 and a *tetM* gene which confers resistance to tetracycline (Da Cunha et al., 2014).

### Phylogenomic and functional analysis

To more accurately reflect the evolutionary relationship within *S. agalactiae* strains, we constructed a phylogenomic tree. Bacterial strains of the same sequence type showed close evolutionary relationships, branches of human ST-17 strains were clustered together, indicating an evolutionary bottleneck. Bovine and human strains were alternately distributed on the evolutionary tree. Although isolated human strains were widely represented in this study (Figure 4).

The PAIs genes of *S. agalactiae* were classified into GO functional groups. After comparing the distribution of functional genes, we found biological process (102 genes), cellular component (105 genes) and molecular function (235 genes) commonly shared by *S. agalactiae* strains. GO biological process terms were mainly related to regulation of transcription, DNA-templated, phosphoenolpyruvate-dependent sugar, phosphotransferase system (PTS), DNA integration, DNA recombination, transmembrane transport, translation, transposition, carbohydrate metabolic process, DNA replication, phosphate ion transport, metal ion transport and cell redox homeostasis. Moreover, only one function (integral component of membrane) was the most present GO cellular component term. In addition, 18 molecular functions were mainly related to DNA binding, ATP binding, sequence-specific DNA binding, transferase activity, kinase activity, hydrolase activity, ATPase activity, metal ion binding, methyltransferase activity, DNA-binding transcription factor activity, phosphorelay sensor kinase activity, peptidase activity, protein-N(Pi)-phosphohistidine-sugar phosphotransferase activity, helicase activity, endonuclease activity, inorganic phosphate transmembrane transporter activity, lyase activity and isomerase activity.

GO analysis demonstrated several annotations functional, which 99 were uncharacterized proteins. Thus, we identified 36 uncharacterized proteins with Pfam and Interpro domains, which 16 belonging unknown function, 3 helix-turn-helix, 3 transposon, 3 protein of unknown function with HXXEE motif, 1 antirestriction protein, 1 nucleoside recognition, 1 PTS, 1 ABC transporter, 1 ATPases associated with diverse cellular activities, 1 N-terminal phage replisome organizer, 1 peptidase C39 family, 1 integron-associated effector binding protein, 1 host cell surface-exposed lipoprotein, 1 phospholipase\_D-nuclease N-terminal and 1 nucleotidyltransferase.

### Discussion

*Streptococcus agalactiae* has been considered the main pathogen causing severe invasive infections worldwide (Périchon et al., 2017). Clinical manifestations include infections as pneumonia, sepsis and meningitis occurring immediately after birth reported as Early Onset Disease (EOD; 0–6 days) or after the 1st week of life referred as Late Onset Disease (LOD; 7–90 days) (Lin et al., 2018; Shabayek and Spellerberg, 2018). In this work, we report the first genome sequencing of a Brazilian *S. agalactiae* ST-17 strain isolated in the metropolitan

area of Rio de Janeiro. *S. agalactiae* GBS90356 showed a genome size and G+C content similar to *S. agalactiae* strains recovered from different hosts (Aguiar et al., 2016; He et al., 2017).

The sequence of a single genome does not reflect how genetic variability drives pathogenesis within a bacterial species (Tettelin et al., 2005). Presently, to make estimates more accurate of the size of the pan-genome of *S. agalactiae*, we included 109 genomes from diverse hosts available at the <https://www.ncbi.nlm.nih.gov/genome>, including ST-17 lineages (n=6) to confirmed the high plasticity of *S. agalactiae*. Similar to previous study of *S. agalactiae* genome sequences (n=24; only COH1 ST-17 strain), the bacterial pan-genome continue vast and new genes will be added to the gene pool of the specie (He et al., 2017).

In this work, the pan-genome of 20 human *S. agalactiae* strains belonging to capsular type III was open ( $\alpha = 0.83$ ), showing a complex set of genetic material and a high level of adaptability by horizontal gene transfer (Donati et al., 2010). Compared to other pathogenic bacteria species, such as *Pseudomonas syringae* with 64% of core genome, *Streptococcus pneumoniae* (74%) and *Listeria monocytogenes* (80%), *S. agalactiae* contains one of the smallest core genomes (14%) described in the literature, suggesting a high genetic variability (Baltrus et al., 2011; Puymège et al., 2015; He et al., 2017; Baraúna et al., 2017).

Although, the higher plasticity could be associated to the multiple hosts of *S. agalactiae* (Tettelin et al., 2005; Lefébure and Stanhope, 2007; He et al., 2017), our study using only one host demonstrated that human ST-17 lineages presented an open pan-genome and a high plasticity. Thus, other factors as genetic mobile elements and gene evolution should influence the genomic plasticity of *S. agalactiae* rather than the adaptation to multiple hosts (Teatero et al., 2016; He et al., 2017). In fact, the host-adaptation is the result of a high plasticity, which is achieved due to the action of mobile genetic elements and gene evolution within the species. Moreover, acquisition of mobile genetic elements appears to correlate with clonal expansion of the bacterial strains (Teatero et al., 2016). Mobile genetic elements are a common feature of *S. agalactiae*, which each strain having its own peculiar burden of transposons, phages, integrase and integrated plasmids (Herbert et al., 2005) with potential virulence role. Thus, this study contributes to understanding of pathogenesis by further delineating the hypervirulence of ST-17 lineages.

PAIs are distinct genetic elements on the chromosomes of a large number of bacterial pathogens. PAIs encode diverse virulence factors normally absent in non-pathogenic strains of the same or closely related species (Gal-Mor and Finlay, 2006). Several tools have been developed to *in silico* detection of potential PAIs, including programs that integrate multiple features of PAIs such as atypical sequence composition and horizontal gene transfer associated genes (Hsiao et al., 2005; Gal-Mor and Finlay, 2006). According Silva (2018), GIPSy presented the best overall results than IslandViewer4, achieving 91% coverage. Currently, GIPSy predicted 10 PAIs and IslandViewer4 identified only 6 PAIs. Thus, GIPSy results were considered in this work.

*S. agalactiae* GBS90356 PAIs contain important functional genes as well as unknown genes. Interestingly, the pathogenicity islands of *S. agalactiae* ST-17 strains were highly conserved. In SagPAI\_1, *S. agalactiae* showed only *cps* and *neu* genes that enable the bacteria to evade host immune defenses and confer the ability to development of meningitis (Pailhorie et al., 2013). Our findings demonstrated that SagPAI\_2 presented a well-known large integrative conjugative element (ICE) Tn916 carrying the *tetM* gene conferring resistance to tetracycline

(Da Cunha et al., 2014), corroborating the tetracycline resistance result observed for GBS90356 strain by antimicrobial susceptibility testing (data not shown). Moreover, resistance genes to mercury (*mer* operon) were found (Sotero-Martins et al., 2008). As these elements are capable of conjugative transposition to various bacterial species, they were previously cited as emergence of multiple antibiotic resistance in Gram-positive pathogens (Campisi et al, 2016).

*S. agalactiae* isolates have different types of *lac* operon (Richards et al., 2013); however, we found *lacR* gene located in SagPAI\_5, a regulator binding to the promoter in the absence of lactose or a similar galactoside inducer, inducing a conformational change that leads to dissociation from DNA (Richards et al., 2013). Other gene identified in SagPAI\_5 was *lmb*, an laminin adhesin (*Lmb*) appears to have a pronounced role in bacterial tropism of the central nervous system, allowing *S. agalactiae* to cross the blood-brain barrier (Shabayek and Spellerberg, 2018). *Lmb*, a 34 kDa surface exposed lipoprotein of *S. agalactiae* mediates its adherence to laminin and this event is important in pathogenesis of microorganism which has been identified as a crucial host factor for bacterial invasion (van Wijk et al., 2017). SagPAI\_2 and 5 presented ABC transporters gene essential for cell viability, virulence, nutrition, acquisition and in the export of other substances, including capsular polysaccharides and hemolysin (Davidson et al., 2008; He et al., 2017). Moreover, both PAIs presented two relevant integrase regions, showing that the mobile genetic elements, as well as the genes they carry may contribute with the hypervirulence and plasticity of ST-17 lineage.

Conjugative transposons are mobile genetic elements capable of intercellular transfer allowing their maintenance, dissemination and the regulation of both processes (Brochet et al, 2009). Dissemination and transfer mechanisms of mobile elements within the *S. agalactiae* remain poorly known. In SagPAI\_8, the integrase gene of a streptococcal conjugative transposon Tn5252 was identified. Tn5252 was originally identified as a chromosome-borne element in *Streptococcus pneumoniae* (Churchward, 1995). Characterization of related elements will reveal the diversity of module combinations within ICE, allowing the study of the interaction between different modules. Furthermore, DNA-array hybridization of a set of diverse strains from both human and animal origins revealed that transposons are frequently present within the *S. agalactiae* population, compensating for the absence of plasmids (Munoz-Najar, 1999; Brochet et al., 2008).

The present work also showed resistance genes to vancomycin (*vex3-vncR/S*) in SagPAI\_9, however, the molecular basis about vancomycin in *S. agalactiae* is essentially unknown (Lin et al., 2011). GBS90356 strain showed susceptible to vancomycin (data not shown). In addition, SagPAI\_10 presented *cfb* gene encodes an extracellular pore-forming protein known as CAMP factor, which has been widely used for phenotypic identification of *S. agalactiae* isolates (Christie et al., 1944; Lang and Palmer, 2003). Furthermore, CAMP pore-forming toxin is responsible for the systemic spread of *S. agalactiae* in the host (Rajagopal et al., 2009). Despite the above, the molecular mechanisms and importance of CAMP factor in bacterial pathogenesis remain unknown.

We predicted several hypothetical proteins and unknown genes may be targeted in the future over studies of disease establishment by *S. agalactiae*. Therefore, other factors as genetic mobile elements and gene evolution should influence the genomic plasticity of the species rather than the adaptation to multiple hosts (Teatero et al., 2016). The results indicated that a large portion of the genome encodes for proteins involved in host-pathogen interaction. Our data also related the presence of transposons with uncharacterized proteins, due to transposon



is integrated into the chromosome, but during conjugation excises itself and then moves to the recipient bacterium and re-integrate into its chromosome. Typically a conjugative transposon encodes only the proteins required for this activity and the proteins that regulate it (Brochet et al., 2009).

We analysed only genes within PAIs and demonstrated that 20 genes have signal peptides, suggesting that proteins involved in the pathogenicity of *S. agalactiae* may not be secreted. Nevertheless, evidence indicates that ST-17 lineages display a specific combination of the secreted and surface-exposed proteins (e.g. SrtA and HvgA) (Brochet et al., 2006; Tazi et al., 2010).

We also analyzed other genes of importance for *S. agalactiae* that were predicted outside the PAIs as *hylB*, gene encoding hyaluronate lyase, a surface exposed protein which cleaves hyaluronan and facilitates the invasion and spreading of bacteria through host tissues (Mello et al., 2002; Oviedo et al., 2013). Other genes found were (i) *cylE* ( $\beta$ -hemolysin), has a central role in balancing the pro- and anti-inflammatory responses of the infected host, necessary to breach the epithelial and endothelial barriers and the phagolysosome membrane (Six et al., 2015); (ii) *sodA*, superoxide dismutase, enzyme essential for the protection of *S. agalactiae* against oxidative stress and for survival in hosts (Poyart et al., 2001); (iii) *fbsA* (Fibronectin-Binding Protein-Fbs) related to interactions between bacterial surface adhesins and components of the host plasma and extracellular matrix (ECM) proteins (Ragunathan & Ponnuraj, 2011). Experiments performed with FbsA mutant demonstrated that the repeat region of FbsA contains the binding site for ligand, suggesting that *fbsA* is a *S. agalactiae* important fibrinogen receptor (Pietrocola et al., 2018; Shabayek and Spellerberg, 2018); (iv) *scpB*, C5a peptidase a surface localized serine protease, cleaves human C5a, leading to reduced neutrophil chemotaxis and decreasing opsonophagocytic killing. Additionally, it mediates binding to fibronectin and contributes to the invasion into epithelial cells (Beckmann et al., 2002; Cheng et al., 2002), and (v) *cas9* protein adapted from a naturally occurring genome editing system in bacteria. The bacteria capture snippets of DNA from invading viruses and use them to create DNA segments known as CRISPR arrays (Lier et al., 2013).

Protein domains often correspond to structural domains which are self-stabilizing. They may occur independently or as part of complex multidomain protein architectures which evolve by domain accretion, domain loss or domain recombination (Pundir and O'Donovan, 2017). We can highlight, among the uncharacterized proteins, the following Pfam and Interpro domains found in this study; the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), a major carbohydrate active-transport system in *Streptococcus*, catalyses the phosphorylation of incoming sugar substrates concomitant with their translocation across the cell membrane, facilitating dissemination this pathogen (Deutscher et al., 2006). Furthermore, Helix-turn-helix (HTH) is a major structural domain capable of binding DNA, which each monomer incorporates two  $\alpha$  helices joined by a short strand of amino acids and bind to the major groove of DNA. It can occur in many proteins that regulate the gene expression of *S. agalactiae* (Aravind et al., 2005). Peptide bacteriocins are exported across the cytoplasmic membrane by a dedicated ATP-binding cassette (ABC) transporter. The ABC transporter is the maturation protease and its proteolytic domain resides in the N-terminal part of the protein (Havarstein et al., 1995), demonstrating its importance for pathogenic bacteria. Moreover, sequence-specific DNA binding, transferase activity, hydrolase activity, integral component of membrane, carbohydrate metabolic process.

All these findings suggest that the virulence of *S. agalactiae* is directly related to its high plasticity and the presence of the mobile genetic elements, consequently the sequencing of new genomes will help to determine a more robust set of virulence factors. Understanding the concept of PAIs can have deeply affected the way we perceive bacterial virulence and microbial evolution. Our results showed have implications for pathogenesis, vaccine design, evolution, and the concept of species and suggest that the research strategies for microbial genomes may need to be reconsidered. All together these observations showed the importance of genomic surveillance of epidemiologically relevant lineages and show the power of the *in silico* analysis of the *S. agalactiae* genome to understand the epidemiology and genome plasticity of *S. agalactiae* population.

## Conclusion

The first ST-17 strain isolated in Brazil from cerebrospinal fluid of the 3-day-old neonate with fatal meningitis is a bacterium Gram positive, non-sporulating having a spherical shape and dimensions ranging from 0.2 to 1.0 microns in diameter. The genomic sequencing was performed and resulted in genome size 2,029,093 bp. The *S. agalactiae* pan-genome is now well advanced but still “open”. In the present study, we sequenced the genome of new *S. agalactiae* strain isolated from human in Brazil. We infer the potential epidemiological distribution of ST-17 in human hosts in Brazil. 10 PAIs which correspond to 15% of the genome size with a total of 352 genes (17% of total genes) were predicted, PAIs 2 and 5 were partially detected only in ST-17 strains and related to the presence of an integrase. Another relevant finding is the similarity of 99% between the genome of *S. agalactiae* COH1 and GBS90356 strains, suggesting the conservation of these PAIs in ST-17 strains. PAIs contain important virulence factors for *S. agalactiae*, the PAIs 2, 4, 5 and 8 are poorly conserved among the *S. agalactiae* strains analyzed in this study. Particularly, in the SagPAI\_2 was identified the transposon Tn916 and *tetM* gene conferring resistance to tetracycline. Based on the Heap’s law, the pan-genome of the specie is open ( $\alpha = 0.7525$ ). We evidence high plasticity in the strain GBS90356, although this is not usual in pathogenic bacteria with small genomes. Most of the genes are involved in *S. agalactiae* pathogenicity process, since PAIs comprise a large part of the genome. GO analysis found 102 genes in biological process, (105) cellular component and (235) molecular function commonly shared by *S. agalactiae* strains. Among the uncharacterized proteins, the following Pfam and Interpro domains found were PTS, HTH, ABC, sequence-specific DNA binding, hydrolase activity and others important for *S. agalactiae*. All these findings suggest that the virulence of *S. agalactiae* could be related to its high plasticity and the presence of the mobile genetic elements.

## Abbreviations

GBS: Group B Streptococcus; LOD: Late-Onset Disease; BHI: Brain Heart Infusion broth; EOD: Early Onset Disease; PAI: Pathogenicity Islands; ICE: Integrative Conjugative Element; PTS: PhosphoTransferase System; HTH: Helix-Turn-Helix; ABC: ATP-Binding Cassette.

## Funding

The authors acknowledge the financial support received from Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil) and Sub Reitoria de Pós-Graduação e Pesquisa da Universidade do Estado do Rio de Janeiro (SR-2/UERJ). This study was financed in part by

the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

### **Data accessibility**

The draft genome sequences of *Streptococcus agalactiae* GBS90356 were submitted to the NCBI. The accession number was LIEE00000000.

### **Competing interests**

No conflict of interest declared.

### **Authors' contributions**

PS, RA, ALC and PE participated in the design and discussion of the research. PS, JN, JF, VV, ALM and RT performed out the genome sequencing, assembly and annotation part of the work. RA, MV, RT and ALC provided software and bioinformatics expertise. PS carried out the *in silico* and experimental part of the work and wrote original draft. PS and RA carried out the analysis of the data and wrote the final manuscript. PS, RA, KS, ALC and PE contributed in literature review and manuscript discussion. All authors have read and approved the final manuscript.

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **References**

Aguiar EL, Mariano DC, Viana MV, Benevides L, Souza FR, Oliveira LC, et al. Complete genome sequence of *Streptococcus agalactiae* strain GBS85147 serotype of type Ia isolated from human oropharynx. *Stand Genomic Sci.* 2016; 3: 1-8

Akter R, Vythilingam I, Khaw LT, Qvist R, Lim YAL, Sitam FT, Venugopalan B, Sekaran SD. Simian malaria in wild macaques: first report from Hulu Selangor district, Selangor, Malaysia. *Malaria Journal.* 2015; 14, 386.

Alikhan N-F, Petty NK, Ben-Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011; 12, 402.

Al Safadi R, Mereghetti L, Salloum M, Lartigue MF, Virlogeux-Payant I, Quentin R, et al. Two-component system *RgfA/C* activates the *fbsB* gene encoding major fibrinogen-binding protein in highly virulent CC17 clone group B *Streptococcus*. *PLoS One.* 2011; doi:10.1371/journal.pone.0014658

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

- Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiology Reviews*. 2005; 29, 231–262.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Zagnitko O. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 2008; 9, 75.
- Baker CJ. The spectrum of perinatal group B streptococcal disease. *Vaccine*. 2013; 31, D3–6.
- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, et al. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog*. 2011; 7:e1002132.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol*. 2012; 19, 455–477.
- Baraúna RA, Ramos RTJ, Veras AAO, Pinheiro KC, Benevides LJ, et al. Assessing the Genotypic Differences between Strains of *Corynebacterium pseudotuberculosis biovar equi* through Comparative Genomics. *PLOS ONE*. 2017; <https://doi.org/10.1371/journal.pone.0170676>
- Bendtsen JD, Nielsen H, Von-Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*. 2004; 340, 783-795.
- Bisharat N, Jones N, Marchaim, D, Block C, Harding RM, et al. Population structure of group B Streptococcus from a low-incidence region for invasive neonatal disease. *Microbiol*. 2005; 151, 1875–1881.
- Bohnsack JF, Whiting A, Gottschalk M, Dunn DM, Weiss R, et al. Population structure of invasive and colonizing strains of *Streptococcus agalactiae* from neonates of six U.S. academic centers from 1995 to 1999. *J. Clin. Microbiol*. 2008; 46, 1285–1291.
- Bolger AM, Lohse M, & Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30, 2114–2120.
- Brochet M, Couve E, Zouine M, Vallaëys T, Rusniok C, Lamy MC, et al. Genomic diversity and evolution within the species *Streptococcus agalactiae*. *Microbes Infect*. 2006; 8, 1227–1243.
- Brochet M, Couvé E, Glaser P, Guedon G, Payot S. Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. *J Bacteriol*. 2008; 190, 6913–6917.
- Brochet M, Da Cunha V, Couve E, Rusniok C, Trieu-Cuot P, Glaser P. Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. *Mol. Microbiol*. 2009; 71, 948–959.
- Campisi E, Rosini R, Ji W, Guidotti S, Rojas-López M, Geng G, Rinaudo CD. Genomic Analysis Reveals Multi-Drug Resistance Clusters in Group B Streptococcus CC17 Hypervirulent Isolates Causing Neonatal Invasive Disease in Southern Mainland China. *Frontiers in Microbiology*. 2016; 7, 1265.

Consortium TU. The UniProt Consortium, Consortium TU Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* 2013; 41, D43–7.

Cheng Q, Stafslie D, Purushothaman SS, Cleary P. The group B streptococcal C5a peptidase is both a specific protease and an invasin. *Infect. Immun.* 2002; 70, 2408–2413.

Churchward JRS GG. Conjugative transposition. *Annual Review in Microbiology.* 1995; 49, 367-397.

Christie R, Atkins NE, Munch-Petersen EA. Note on a lytic phenomenon shown by group B streptococci. *Aust J Exp Biol Med Sci.* 1944; 22:197–200.

Da Cunha V, Davies MR, Douarre PE, Rosinski-Chupin I, Margarit I, Spinali S, et al. *Streptococcus agalactiae* clones infecting humans were selected and fixed through the extensive use of tetracycline. *Nat. Commun.* 2014; 5:45-44

Darling AE, Mau B, Perna NT. Progressive Mauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS ONE.* 2010;doi.org/10.1371/journal.pone.0011147

Davidso AL, Dassa E, Orelle C, Chen J. Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiology and Molecular Biology Reviews.* 2008; 72, 317–364.

Deutscher J, Francke C & Postma PW. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and molecular biology reviews.* *MMBR.* 2006; 70: 939–1031.

Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, et al. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biol.* 2010; 11:107.

Doran KS, Nizet V. Molecular pathogenesis of neonatal group B streptococcal infection: no longer in its infancy. *Mol Microbiol* 2004; 54:23-31.

D'Urzo N, Martinelli M, Pezzicoli A, De Cesare V, Pinto V, Margarit I, et al. Acidic pH Strongly Enhances In Vitro Biofilm Formation by a Subset of Hypervirulent ST-17 *Streptococcus agalactiae* Strains. *Applied and Environmental Microbiology* 2014; 7:2176-2185.

Edwards MS, Baker CJ. Group B streptococcal infections in elderly adults. *Clin Infect Dis.* 2005; 41, 839–47.

Emonet S, Schrenzel J, Tejada BM. Molecular-based screening for perinatal group B streptococcal infection: implications for prevention and therapy. *Mol Diagn Ther.* 2013; 17:355–361.

Gal-Mor O1, Finlay BB. Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cell Microbiol.* 2006; 11:1707-19.

Gibbs RS, Schrag S, Schuchat A. Perinatal infections due to group B streptococci. *Obstet Gynecol.* 2004. 104, 1062–1076.

- Havarstein LS, Diep DB, Nes IF. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol* 1995; 16, 229-240.
- He EM, Chen CW, Guo Y, Hsu MH, Zhang L, et al. The genome of serotype VI *Streptococcus agalactiae* serotype VI and comparative analysis. *Gene*. 2017; 597, 59–65.
- Herbert MA, Beveridge CJ, McCormick D, Aten E, Jones N, Snyder LA, Saunders NJ. Genetic islands of *Streptococcus agalactiae* strains NEM316 and 2603VR and their presence in other Group B Streptococcal strains. *BMC Microbiology*. 2005; 5, 31.
- Hsiao WW, Ung K, Aeschliman D, Bryan J, Finlay BB, and Brinkman FS. Evidence of a large novel gene pool associated with prokaryotic genomic islands. *PLoS Genet*. 2005; 1: e62.
- Johri AK, Paoletti LC, Glaser P, Dua M, Sharma PK, Grandi G, Rappuoli R. Group B Streptococcus: global incidence and vaccine development. *Nat Rev Microbiol*. 2006; 4, 932–942.
- Jones N, Bohnsack JF, Takahashi S, Oliver KA, Chan MS, Kunst F, et al. Multilocus sequence typing system for group B streptococcus. *J Clin Microbiol*. 2003; 41, 2530–2536.
- Jones N, Oliver KA, Barry J, Harding RM, Bisharat N, Spratt BG, Derrick TPW. Enhanced Invasiveness of Bovine-Derived Neonatal Sequence Type 17 Group B Streptococcus Is Independent of Capsular Serotype. *Oxford Group B Streptococcus Consortium. Clinical Infectious Diseases*. 2006; 42-7, 915–924.
- Lamy MC, Dramsi S, Billoet A, Reglier-Poupet H, Tazi, A, Raymond J, et al. Rapid detection of the “highly virulent” group B Streptococcus ST-17 clone. *Microbes Infect*. 2006; 8, 1714–1722.
- Lancefield RC. A serological differentiation of specific types of bovine hemolytic streptococci (group B). *J Exp Med*. 1934; 59, 441–458.
- Lang S, Palmer M. Characterization of *Streptococcus agalactiae* CAMP factor as a pore-forming toxin. *J. Biol. Chem*. 2003; 278:38167–38173.
- Lefébure T and Stanhope MJ. Evolution of the core and pan-genome of Streptococcus: positive selection, recombination, and genome composition. *Genome Biology*. 2007; 8: 71.
- Li W, Wang HQ, He RZ, Li YW, Su YL, Li AX. Major surfome and secretome profile of *Streptococcus agalactiae* from Nile tilapia (*Oreochromis niloticus*): Insight into vaccine development. *Fish Shellfish Immunol*. 2016; 55, 737-46.
- Lin FP-Y, Lan R, Sintchenko V, Gilbert GL, Kong F, et al. Computational Bacterial Genome-Wide Analysis of Phylogenetic Profiles Reveals Potential Virulence Genes of *Streptococcus agalactiae*. *PLoS ONE*. 2011; 6(4): e17964.
- Lin SM, Zhi Y, Ahn KB, Lim S, Seo HS. Status of group B streptococcal vaccine development. *Clin Exp Vaccine Res*. 2018; 7, 76-81.
- Lione VO, Santos GS, Hirata Júnior R, Mattos-Guaraldi AL, Nagao PE. Involvement of intercellular adhesion molecule-1 and  $\beta 1$  integrin in the internalization process to human

endothelial cells of group B *Streptococcus* clinical isolates. *International Journal of Molecular Medicine*. 2005; 15, 153-157.

Liu<sup>a</sup> J, Yang L, Peters BM, et al. Complete sequence of pBM413, a novel multidrug-resistance megaplasmid carrying qnrVC6 and blaIMP-45 from *Pseudomonas aeruginosa*. *Int. J. Antimicrob. Agents*. 2017; 17: 145–150.

Liu J, Xua R, Zhong H, Zhong Y, Xie Y, Lia LL, Bing CD. Prevalence of GBS serotype III and identification of a ST 17-like genotype from neonates with invasive diseases in Guangzhou, China Zhenbo Xua. *Microbial Pathogenesis*. 2018; 120, 213–218.

Luan SL, Granlund M, Sellin M, Lagergård T, Spratt BG, Norgren M. Multilocus sequence typing of Swedish invasive group B streptococcus isolates indicates a neonatally associated genetic lineage and capsule switching. *J. Clin. Microbiol.* 2005; 43, 3727–3733.

Manning SD, Springman AC, Lehotzky E, Lewis MA, Whittam TS, Davies HD. Multilocus sequence types associated with neonatal group B streptococcal sepsis and meningitis in Canada. *J. Clin. Microbiol.* 2009; 47, 1143–1148.

Munoz-Najar UVM. An operon that confers UV resistance by evoking the SOS mutagenic response in streptococcal conjugative transposon Tn5252. *Journal of Bacteriology*. 1999; 181(9), 2782-2788.

Musser JM, Mattingly SJ, Quentin R, Goudeau A, Selander RK. Identification of a high-virulence clone of type III *Streptococcus agalactiae* (group B *Streptococcus*) causing invasive neonatal disease. *Proc. Natl. Acad. Sci.* 1989; 86, 4731– 4735.

Nielsen H. Predicting Secretory Proteins with SignalP. In: Kihara D. *Protein Function Prediction*. Humana Press, New York, NY. *Methods in Molecular Biology*. 2017; 16:11.

Otaguiri ES, Morguette AEB, Morey AT, Tavares ER, Kerbauy G, de Almeida Torres, RSL, Yamada-Ogatta SF. Development of a melting-curve based multiplex real-time PCR assay for simultaneous detection of *Streptococcus agalactiae* and genes encoding resistance to macrolides and lincosamides. *BMC Pregnancy and Childbirth*. 2018; 18, 126.

Pailhories H, Quentin R, Lartigue M-F. The transcription of the *neuD* gene is stronger in serotype III group B streptococci strains isolated from cerebrospinal fluid than in strains isolated from vagina. *FEMS Microbiology Letters*. 2013; 349, 71–75.

Pantoja Y, Pinheiro K, Veras A, Araújo F, Lopes de Sousa A, Guimarães LC, Silva A, Ramos RTJ. PanWeb: A web interface for pan-genomic analysis. *PLoS ONE*. 2017; 12(5): e0178154.

Peotta VA, Ribeiro FL, Arnholdt ACV, Nagao PE. Human monocytes and monocyte-derived macrophages phagocytosis of serotype III group B streptococci strains. *Curr. Microbiol.* 2001; 43, 64-68.

Périchon B, Szili N, du Merle L, RosinskiChupin I, Gominet M, Bellais S, Poyart C, Trieu-Cuot P, Dramsi S. Regulation of PI-2b Pilus Expression in Hypervirulent *Streptococcus agalactiae* ST-17 BM110. *PLoS ONE*. 2017; 12(1), e0169840.

Phares CR, Lynfield R, Farley MM, Mohle-Boetani J, Harrison LH, Petit S, et al. Epidemiology of invasive group B streptococcal disease in the United States, 1999–2005. *JAMA*. 2008; 299, 2056–2065.

Phillips EA, Tapsall JW, Smith DD. Rapid tube CAMP test for identification of *Streptococcus agalactiae* (Lancefield group B). *J. Clin. Microbiol.* 1980; 12:135–137.

Pietrocola G, Arciola CR, Rindi S, Montanaro L and Speziale P. *Streptococcus agalactiae* Non-Pilus, Cell Wall-Anchored Proteins: Involvement in Colonization and Pathogenesis and Potential as Vaccine Candidates. *Front. Immunol.* 2018; 9:602.

Poyart C, Tazi A, Réglier-Poupet H, Billoët A, Tavares N, Raymond J, Trieu- Cuot P. Multiplex PCR Assay for Rapid and Accurate Capsular Typing of Group B Streptococci. *J Clin Microbiol.* 2007; 1985–1988.

Poyart C, Réglier-Poupet H, Tazi A, Billoët A, Dmytruk N, Bidet P, et al. Invasive group B streptococcal infections in infants, France. *Emerg. Infect. Dis.* 2008; 14, 1647–1649.

Pundir S, Martin MJ, O'Donovan C. UniProt Protein Knowledgebase. *Methods Mol. Biol.* 2017; 1558, 41-55.

Puymège A, Bertin S, Guédon G, Payot S. Analysis of *Streptococcus agalactiae* pan-genome for prevalence, diversity and functionality of integrative and conjugative or mobilizable elements integrated in the tRNALys CTT gene. *Mol. Genet. Genomics.* 2015; doi:10.1007/s00438-015-1031-9.

Rajagopal L. Understanding the regulation of group B streptococcal virulence factors. *Future Microbiol.* 2009; 4, 201–221.

Richards VP, Choi SC, Bitar PDP, Gurjar AA, Stanhope MJ. Transcriptomic and genomic evidence for *Streptococcus agalactiae* adaptation to the bovine environment. *BMC Genomics.* 2013; 14, 920.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. Artemis: sequence visualization and annotation. *Bioinforma Oxf. Engl.* 2000; 16, 944-5.

Saad EJ, Baenas DF, Boisseauc CS, García MJ, Núñez SA, Sanchez PE, Balderramo DC, Hernández D, Caeiro JP. *Streptococcus agalactiae* bacteremia in non-pregnant adult patients at two teaching hospitals. *Rev. Argent. Microbiol.* 2017; doi.org/10.1016/j.ram.2017.08.002

Santos GS, Lione VO, Costa E Silva Filho F, Nagao PE. Signal transduction in human endothelial cells induced by their interaction with group B Streptococci. *International Journal of Molecular Medicine.* 2005; 15, 859-863.

Schuchat A. Epidemiology of group B streptococcal disease in the United States: shifting paradigms. *Clin. Microbiol. Rev.* 1998; 11:497-513.

Shabayek S, Spellerberg B. Group B Streptococcal Colonization, Molecular Characteristics and Epidemiology. *Front. Microbiol.* 2018; 9, 437.



Silva Filho, A. C., Raittz, R. T., Guizelini, D., De Pierri, C. R., Augusto, D. W., Dos Santos-Weiss, I., & Marchaukoski, J. N. Comparative Analysis of Genomic Island Prediction Tools. *Frontiers in genetics*. 2018; 9, 619.

Slotved HC, Kong F, Lambertsen L, Sauer S, Gilbert GL. Serotype IX, a proposed new *Streptococcus agalactiae* serotype. *J. Clin. Microbiol.* 2007; 45, 2929–2936.

Soares GC, Santos MHB, Costa AF, Santos AL, Morandi V, Nagao PE. Metallopeptidases produced by group B *Streptococcus*: influence of proteolytic inhibitors on growth and on interaction with human cell lineages. *Int. J. Mol. Med.* 2008; 22, 119–125.

Soares SC, Geyik H, Ramos RT, de Sá PH, Barbosa EG, Baumbach J, et al. GIPSy: Genomic island prediction software. *Journal of Biotechnology*. 2015; 232, 2-11.

Sotero-Martins A, de Jesus MS, Lacerda M, Moreira JC, Filgueiras ALL, Barrocas PRG. A conservative region of the mercuric reductase gene (*merA*) as a molecular marker of bacterial mercury resistance. *Brazilian Journal of Microbiology*. 2008; 39(2): 307–310.

Sullivan MJ, Forde BM, Prince DW, Ipe DS, Ben Zakour NL, Davies MR, Dougan G, Beatson SA, Ulett GC. Complete genome sequence of serotype III *Streptococcus agalactiae* sequence type 17 strain 874391. *Genome Announc.* 2017; doi.org/10.1128/genomeA.01107-17.

Sunde RA. mRNA transcripts as molecular biomarkers in medicine and nutrition. *The Journal of nutritional biochemistry*, 2010; 21: 665–670.

Tazi A, Disson O, Bellais S, Bouaboud A, Dmytruk N, Dramsi S, Mistou MY, Khun H, Mechler C, Tardieux I, Trieu-Cuot P, Lecuit M, Poyart C. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. *J. Exp. Med.* 2010; 207: 2313–2322.

Teatero S, Ramoutar E, McGeer A, Li A, Melano RG, Wasserscheid J, Fittipaldi N. Clonal Complex 17 Group B *Streptococcus* strains causing invasive disease in neonates and adults originate from the same genetic pool. *Scientific Reports*. 2016; 6:20047.

Teatero S, Athey TBT, Van Caesele P, Horsman G, Alexander DC, Melano RG, et al. Emergence of Serotype IV Group B *Streptococcus* Adult Invasive Disease in Manitoba and Saskatchewan, Canada, Is Driven by Clonal Sequence Type 459 Strains. *Journal of Clinical Microbiology*. 2015; 53(9): 2919–2926.

Tettelin H, Massignani V, Cieslewicz MJ, Eisen JA, Peterson S, Wessels MR, et al. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc. Natl. Acad. Sci. USA*. 2002; 99: 12391–12396.

Tettelin H, Massignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial ‘pan-genome’. *Proc. Natl. Acad. Sci. USA*. 2005; 102:13950–13955.

van Wijk, XM, Döhrmann, S, Hallström, BM, Li, S, Voldborg, BG, Meng, BX, ... & Lewis, NE. Whole-Genome Sequencing of Invasion-Resistant Cells Identifies Laminin  $\alpha 2$  as a Host Factor for Bacterial Invasion. *Mol bio.* 2017; 8: e02128-16.

Veltri D, Malapi-Wight M, Crouch JA. Simple Synteny: a web-based tool for visualization of microsynteny across multiple species. *Nucleic Acids Res.* 2016; 44: W41–W45.

Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. PGAP: pan-genomes analysis pipeline. *Bioinformatics.* 2012; 28: 416–418.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart, DS. PHAST: a fast phage search tool. *Nucleic Acids research.* 2011; 39:W347–W352.

## FIGURE LEGENDS

**FIGURE 1 | Circular map of the genomes to compare the nucleotide identity between *S. agalactiae* strains and to determine the conservation of PAIs.** The genomes were compared by BLASTn, and the percent identity between them was determined by the intensity of color in the circular rings. The genomes of *S. agalactiae* were identified by the name of the strain. The innermost ring to the outermost is presented in this figure, as follows: the GC contents and the GC skew of the GBS90356 strain, the genomes of the *Streptococcus agalactiae* strains GBS90356; NGBS572; 2603V/R; 874391; A909; BM110; COH-1; GBS-NY; GBS-M002; NEM316; NGBS128; Sag37; SG-M6, respectively and the three outermost rings comprise the location of the PAIs and CDSs (red) of *S. agalactiae* GBS90356.

**FIGURE 2 | Synteny analysis of *S. agalactiae* pathogenicity islands 2 (A) and 5 (B).** From top to bottom of the figure, the genomes of strains GBS90356, COH1, NGBS128, A909 and NEM316 are presented. Genes are represented by arrows of different colors. Genes that were found only in ST-17 strains are highlighted by red rectangles in *S. agalactiae* NGBS128 genomes. These genes are described in Table 1.

**FIGURE 3 | Prediction of the pan-genome of *Streptococcus agalactiae*.** The graph shows the size of the pan-genome (blue graphic bars) and size of the core genome (red bar graphs), the dotted lines indicate the standard deviation and white circles outside the dotted lines indicate discordant values of the analysis, where in the Y axis we have the size of the pan-genome and core genome and in the X axis we have the number of *S. agalactiae* strains used in the comparative analysis performed in the PGAP software. The calculation of the pan-genome and core genome size showed very similar values using either the mean (yellow line) or median (green line) of the possible combinations.

**FIGURE 4 | Phylogenomic tree of 109 *S. agalactiae* strains based on genome.** The phylogenomic relationships among *S. agalactiae* strains of different hosts. On the left is the phylogenomic tree of 109 *S. agalactiae* strains (Table S1). On the right is the information about host, capsular type and sequence type of the strains.

## TABLES

**TABLE 1.** Genes of PAIs 2 and 5 that were found only in ST-17 strains.

Gene size (bp)	Gene name	Product name
<b><i>S. agalactiae</i> Pathogenicity Island 2 (30,193 bp)</b>		
1194	–	Integrase, superantigen-encoding pathogenicity islands SaPI
201	–	Hypothetical protein
231	–	Hypothetical protein
423	–	Hypothetical protein
354	–	Hypothetical protein
1920	<i>tetM</i>	Tetracycline resistance protein TetM
276	–	Hypothetical protein
597	–	Hypothetical protein
1023	–	Lipoprotein, NLP/P60 family
2028	–	membrane protein, putative
2454	–	Hypothetical protein
396	–	Hypothetical protein
1026	–	Hypothetical protein
501	–	Antirestriction protein
762	–	Hypothetical protein
222	–	Conjugation related protein
273	–	Hypothetical protein
1185	–	Tn916, transcriptional regulator putative
1098	–	Hypothetical protein
159	–	Hypothetical protein
774	–	Hypothetical protein
378	–	Hypothetical protein
<b><i>S. agalactiae</i> Pathogenicity Island 5 (124,906 bp)</b>		
1203	–	Integrase
258	–	Hypothetical protein
870	–	Hypothetical protein
1161	–	FtsK-like DNA segregation ATPase, YDCQ <i>B.subtilis</i> ortholog
426	–	Hypothetical protein
162	–	Hypothetical protein
1161	–	Hypothetical protein
1197	–	serine/threonine protein kinase
810	–	type II restriction enzyme <i>eco47ii</i>