



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

CENTRO DE CIÊNCIAS DA SAÚDE

PÓS-GRADUAÇÃO EM BIOTECNOLOGIA VEGETAL E BIOPROCESSOS

JOÃO RICARDO VIDAL AMARAL

GENÔMICA E BIOPROSPECÇÃO DE SUBSTÂNCIAS

ANTIMICROBIANAS EM *Streptococcus agalactiae* ISOLADOS DE

MASTITE BOVINA

Rio de Janeiro

2021

JOÃO RICARDO VIDAL AMARAL

**GENÔMICA E BIOPROSPECÇÃO DE SUBSTÂNCIAS
ANTIMICROBIANAS EM *Streptococcus agalactiae* ISOLADOS DE
MASTITE BOVINA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Biotecnologia Vegetal e Bioprocessos da Universidade Federal do Rio de Janeiro, como parte dos requisitos para a obtenção do grau de Mestre em Biotecnologia Vegetal e Bioprocessos

Orientador: Alexandre Soares Rosado

Coorientadora: Selma Soares de Oliveira

Rio de Janeiro

2021

Ficha catalográfica :

VIDAL AMARAL, João Ricardo

Genômica e bioprospecção de substâncias antimicrobianas em *Streptococcus agalactiae* isolados de mastite bovina

154 f.

Dissertação de Mestrado em Biotecnologia Vegetal e Bioprocessos

1 – Peptídeos antimicrobianos

2 – *Streptococcus agalactiae*

3 – Resistência antimicrobiana

4 - Genômica

I. Universidade Federal do Rio de Janeiro

II. Mestrado em Biotecnologia Vegetal e Bioprocessos

AGRADECIMENTOS

Aos meus pais, Paulo Roberto Amaral e Elizabete Vidal Amaral, agradeço infinitamente por todo amor a mim dedicado e por serem meus grandes exemplos de determinação e força. Ao meu pai, agradeço por sua serenidade e apoio incondicional nas minhas escolhas. À minha mãe, agradeço pelos pequenos gestos de carinho e preocupação que me mantiveram forte durante os momentos mais difíceis dessa minha jornada.

À minha melhor amiga, Yngrid Synara de Sena Silva, eu não tenho palavras para descrever toda minha gratidão pelo seu companheirismo, preocupação e carinho para comigo pelos já quase 2 séculos de vida. Uma das poucas certezas da minha vida é que nossa amizade ainda perdurará por muitos anos.

À minha prima Alessandra Vidal de Farias, a qual considero como uma irmã mais nova, por sua personalidade sonhadora e sua incrível capacidade de sempre me fazer rir, até nos momentos mais impróprios. Poder compartilhar tantos momentos contigo é um grande presente para mim.

Aos meus amigos de faculdade, Marcelle Pita e David Silva, os quais conheço desde o início da minha graduação, por sempre mantermos contato mesmo trilhando caminhos diferentes. Nossas conversas sempre trazem humor para meus dias e me mantêm focado em meus objetivos. Desejo e sempre desejarei o melhor para vocês.

Aos meus amigos e companheiros de laboratório por toda ajuda e todos os momentos felizes que compartilhamos nesses anos: Sarah Mederos, Ana Carolina, Júnia Schultz, Barbara Peckle, Renato, José Roberto, Felipe Diniz, Antônio e muitos outros que participaram dessa história. Um agradecimento especial à Sheila da Silva por todo conhecimento a mim passado ainda quando na função de minha coorientadora de graduação. Eles foram essenciais para meu trabalho de mestrado. Sei que tenho em ti uma grande amiga que sempre está disposta a me auxiliar. Desejo todo sucesso do mundo na sua carreira.

Ao meu orientador, Alexandre Soares Rosado, e a minha coorientadora, Selma Soares de Oliveira, por terem aberto seus braços e me dado a oportunidade de trabalhar e aprender com os senhores. Agradeço pelo que representam e pelo exemplo que são para mim e também pelo apoio e pela paciência que demonstraram.

Ao Programa de Pós-Graduação em Biotecnologia Vegetal e Bioprocessos (PBV) da UFRJ e ao Instituto de Microbiologia Paulo de Góes também da UFRJ, pelos conhecimentos que adquiri nesses últimos anos, pelas incríveis pessoas que tive a oportunidade de conhecer e pelas experiências que me tornaram o cientista que sou hoje.

Gostaria de homenagear e agradecer também à ciência e aos pesquisadores de todo o mundo, assim como à Fiocruz e aos trabalhadores de saúde, por seus incansáveis esforços no combate à pandemia da COVID-19 nesses longos e penosos meses. Viva o SUS!

Por fim, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de mestrado a mim confiada, a qual permitiu que eu me dedicasse total e exclusivamente a minha pesquisa de mestrado.

RESUMO

A resistência antimicrobiana representa um dos grandes desafios mundiais da atualidade. A acelerada notificação de estirpes resistentes a um ou múltiplos antimicrobianos coloca em risco a qualidade de vida da sociedade atual, uma vez que caminhamos em direção à era pós-antibióticos. Nesse cenário, em decorrência de um esgotamento de formas de combater infecções por microrganismos, as mesmas podem tornar-se intratáveis. Os peptídeos antimicrobianos constituem uma alternativa aos tradicionais antibióticos, já que, possuem forte capacidade inibitória contra bactérias, fungos e vírus e são produzidos pelos mais diversos tipos de organismos. Em trabalho anterior a presente dissertação, foi observado um espectro de inibição com foco em espécies do gênero *Streptococcus*, tais quais *S. pyogenes*, *S. agalactiae*, *S. bovis* e *S. uberis*. No presente trabalho, é investigada a atividade antimicrobiana de 4 estirpes de *Streptococcus agalactiae* isolados de mastite bovina, oriundos de fazendas de gado de São Paulo e Minas Gerais, por meio de testes de antibiose contra *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Cellulomonas fimi* e *Micrococcus* sp., além da realização de sequenciamento genômico Illumina Hi-Seq 2500 e análises de bioinformática. Três estirpes tiveram seus genomas sequenciados com êxito e sua taxonomia confirmada via análise por sequências multilocus com inferência filogenética pela metodologia de máxima verossimilhança. Análises do potencial de virulência e resistência antimicrobiana sugeriram que a estirpe paulista LGMAI_St_08 é levemente mais patogênica que as outras, além de identificar o gene *mreA*, o qual está associado à resistência à eritromicina, azitromicina e espiramicina. Buscas por genes de metabólitos secundários e peptídeos antimicrobianos detectaram o gene *zooA*, responsável por codificar a bacteriocina zoocina A, a qual se caracteriza como uma exoenzima lítica que atua na parede celular bacteriana. Genômica comparativa das estirpes de *Streptococcus agalactiae* LGMAI evidenciou, como esperado, alto nível de similaridade entre os genomas, com as estirpes mineiras sendo um pouco mais similares entre si. Dessa forma, as estirpes de *Streptococcus agalactiae* isolados de mastite bovina apresentam interessante potencial aplicação biotecnológica no combate de *Streptococcus* spp. patogênicas, responsáveis por zoonoses e infecções em seres humanos.

Palavras-chave: peptídeos antimicrobianos; *Streptococcus agalactiae*; resistência antimicrobiana; genômica.

ABSTRACT

Antibiotic resistance is one of the biggest challenges of our time. The accelerated identification rate of strains resistant to one or multiple antimicrobials threatens today's quality of life. In this scenario, microorganisms' infections could become untreatable as a result of a depletion in ways to fight them. Antimicrobial peptides represent an alternative to traditional antibiotics since they hold inhibitory activity against bacteria, fungi, and viruses and are produced by many organisms. Previous work with these strains revealed an inhibition spectrum focused on species of the genus *Streptococcus*, namely *S. pyogenes*, *S. agalactiae*, *S. bovis*, and *S. uberis*. Here, we assessed the antimicrobial activity of 4 *Streptococcus agalactiae* strains isolated from bovine mastitis from farms in São Paulo and Minas Gerais. For this, we performed antibiosis assays against *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Cellulomonas fimi*, and *Micrococcus* sp.. In addition, Illumina HiSeq 2500 whole-genome sequencing, and bioinformatic analyses followed. Three strains had their genomes successfully sequenced, allowing for their taxonomic confirmation via an MLSA analysis using the maximum likelihood method. Virulence potential and antibiotic resistance assessments suggested that strain LGMAI_St_08 is slightly more pathogenic than the others and identified the *mreA* gene in the three isolates. This gene is associated with resistance against erythromycin, azithromycin, and spiramycin. Analyses for secondary metabolites and antimicrobial peptides detected the gene *zooA*, coding for the bacteriocin zoocin A. This peptide is a lytic exoenzyme that targets bacterial cell wall. Finally, comparative genomics evidenced, as expected, high similarity among the genomes, with a greater similarity between strains LGMAI_St_11 and LGMAI_St_14. In this manner, the present work shows promising antimicrobial and biotechnological potential for the *Streptococcus agalactiae* bovine isolates.

Keywords: antimicrobial peptides; *Streptococcus agalactiae*; antimicrobial resistance; genomics.

LISTA DE FIGURAS

| | |
|--|----|
| FIGURA 1: Linha do tempo dos principais eventos de resistência antimicrobiana..... | 15 |
| FIGURA 2: Gráfico de barras da distribuição de resistência antimicrobiana global, baseada em renda..... | 17 |
| FIGURA 3: Morte em decorrência de resistência antimicrobiana por ano, comparada a outras grandes causas de morte | 19 |
| FIGURA 4: Publicações e patentes por ano para peptídeos antimicrobianos | 21 |
| FIGURA 5: Evolução de resistência contra peptídeos antimicrobianos | 23 |
| FIGURA 6: Status de desenvolvimento de peptídeos antimicrobianos | 23 |
| FIGURA 7: Peptídeos antimicrobianos atualmente em testes clínicos e os organismos dos quais eles são derivados | 24 |
| FIGURA 8: Exemplos diversos da microbiota produtora de antibióticos em diferentes sítios do corpo | 26 |
| FIGURA 9: Mecanismos de formação de poros em bicamadas fosfolipídicas por meio da ação de bacteriocinas | 30 |
| FIGURA 10: Distribuição geográfica da prevalência de <i>Streptococcus</i> grupo B em mulheres grávidas no Brasil | 38 |
| FIGURA 11: Resultados da análise de completude por busca de ortólogos universais realizada com a ferramenta Busco..... | 61 |
| FIGURA 12: Mapa de calor dos resultados obtidos pela análise de ANI | 63 |
| FIGURA 13: Árvore filogenética com o 16S rRNA | 65 |
| FIGURA 14: Árvore filogenética, inferida por MLSA, com os genes <i>aroE</i> , <i>gki</i> , <i>pheS</i> e <i>recA</i> e o gene para o 16S rRNA | 67 |
| FIGURA 15: Resultados da sorotipagem <i>in silico</i> com dados de sequenciamento genômico para <i>Streptococcus</i> grupo B | 68 |
| FIGURA 16: Alinhamento com Clustal X das sequências de aminoácidos do peptídeo antimicrobiano LGMAI zoocina A..... | 78 |
| FIGURA 17: Alinhamento da sequência da zoocina A LGMAI com sequências de zoocina A do Uniprot | 79 |
| FIGURA 18: Alinhamento da sequência da zoocina A LGMAI com sequências de zoocina A do Uniprot | 80 |

| | |
|--|----|
| FIGURA 19: Alinhamento com Clustal X das sequências de aminoácidos do fator de imunidade zif | 81 |
| FIGURA 20: Anotação funcional de grupos ortólogos para as estirpes LGMAI..... | 84 |
| FIGURA 21: Anotação funcional de grupos ortólogos KEGG para as estirpes LGMAI..... | 86 |
| FIGURA 22: Diagrama de Venn, mostrando os resultados gerais da análise de genômica funcional comparativa das estirpes LGMAI | 87 |
| FIGURA 23: Distribuição da contagem de clusters e de proteínas identificados pela ferramenta de comparação e anotação de genomas OrthoVenn2 | 88 |
| FIGURA 24: Mapa hierárquico dos grupos e subgrupos funcionais anotados (A) exclusivamente para a estirpe LGMAI_St_08 e (B) compartilhados entre as estirpes LGMAI_St_08 e LGMAI_St_11 | 89 |
| FIGURA 25: Mapa hierárquico dos grupos e subgrupos funcionais anotados (A) compartilhados entre as estirpes LGMAI_St_08 e LGMAI_St_14 e (B) compartilhados entre as estirpes LGMAI_St_11 e LGMAI_St_14..... | 90 |
| FIGURA 26: Mapa hierárquico dos grupos e subgrupos funcionais anotados que são compartilhados pelas três estirpes LGMAI..... | 91 |
| FIGURA 27: Mapa genômico circular contendo os 3 genomas das estirpes LGMAI e outros 15 genomas de <i>Streptococcus agalactiae</i> | 92 |
| FIGURA 28: Mapa circular do genoma da estirpe LGMAI_St_08..... | 93 |
| FIGURA 29: Mapa circular do genoma da estirpe LGMAI_St_11..... | 94 |
| FIGURA 30: Mapa circular do genoma da estirpe LGMAI_St_14..... | 95 |

LISTA DE QUADROS

| | |
|--|----|
| QUADRO 1: Peptídeos antimicrobianos de origem bacteriana em testes clínicos | 24 |
| QUADRO 2: Exemplos de bacteriocinas isoladas de alimentos processados ou usadas na biopreservação dos mesmos..... | 33 |
| QUADRO 3: Características e agrupamentos das principais bactérias pertencentes ao gênero <i>Streptococcus</i> , notórios causadores de infecções em seres humanos e animais ... | 34 |
| QUADRO 4: Estirpes de <i>Streptococcus agalactiae</i> LGMAI | 39 |
| QUADRO 5: Painel de bactérias indicadoras usadas por Silva (2003)..... | 40 |
| QUADRO 6: Estirpes bacterianas produtoras | 45 |
| QUADRO 7: Estirpes bacterianas indicadoras | 46 |
| QUADRO 8: Genomas usados nas análises filogenéticas e seus respectivos números de acesso no NCBI..... | 50 |
| QUADRO 9: Detalhamento das categorias funcionais COG usadas pela anotação da ferramenta <i>eggNOG-mapper</i> | 56 |
| QUADRO 10: Espectro de ação das estirpes produtoras de substância antimicrobiana | 59 |

LISTA DE TABELAS

| | |
|--|----|
| TABELA 1: História da descoberta das principais classes de antibióticos | 18 |
| TABELA 2: Exemplos mais representativos de peptídeos antimicrobianos | 22 |
| TABELA 3: Países em que a nisina é usada e suas áreas de aplicação | 27 |
| TABELA 4: Pipeline para sorotipagem WGS de <i>Streptococcus</i> grupo B com sequências marcadoras de referência e os requisitos de identidade necessários para os 10 sorotipos capsulares | 36 |
| TABELA 5: Estimativa de resistência antimicrobiana para infecções por <i>Streptococcus agalactiae</i> no ano de 2019 | 38 |
| TABELA 6: Entradas das sequências proteicas usadas no alinhamento do UniProt | 54 |
| TABELA 7: Genomas completos de <i>Streptococcus agalactiae</i> usados na comparação contra as estirpes LGMAI por meio do software BRIG | 58 |
| TABELA 8: Estatísticas dos três genomas montados | 60 |
| TABELA 9: Recorte dos resultados da análise de ANI, contendo as 5 melhores identidades.... | 64 |
| TABELA 10: Genes de resistência antimicrobiana identificados pela ferramenta ABRicate nos bancos de dados CARD e Resfinder | 69 |
| TABELA 11: Recorte dos resultados da análise de busca por fatores de virulência no banco de dados VFDB..... | 70 |
| TABELA 12: Recorte dos resultados obtidos com o programa Gipsy | 71 |
| TABELA 13: Resultados da análise de busca por sequências de fagos realizada com o servidor Phaster..... | 74 |
| TABELA 14: Resultados dos clusters de metabólitos secundários obtidos na análise com a ferramenta antiSMASH | 76 |
| TABELA 15: Resultados gerais da busca por clusters de peptídeos antimicrobiano, realizada com a ferramenta BAGEL4 | 77 |
| TABELA 16: Recorte dos resultados da anotação de clusters gênicos de peptídeos antimicrobianos, realizada na ferramenta BAGEL4 | 77 |
| TABELA 17: Resultados da análise com o banco de dados de bacteriocinas BACTIBASE | 77 |
| TABELA 18: Resultados da análise de busca por domínios e região de peptídeo sinal identificados com o Interpro..... | 78 |

| | |
|--|----|
| TABELA 19: Resultados do <i>blastx</i> entre os genomas e a sequência proteica do fator de imunidade zif de <i>Streptococcus equi</i> subsp. <i>zooepidemicus</i> | 81 |
| TABELA 20: Regiões de superfamília e família proteicas, sítio de ligação com braço de tRNA e estrutura de α -hélice identificadas na sequência do fator de imunidade putativo zif pelo banco de dados Interpro | 82 |
| TABELA 21: Ocupação dos grupos de ortólogos identificados pelo eggNOG-mapper nas 4 grandes categorias | 82 |
| TABELA 22: Ocupação dos grupos de ortólogos KEGG identificados pelo servidor KAAS nas 6 grandes categorias | 85 |

SUMÁRIO

| | |
|---|----|
| 1 INTRODUÇÃO | 14 |
| 1.1 ERAS PRÉ E PÓS-ANTIBIÓTICOS | 14 |
| 1.2 PEPTÍDEOS ANTIMICROBIANOS | 20 |
| 1.3 BACTERIOCINAS | 25 |
| 1.3.1 Classificação | 29 |
| 1.3.2 Mecanismos de ação | 30 |
| 1.4 BACTÉRIAS ÁCIDO-LÁCTICAS | 32 |
| 1.4.1 O gênero <i>Streptococcus</i> | 33 |
| 1.4.1.1 <i>Streptococcus agalactiae</i> | 36 |
| 1.4.1.2 Histórico das estirpes de <i>Streptococcus agalactiae</i> do LGMAI..... | 39 |
| 1.5 O POTENCIAL DA GENÔMICA..... | 40 |
| 1.6 JUSTIFICATIVA | 43 |
| 2 OBJETIVOS | 44 |
| 2.1 OBJETIVO GERAL..... | 44 |
| 2.2 OBJETIVOS ESPECÍFICOS | 44 |
| 3 METODOLOGIA | 45 |
| 3.1 ESTIRPES BACTERIANAS, MEIOS DE CULTURA E CONDIÇÕES DE CRESCIMENTO..... | 45 |
| 3.1.1 Estirpes bacterianas | 45 |
| 3.1.2 Meios de cultura | 46 |
| 3.1.3 Condições de crescimento | 46 |
| 3.2 TESTES DE DETECÇÃO DE ATIVIDADE ANTIMICROBIANA..... | 47 |
| 3.3 EXTRAÇÃO DE DNA GENÔMICO, SEQUENCIAMENTO, MONTAGEM DE GENOMAS E ANÁLISES DE BIOINFORMÁTICA..... | 48 |

| | |
|--|------------|
| 3.3.1 Extração de DNA e sequenciamento de genomas..... | 48 |
| 3.3.2 Controle de qualidade e montagem dos genomas..... | 48 |
| 3.3.3 Controle de qualidade pós-montagens..... | 49 |
| 3.3.4 Identificação filogenética e sorotipagem <i>in silico</i> com dados de sequenciamento genômico das estirpes produtoras de antimicrobianos..... | 50 |
| 3.3.5 Busca por genes e clusteres gênicos de peptídeos antimicrobianos | 53 |
| 3.3.6 Análise dos genes e clusteres gênicos de peptídeos antimicrobianos | 53 |
| 3.3.7 Avaliação de fatores de virulência, genes de resistência a antibióticos, ilhas genômicas de patogenicidade e resistência e sequências de fagos | 55 |
| 3.3.8 Anotação funcional e genômica comparativa | 56 |
| 4 RESULTADOS | 59 |
| 4.1 TESTES DE DETECÇÃO DE ATIVIDADE ANTIMICROBIANA..... | 59 |
| 4.2 GENOMAS | 59 |
| 4.3 IDENTIFICAÇÃO FILOGENÉTICA E SOROTIPAGEM <i>IN SILICO</i> COM DADOS DE SEQUENCIAMENTO GENÔMICO DAS ESTIRPES PRODUTORAS DE ANTIMICROBIANOS..... | 62 |
| 4.4 AVALIAÇÃO DE FATORES DE VIRULÊNCIA, GENES DE RESISTÊNCIA A ANTIBIÓTICOS, ILHAS GENÔMICAS DE PATOGENICIDADE E RESISTÊNCIA E SEQUÊNCIAS DE FAGOS..... | 69 |
| 4.5 BUSCA POR GENES E CLUSTERES GÊNICOS ASSOCIADOS A PEPTÍDEOS ANTIMICROBIANOS..... | 76 |
| 4.6 ANÁLISE DOS GENES E CLUSTERES GÊNICOS ASSOCIADOS A PEPTÍDEOS ANTIMICROBIANOS..... | 78 |
| 4.7 ANOTAÇÃO FUNCIONAL E GENÔMICA COMPARATIVA | 82 |
| 5 DISCUSSÃO | 96 |
| 6 CONCLUSÃO..... | 110 |
| REFERÊNCIAS | 113 |

1 INTRODUÇÃO

1.1 ERAS PRÉ E PÓS-ANTIBIÓTICOS

Os antibióticos desempenham um papel muito importante no tratamento de doenças infecciosas e na profilaxia antimicrobiana de diversos procedimentos, como cirúrgicos e na pecuária. No entanto, não muito tempo atrás, antes do advento dos mesmos, a qualidade e a expectativa de vida das populações eram severamente ameaçadas pelas infecções por microrganismos e suas complicações potenciais. Por volta de 1900, por exemplo, a maioria dos cidadãos estadunidenses vivia somente até os quarenta anos, em decorrência de doenças infecciosas, como cólera, febre tifóide e tuberculose (Streicher, 2021).

Alexander Fleming descobriu, no ano de 1928, a primeira substância antimicrobiana, a qual posteriormente veio a ser conhecida como penicilina. Essa substância revolucionou o campo de tratamento de doenças infecciosas, uma vez que, após sua introdução no mercado no ano de 1940, as taxas de mortalidade causada por sífilis reduziram drasticamente. O sucesso da penicilina estimulou o interesse pela busca de outros compostos semelhantes, o que resultou na descoberta de substâncias como as sulfonamidas, as quais contribuíram para a redução de mortalidade por escarlatina, e a estreptomicina, responsável pela grande queda na mortalidade por tuberculose. Esse período ficou conhecido como a era de ouro dos antibióticos (Aminov, 2017).

Infelizmente, com a mesma velocidade que novos antibióticos foram descobertos, bactérias resistentes a eles também foram detectadas. Esse fenômeno acontece, porque grande parte dos antibióticos está presente na natureza há anos, participando da relação ecológica de competição entre microrganismos e seus competidores. Essa relação ecológica ocorre de modo que os microrganismos evoluem novas estratégias para inibir seus competidores, ao passo que seus competidores evoluem novas formas de anular a efetividade de tais estratégias (Streicher, 2021). Sendo assim, a maioria dos antibióticos originais, assim como suas variantes produzidas pela indústria farmacêutica, perderam efetividade de uso no tratamento de doenças infecciosas nos últimos anos (figura 1).

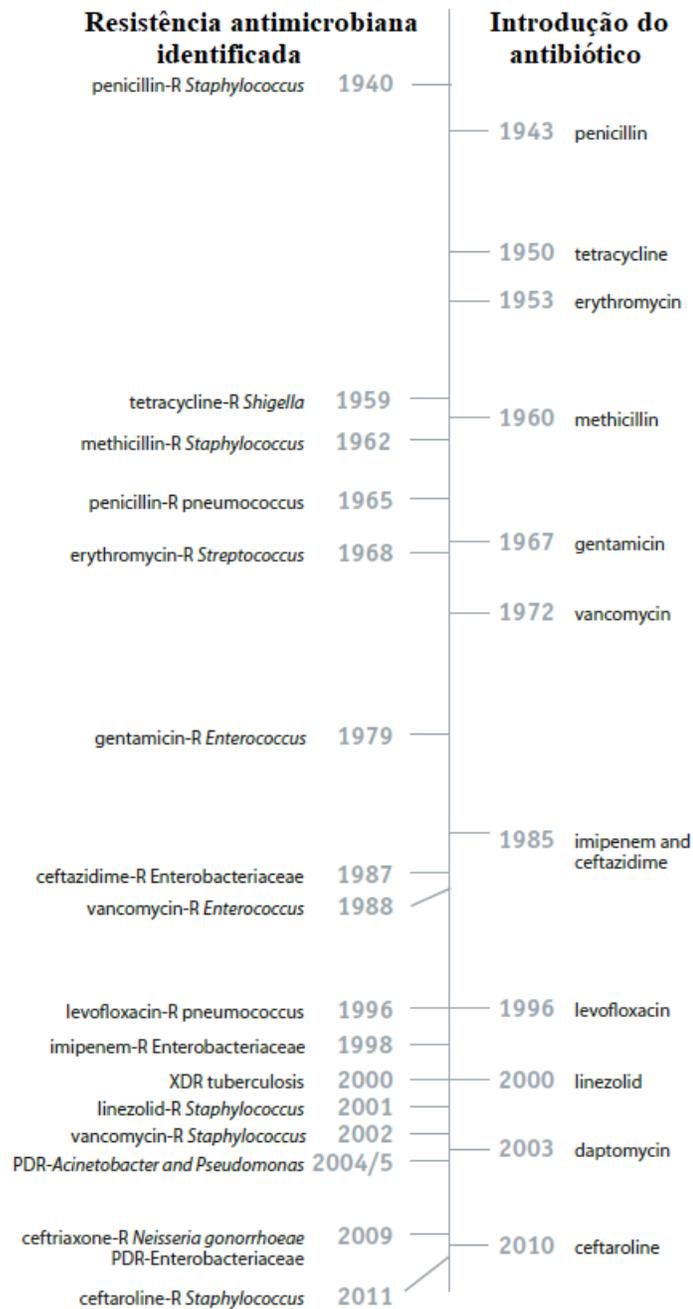


Figura 1: Linha do tempo da introdução de novos antibióticos no mercado e notificação dos principais eventos de resistência antimicrobiana. As datas são baseadas nas primeiras notificações de resistência na literatura. Nos casos de *Acinetobacter* e *Pseudomonas* resistentes a pan-drogas, as datas são baseadas em transmissões e surtos em clínicas e hospitais. Obs: penicilina estava em uso limitado antes do uso difundido na população em 1943.

Fonte: CDC, 2013 (adaptado).

Nas últimas décadas, o registro do número de patógenos bacterianos com aquisição de resistência a um ou múltiplos agentes antimicrobianos aumentou rapidamente, seja por abuso ou uso indevido de antibióticos administrados em pacientes e na pecuária ou pelo descarte indiscriminado no meio ambiente (Roca *et al.*, 2015). Segundo relatório emitido em 2018 pela Organização Mundial da Saúde (WHO, 2018), intitulado *Antimicrobial resistance and primary health care*, sobre o tema de resistência antimicrobiana e cuidados primários à saúde, o consumo de antibióticos em escala global aumentou em 65% entre os anos de 2000 e 2015. Para países de baixa e média renda, os dados são ainda mais alarmantes, já que a estimativa é de que o consumo de antibióticos tenha dobrado nesse mesmo período (Laxminarayan *et al.*, 2020).

Seis anos após a comissão de *Lancet Infectious Diseases* de 2013, a qual abordou as principais implicações do aumento da resistência antimicrobiana em um nível global, a preocupação com o problema subiu em prioridade nas agendas de saúde pública mundiais. Tal fenômeno resultou em progressos no monitoramento do surgimento de estirpes resistentes, ações nacionais e iniciativas globais, como o banimento de combinações irracionais de drogas na Índia e a cessação do uso de colistina como fator de crescimento de animais, e também trazendo o reconhecimento da importância de estudos de *One Health* para enfrentar o problema mundial de resistência antimicrobiana e seus desdobramentos (Laxminarayan *et al.*, 2020). Não obstante, o progresso cresceu de forma inconsistente globalmente, com uma tendência a aumento dos níveis de resistência antimicrobiana em países de baixa e média renda e uma leve estagnação de tais níveis em países de alta renda (figura 2).

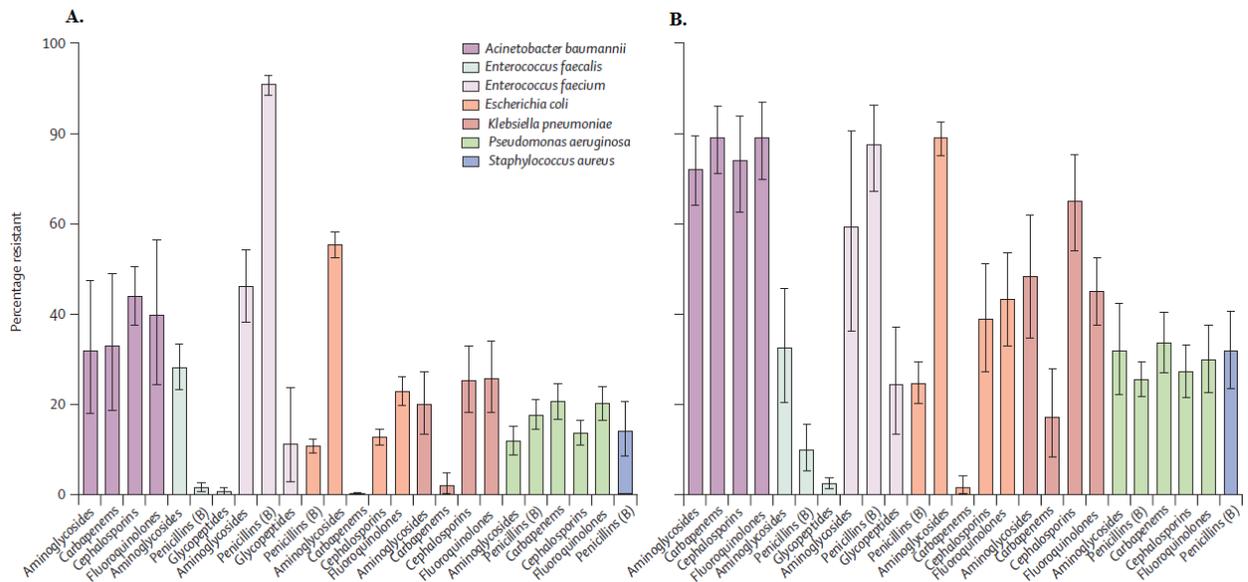


Figura 2: Gráfico de barras da distribuição de resistência antimicrobiana global, baseada em renda. Média ponderada das taxas de resistência antimicrobiana global para as diferentes combinações de antibióticos-patógenos em (A) países de alta renda e (B) países de baixa e média renda. O gráfico usa os dados mais recentes de 2012-2015 e baseia-se nas taxas de resistência antimicrobiana mundiais para os organismos causadores de doenças considerados como patógenos prioritários pela Organização Mundial de Saúde.

Fonte: Laxminarayan *et al.*, 2020 (adaptado).

A Agência Nacional de Vigilância Sanitária (Anvisa), juntamente à Organização Pan-Americana da Saúde (OPAS/OMS) e à Coordenação Geral dos Laboratórios de Saúde Pública (CGLAB/SVS/MS) criou, em 2005, a Rede Nacional de Monitoramento da Resistência Microbiana em Serviços de Saúde (Rede RM). A Rede RM tem o objetivo de melhorar a efetividade dos serviços de saúde por meio do monitoramento e do controle mais apropriado da emergência de resistência antimicrobiana (Anvisa, 2014).

De acordo com o boletim informativo de número 9 sobre Segurança do Paciente e Qualidade em Serviços de Saúde, publicado em dezembro de 2014 pela Rede RM, observou-se que, dentre os cocos Gram positivos mais frequentemente identificados como agentes etiológicos de infecções primárias em pacientes adultos internados em unidades de terapia intensiva no ano de 2013, as taxas de resistência à oxacilina foram de 78% em *Staphylococcus coagulase* negativo e de 60,9% em *Staphylococcus aureus*. Já em relação aos bacilos Gram negativos isolados nas mesmas condições deste levantamento, observou-se que as taxas de resistência aos carbapenêmicos foram de 80,7% em *Acinetobacter* spp., 37,4% em *Pseudomonas aeruginosa* e 33% em *Klebsiella pneumoniae*. Além disso, as taxas de resistência às cefalosporinas de amplo espectro foram de

31,7% em *K. pneumoniae*, 29,9% em *Serratia* spp., 27,4% em *Escherichia coli* e finalmente 27,2% em *Enterobacter* spp..

Para piorar ainda mais o problema da resistência antimicrobiana, as principais classes de antimicrobianos foram descobertas na era de ouro dos antibióticos (tabela 1). Desde então, novas classes não foram encontradas e os esforços concentraram-se em modificações das drogas já existentes e em síntese de drogas sintéticas (Aminov, 2017).

Tabela 1: História da descoberta das principais classes de antibióticos.

| Classes | Ano |
|---------------------|------|
| Arsfenaminas | 1910 |
| β -lactâmicos | 1929 |
| Sulfonamidas | 1935 |
| Polipeptídeos | 1939 |
| Aminoglicosídeos | 1943 |
| Tetraciclina | 1945 |
| Anfenicóis | 1947 |
| Lipopeptídeos | 1947 |
| Macrolídeos | 1950 |
| Oxazolidinonas | 1952 |
| Glicopeptídeos | 1953 |
| Streptograminas | 1953 |
| Ansamicinas | 1957 |
| Quinolonas | 1962 |
| Lincosamidas | 1963 |

Fonte: Aminov, 2017.

Sendo assim, a resistência antimicrobiana é um dos maiores desafios de saúde global da atualidade (O'Neill, 2014). O estudo estatístico liderado por Jim O'Neill na revisão sobre resistência antimicrobiana de 2014, mostrou que, se as taxas de resistência antimicrobiana continuarem a aumentar, a expectativa para o ano de 2050 é de que 10 milhões de pessoas morram por ano em decorrência de infecções microbianas não tratáveis (figura 3).

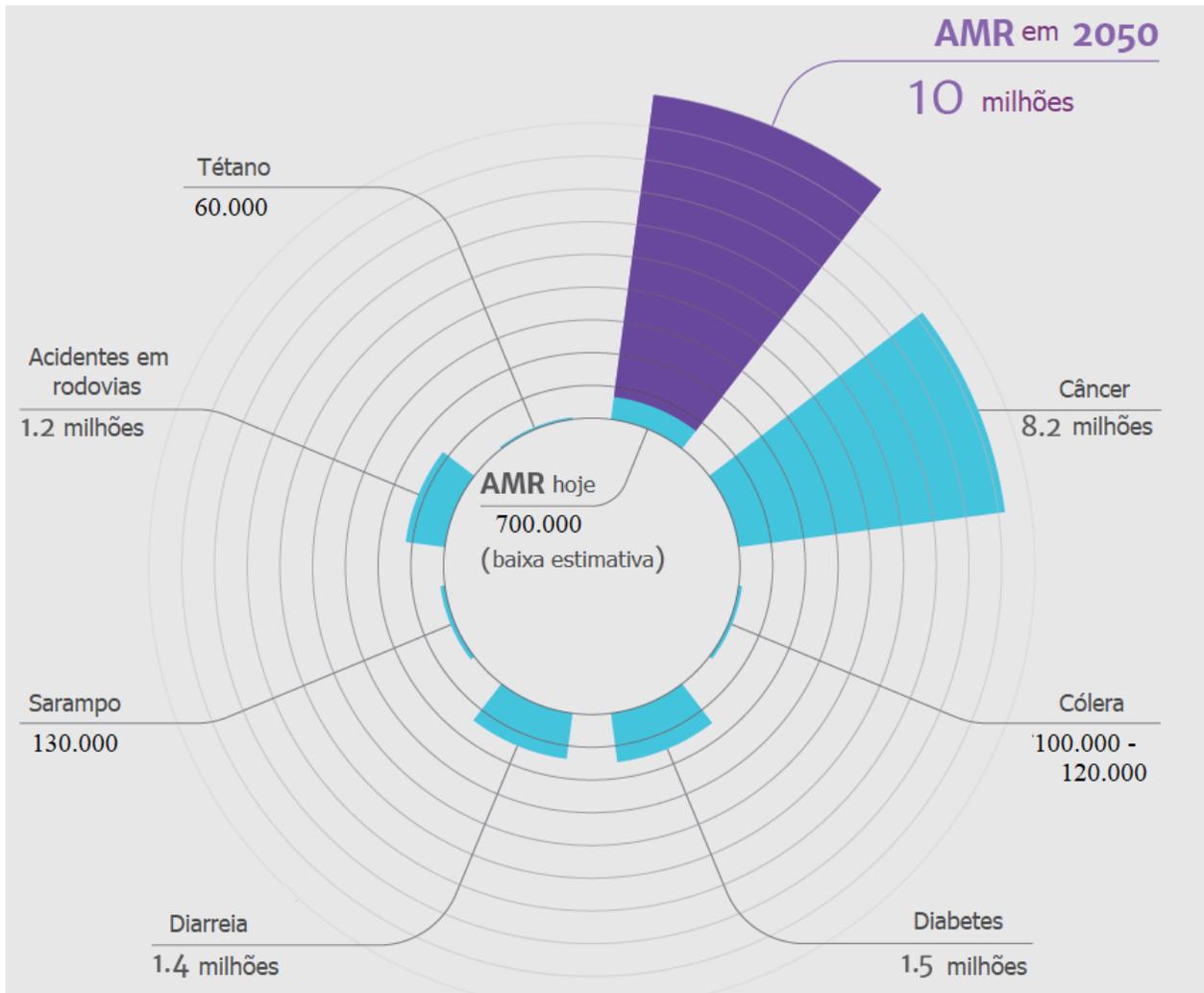


Figura 3: Morte em decorrência de resistência antimicrobiana por ano, comparada a outras grandes causas de morte. Em azul claro, dados atuais. Em roxo, estimativa para 2050. As estimativas de morte por AMR hoje, no estudo, são consideradas baixas, pois representam apenas uma parcela das bactérias com altos níveis de resistência (*Klebsiella pneumoniae*, *Escherichia coli* e *Staphylococcus aureus*).

Fonte: Adaptado de O'Neill, 2014.

Urge, desse modo, a busca por formas de contornar o problema de resistência antimicrobiana atualmente. Dentre as alternativas, as mais promissoras e que mais têm chamado atenção dos pesquisadores são aquelas que buscam soluções a longo prazo, como é o caso da terapia com fagos, oligonucleotídeos antimicrobianos e anticorpos monoclonais, as quais vislumbram possibilidades de tratamento para um futuro, no qual antibióticos são completamente ineficazes. Outras áreas de pesquisa também importantes envolvem o estudo de probióticos, imunomodulação, peptídeos antimicrobianos, entre outros (Streicher, 2021).

Além da busca tradicional por novas alternativas, as estratégias devem fazer uso da grande área das multiômicas da bioinformática, tanto para a identificação de possíveis novos antimicrobianos, como para o monitoramento e prevenção de resistência antimicrobiana nas populações e em estudos de *One Health*. Por sua vez, *One Health* representa um trabalho multisetorial que considera a perspectiva de profissionais de áreas distintas do conhecimento científico, visando cobrir todas as vertentes (humana, animal e ambiental) relacionadas a um problema em comum (McEwen; Collignon, 2017), como a resistência antimicrobiana. Sendo assim, os estudos de multiômicas trazem grande variedade de dados, permitindo exploração de novos domínios, como a microbiota intestinal e, em particular, a diversidade de antimicrobianos, como as bacteriocinas (peptídeos antimicrobianos de bactérias) lá presentes. A descoberta de novas bacteriocinas é um bom exemplo para descoberta de novas drogas fazendo uso de estudos de multiômicas (Chernov *et al.*, 2019).

Com base no elevado número de casos de bactérias multirresistentes e a escassez de novas famílias de antibióticos encontradas pela indústria farmacêutica, é notória a necessidade de se pesquisar e desenvolver novas substâncias antimicrobianas que possam substituir o nicho previamente ocupado pelos antibióticos (Cotter; Ross; Hill, 2013). Uma classe de substâncias antimicrobianas que apresenta grande potencial de aplicação no controle de microrganismos nocivos são os peptídeos antimicrobianos (Ghosh *et al.*, 2019; Yang *et al.*, 2014; Cotter; Ross; Hill, 2013; Parada *et al.*, 2007).

1.2 PEPTÍDEOS ANTIMICROBIANOS

Peptídeos antimicrobianos são pequenas moléculas de caráter proteico, que são produzidas pelos mais diversos tipos de organismos, de eucariotos a procariotos, e caracterizam-se por apresentarem forte atividade inibitória contra bactérias, vírus e fungos (Lazzaro; Zasloff; Rolff, 2020). Essas moléculas têm papel importante no sistema imune inato dos organismos que as produzem, uma vez que atuam protegendo-os de infecções por microrganismos indesejados por meio de uma variada gama de mecanismos de ação, como pela ruptura de membrana plasmática, penetração intracelular e interferência em vias metabólicas e também por meio de imunomodulação (Koo; Seo, 2019).

Embora alguns peptídeos antimicrobianos naturais tenham sido descobertos na primeira metade do século XX, tais quais a gramicidina e a defensina, o grande interesse nessas substâncias aumentou com o surgimento do problema de resistência antimicrobiana global (Koo; Seo, 2019). A conjuntura da necessidade por novos antimicrobianos com as vantagens dos peptídeos antimicrobianos em relação aos antibióticos tradicionais tornou-se grande incentivo para a pesquisa dos peptídeos antimicrobianos (figura 4).

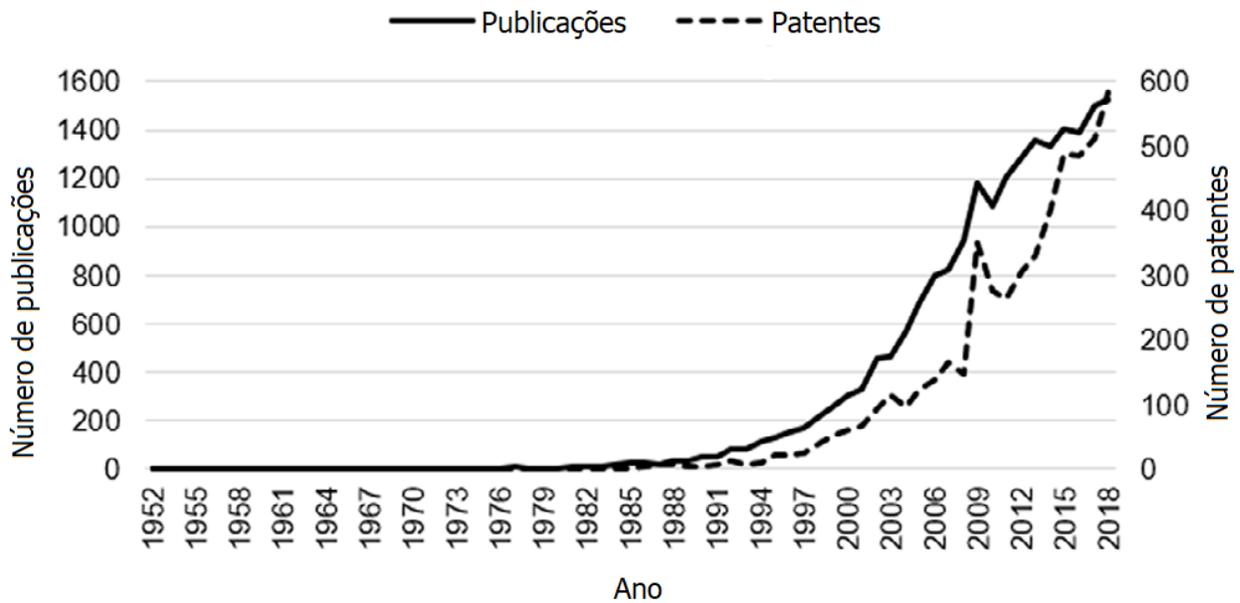


Figura 4: Publicações e patentes por ano para peptídeos antimicrobianos (SciFinder, fevereiro de 2019).

Fonte: Adaptado de Koo; Seo, 2019.

Com o aumento no estudo dessas substâncias como alternativas aos tradicionais antibióticos, peptídeos antimicrobianos aprovados em testes clínicos passaram, então, a ser utilizados no controle de microrganismos indesejados, tanto na área da conservação alimentícia, como na esfera clínica. Dentre os principais representantes de peptídeos antimicrobianos, 5 são de origem bacteriana: nisina A, gramicidina S, polimixina, daptomicina e teixobactina (tabela 2).

Tabela 2: Exemplos mais representativos de peptídeos antimicrobianos.

| Peptídeo | Descoberta | Peso molecular | Especificidade^a | Estrutura^b |
|-------------------------|-------------------|-----------------------|--|------------------------------|
| antimicrobiano | | (Da) | | |
| Nisina A | 1928 | 3.354,1 | <i>E. coli</i> , <i>S. aureus</i> | Policíclico (CP) |
| Gramicidina S | 1944 | 1.140,7 | <i>S. aureus</i> , <i>E. coli</i> | Policatiônico (CP) |
| Polimixina | 1947 | 1.203,5 | <i>K. pneumoniae</i> , <i>A. baumannii</i> | Policatiônico (CP) |
| Daptomicina | 1986 | 1.619,7 | MRSA, VRE | Lipopeptídeo (CP) |
| Teixobactina | 2015 | 1.242,5 | MRSA, VISA | CP |
| Melitina | 1967 | 2.846,5 | MRSA | α -hélice (CAP) |
| Magainina | 1987 | 2.409,9 | <i>E. coli</i> | α -hélice (CAP) |
| Catelicidina (LL-37) | 1996 | 4.493,3 | <i>E. cloacae</i> , <i>K. pneumoniae</i> , <i>P. aeruginosa</i> | α -hélice (CAP) |
| Buforina | 1996 | 2.434,9 | <i>E. coli</i> | α -hélice (CAP) |

^a MRSA – *Staphylococcus aureus* resistente à meticilina; VISA – *Staphylococcus aureus* de suscetibilidade intermediária à vancomicina; VRE – *Enterococci* resistente à vancomicina.

^b CAP – peptídeo linear anfipático catiônico; CP – peptídeo cíclico.

Fonte: Adaptado de Koo; Seo, 2019.

Grande parte dos peptídeos antimicrobianos interage com a superfície celular e não são diretamente mutagênicos, enquanto que muitos antibióticos, por outro lado, aumentam diretamente as taxas de mutação nas células bacterianas, uma vez que ativam respostas ao estresse, como as vias SOS (*Salt Overly-Sensitive*) e *rpoS*. Além disso, peptídeos antimicrobianos matam mais rapidamente, numa escala de minutos em vez de horas, de modo que existem muito menos gerações bacterianas, nas quais estirpes resistentes podem ser selecionadas. Não obstante, resistência a peptídeos antimicrobianos é um fenômeno já registrado e pode surgir em decorrência de fatores não específicos, como secreção de proteases e exopolímeros, formação de biofilmes e ativação de bombas de efluxo (Lazzaro; Zasloff; Rolff, 2020). Sendo assim, além da grande variedade de peptídeos antimicrobianos, a farmacodinâmica dos mesmos é uma de suas principais características responsável pela sua notória reduzida probabilidade de evolução de resistência (figura 5).

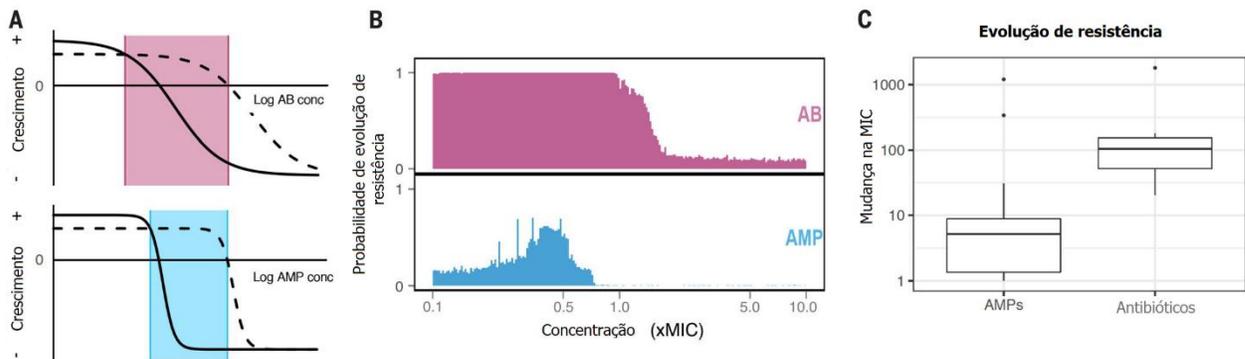


Figura 5: Evolução de resistência contra peptídeos antimicrobianos. A cor rosa representa antibióticos (AB), e a cor azul representa peptídeos antimicrobianos (AMP). Em (A), a linha sólida representa estirpes sensíveis, enquanto que a linha tracejada representa estirpes resistentes. O eixo horizontal representa os respectivos valores de concentração inibitória mínima (MIC), enquanto que as caixas coloridas representam o intervalo no qual mutantes resistentes podem ser selecionados. Em (B), simulação computacional prevendo mais baixa probabilidade de evolução de resistência em relação a concentrações de MIC para peptídeos antimicrobianos (azul). Em (C), experimentos de evolução de resistência em *E. coli* contra 15 peptídeos antimicrobianos com taxas de resistência mais baixas do que as do experimento contra 12 antibióticos.

Fonte: Adaptado de Lazzaro; Zasloff; Rolff, 2020.

Para se ter uma noção do avanço da pesquisa nessa área, em 2018, cerca de 70 peptídeos antimicrobianos estavam sendo testados como possíveis candidatos a drogas (Koo; Seo, 2019). Desses, 27 estavam em estágio de testes clínicos, 34 em testes pré-clínicos, 10 reprovados e 1 aprovado (daptomicina) pela agência federal do departamento de saúde e serviços humanos americano FDA (figura 6).

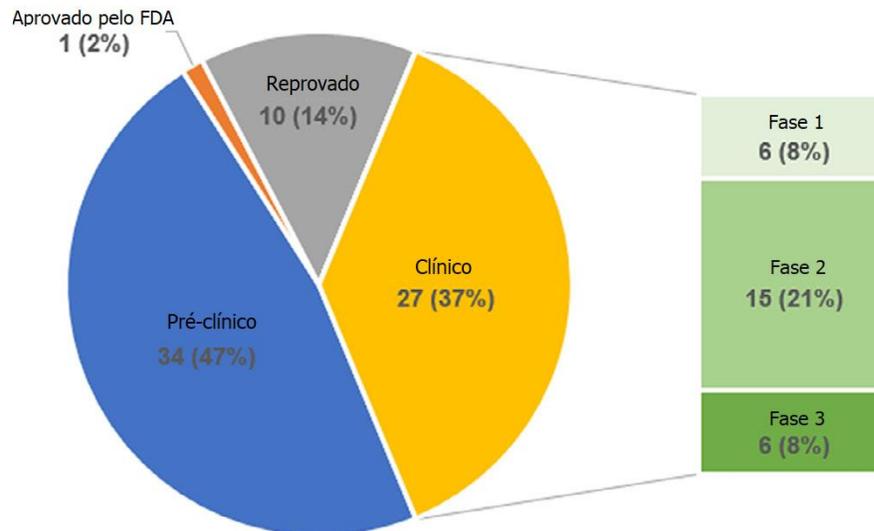


Figura 6: Status de desenvolvimento de peptídeos antimicrobianos (2018).

Fonte: Adaptado de Koo; Seo, 2019.

A maioria dos peptídeos antimicrobianos sendo testados é de origem humana, tendo em vista a baixa toxicidade que eles apresentam para aplicação no tratamento de infecções em humanos, sendo seguidos pelos peptídeos antimicrobianos sintéticos (figura 7).

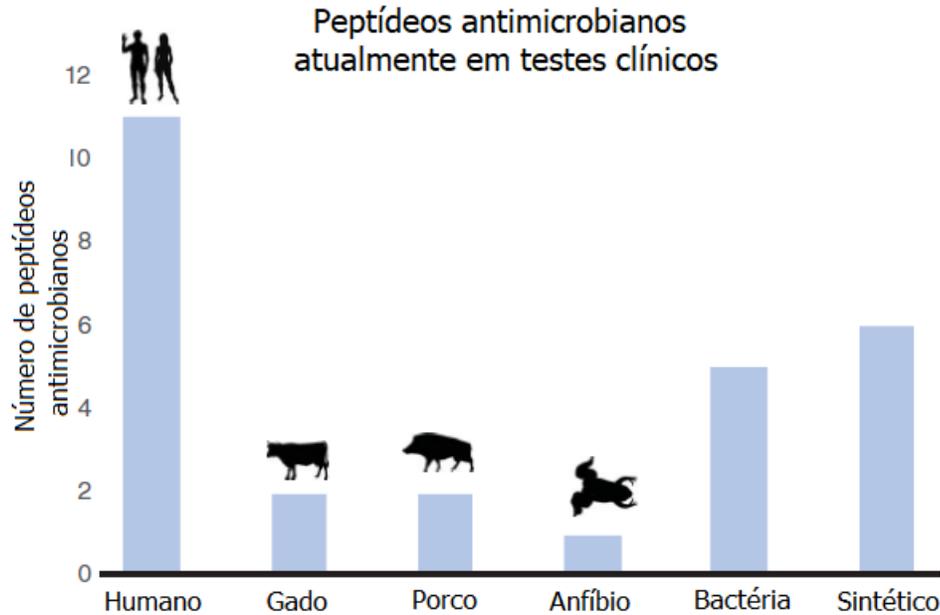


Figura 7: Peptídeos antimicrobianos atualmente em testes clínicos e os organismos dos quais eles são derivados.

Fonte: Adaptado de Lazzaro; Zasloff; Rolff, 2020.

Cinco dos peptídeos antimicrobianos que estavam sendo pesquisados como potenciais drogas eram de origem bacteriana (quadro 1).

Quadro 1: Peptídeos antimicrobianos de origem bacteriana em testes clínicos (fases I - III).

| Peptídeo antimicrobiano | Origem | Alvo | Fase | Administração |
|--------------------------|--------------------------|---|------|---------------|
| NVB-302 | Lantibiótico | <i>Clostridium difficile</i> | I | Oral |
| Wap-8294A2 (lotilibcina) | <i>Lysobacter</i> spp. | VRE, MRSA | I/II | Tópica |
| Surotomicina | Lipopeptídeo cíclico | <i>Clostridium difficile</i> (diarreia) | III | Oral |
| Ramoplanina (NTI-851) | <i>Actinoplanes</i> spp. | VRE, <i>Clostridium difficile</i> | III | Oral |
| Friulimicina B | Lipopeptídeo cíclico | Pneumonia, MRSA | I | Intravenosa |

Fonte: Adaptado de Koo; Seo, 2019.

1.3 BACTERIOCINAS

As bacteriocinas são uma classe de peptídeos antimicrobianos sintetizadas nos ribossomos de bactérias e que apresentam caráter anfipático e catiônico, sendo capazes de inibir o crescimento de outras bactérias, tanto Gram positivas, quanto Gram negativas (Kumariya *et al.*, 2019). As características anfipática e catiônica desses peptídeos permitem que eles interajam com as membranas citoplasmáticas das células bacterianas, causando, por exemplo, sua desintegração por meio da formação de poros (Bahar; Ren, 2013). Bacteriocinas produzidas por bactérias ácido-lácticas, apresentam um *status* de “geralmente reconhecidas como seguras” (do inglês *Generally Recognized as Safe*), visto que elas são peptídeos produzidos por bactérias naturalmente presentes em alimentos que consumimos, como queijos (Ogaki; Furlaneto; Maia, 2015).

Em relação aos tradicionais antibióticos, esses peptídeos antimicrobianos apresentam a vantagem de serem espécie-específicos, garantindo uma atuação pontual contra o microrganismo indesejado e evitando danos e pressão seletiva sobre bactérias benéficas. Além disso, embora já tenham sido reportados casos de resistência observados em laboratório, na prática, eles são escassos. A nisina, por exemplo, é uma bacteriocina usada na indústria de conservação alimentícia há muitos anos e ainda não foi reportado surgimento de resistência contra ela (Ghosh *et al.*, 2019).

Bacteriocinas também são produzidas pelas bactérias que compõem a microbiota humana e são imprescindíveis para as relações ecológicas de competição entre elas, assim como para a proteção do corpo humano contra bactérias maléficas e para o bem estar geral do organismo (figura 8). Sendo assim, a microbiota humana é uma das novas fronteiras de descoberta de novos antimicrobianos, também em decorrência da sua esperada baixa toxicidade (Mousa *et al.*, 2017).

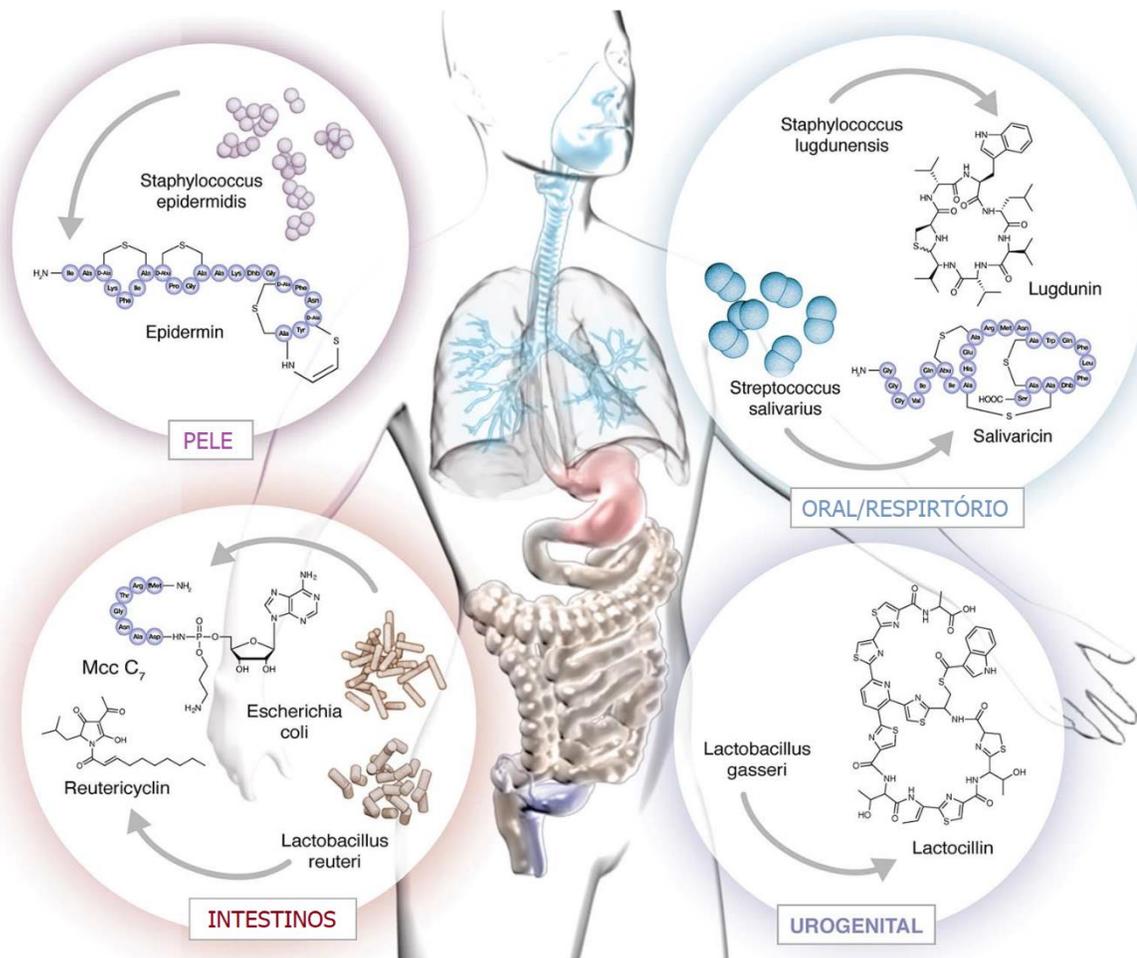


Figura 8: Exemplos diversos da microbiota produtora de antibióticos em diferentes sítios do corpo.

Fonte: Adaptado de Mousa *et al.*, 2017.

Historicamente, a colicina, isolada de *Escherichia coli*, foi a primeira bacteriocina descrita na literatura científica, identificada pelo cientista belga André Gratia em 1925, em seu trabalho intitulado “*Sur un remarquable exemple d’antagonisme entre deux souches de colibacille*” (em tradução livre: sobre um exemplo notável de antagonismo entre duas cepas de colibacilo), publicado em *Comptes Rendus de Séances et Mémoires de la Société de Biologie*, 93, 1040-1041 (Balciunas *et al.*, 2013). Essas substâncias atraíram a atenção dos pesquisadores da época, por causa de sua capacidade de interferir, de maneira bactericida ou bacteriostática, no crescimento de outras espécies bacterianas filogeneticamente próximas (Ogaki; Furlaneto; Maia, 2015).

Alguns anos depois, a busca por novas substâncias de mesmo caráter resultou a descoberta da nisina pelos pesquisadores L. A. Rogers e E. O. Whittier no ano de 1928. Essa substância foi isolada de *Lactococcus lactis* subsp. *lactis* e era capaz de inibir *Lactobacillus bulgaricus*. Ela é o exemplo de bacteriocina com maior sucesso de aplicação na indústria alimentícia, sendo a primeira

a ser utilizada como conservante alimentício a partir do ano de 1969 e atualmente implementada em mais de 50 países (Ogaki; Furlaneto; Maia, 2015; Balciunas *et al.*, 2013; tabela 3).

Tabela 3: Países em que a nisina é usada e suas áreas de aplicação.

| Região | Alimentos | Adição máxima |
|------------------|---|----------------------|
| Abu Dhabi | Leite pasteurizado, leite aromatizado, leite de longa vida, queijo processado, queijo, outros laticínios, vegetais enlatados | Sem limite |
| Argentina | Queijo processado | 500 ui/g |
| Austrália | Queijo processado, queijo, tomate em lata, massa de tomate em lata, sopas em lata | Sem limite |
| Bahrein | Leite pasteurizado, leite aromatizado, leite de longa vida, queijo processado, outros laticínios, vegetais enlatados | Sem limite |
| Bélgica | Queijo | 100 ui/g |
| Bolívia | Uso não proibido em alimentos | Sem limite |
| Bulgária | Queijo, gelo para armazenamento de peixe fresco | 200 mg/kg |
| Chile | Produtos alimentícios | 4.000 ui/g |
| Colômbia | Queijo processado | 500 ui/g |
| Costa Rica | Produtos de queijo | Sem limite |
| Chipre | Queijo, creme coagulado, vegetais enlatados | Sem limite |
| República Tcheca | Produtos de panificação e recheios, maionese, queijo processado, alimentos processados e semiprocessados, vegetais enlatados, alimentos para bebês, | 500 ui/g |
| Eslováquia | laticínios e vegetais | |
| Dubai | Queijo, leite pasteurizado, leite aromatizado, leite de longa vida, outros laticínios | Sem limite |
| Egito | Queijo processado | 500 ui/g |
| Irlanda | Queijo processado | 500 ui/g |
| União Europeia | Pudins de sêmola e tapioca e produtos semelhantes | 3 mg/kg |
| | Queijo maturado e queijo processado | 12,5 mg/kg |
| | Queijo mascarpone | 10 mg/kg |
| | Ovos pasteurizados | 6,25 mg/kg |
| Finlândia | Queijo processado | 480 ui/g |
| França | Queijo processado | Sem limite |
| Gibraltar | Queijo, comida enlatada, creme coagulado | Sem limite |
| Hong Kong | Queijo, comida enlatada, creme coagulado | Sem limite |
| Índia | Queijo, queijo processado | 1000 ui/g |
| Itália | Queijos, vegetais enlatados, cremes para confeitaria | 100 ui/g |

Continua

| Conclusão | | |
|-------------------|--|----------------------|
| Região | Alimentos | Adição máxima |
| Jordânia | Queijo processado | 500 ui/g |
| Kuwait | Queijo processado | 4000 ui/g |
| Malásia | Comida enlatada e queijo | Sem limite |
| Malta | Queijo, comida enlatada, creme coagulado | Sem limite |
| México | Nisina é um aditivo permitido | 500 ui/g |
| Países Baixos | Queijo de fábrica, quarg, certos queijos processados, queijo em pó | 500 ui/g |
| Nova Zelândia | Queijo processado | 500 ui/g |
| Peru | Nisina é um aditivo permitido | Sem limite |
| Filipinas | Queijo processado | 4.000 ui/g |
| Polônia | Queijo maturado | 500 ui/g |
| Catar | Leite, derivados de leite | Sem limite |
| Arábia Saudita | Alguns alimentos e laticínios | Sem limite |
| Singapura | Queijo, comidas enlatadas | Sem limite |
| África do Sul | Queijo processado, outros queijos | 500 ui/g |
| Espanha | Queijo processado | 500 ui/g |
| Suécia | Queijo e margarina, queijo com carne fresca, queijo processado | 500 ui/g |
| Taiwan | Queijo | 1.000 ui/g |
| Tailândia | Queijo processado | 4.000 ui/g |
| Trindade e Tobago | Comidas enlatadas, queijo, creme coagulado | Sem limite |
| Reino Unido | Queijo, comidas enlatadas, creme coagulado | Sem limite |
| EUA | Certos cremes pasteurizados de queijo processado | 10.000 ui/g |
| Uruguai | Queijo processado | 4.000 ui/g |
| Antiga URSS | Queijo processado dietético | 8.000 ui/g |
| | Produtos vegetais enlatados | 4.000 ui/g |

ui: unidade internacional de farmacologia

Fonte: Adaptado de Auffermann *et al.*, 2015.

A nisina é composta por uma cadeia de 34 resíduos de aminoácidos, totalizando um peso molecular de 3.500 Da, e apresenta atividade antimicrobiana de amplo espectro contra estirpes dos gêneros *Lactococcus*, *Streptococcus*, *Staphylococcus*, *Micrococcus*, *Pediococcus*, *Lactobacillus*, *Listeria* e *Mycobacterium*. Há duas variantes disponíveis desse composto: a nisina A que contém o aminoácido histidina na posição 27; e a nisina Z que contém o aminoácido asparagina na posição 27, o que confere as características comercialmente favoráveis de maior solubilidade e difusão a essa última (Arauz *et al.*, 2009).

1.3.1 Classificação

Desde sua primeira descrição por Klaenhammer (1993), a classificação das bacteriocinas sofreu diversas modificações, especialmente com a remoção da classe IV (Ogaki; Furlaneto; Maia, 2015). Atualmente, os autores chegaram a um consenso de que as bacteriocinas são classificadas em 3 classes (Kumariya *et al.*, 2019).

Em primeiro lugar, a classe I tem como principal exemplo a nisina comercial e é composta por peptídeos pequenos com menos de 28 resíduos de aminoácidos e peso molecular menor do que 5 Kda, os quais sofrem modificações pós-traducionais, resultando na presença de aminoácidos não convencionais em sua constituição, como a lantionina, a β -metil-lantionina, a desidrobutirina, a desidroalanina e a labirintina. Essa classe é ainda subdividida em: subclasse Ia, chamada de lantibióticos; subclasse Ib, conhecida por labirintopeptinas; e subclasse Ic, denominada sactibióticos (Kumariya *et al.*, 2019).

Em seguida, a classe II contém peptídeos pequenos, constituídos de 30 a 60 resíduos de aminoácidos e com peso molecular menor do que 10 KDa, termoestáveis e que não sofrem modificações pós-traducionais. O principal representante dessa classe é a pediocina. A classe II pode ser ainda subdividida em subclasse IIa, composta por bacteriocinas tipo-pediocina; IIb, constituída de dois peptídeos estruturalmente diferentes que atuam em conjunto; IIc, contendo bacteriocinas circulares; e IId, compreendendo peptídeos lineares, não-modificados e não tipo-pediocina (Kumariya *et al.*, 2019).

Por fim, a classe III representa os peptídeos grandes e termolábeis, composta por bacteriocinas de atividade bacteriolítica. Colicinas são os principais exemplos de bacteriocinas dessa classe. Outros peptídeos descritos são a helvecitina M, produzida por *Lactococcus crispatus*, helvecitina J, de *Lactococcus helveticus* e a enterolisina A, produzida por *Enterococcus faecalis* (Kumariya *et al.*, 2019). Os mecanismos de ação e subdivisão dessa classe, no entanto, são escassos na literatura (Hécharde; Sahl, 2002).

1.3.2 Mecanismos de ação

A maior parte das bacteriocinas atua permeabilizando a membrana citoplasmática das células bacterianas por meio da formação de poros (figura 9), desestabilizando, dessa forma, o balanço de cargas dessa estrutura e promovendo a dissipação da força próton motora dela com consequente esgotamento do ATP intracelular e eventual extravasamento do conteúdo intracelular para o meio extracelular (Ogaki; Furlaneto; Maia, 2015).

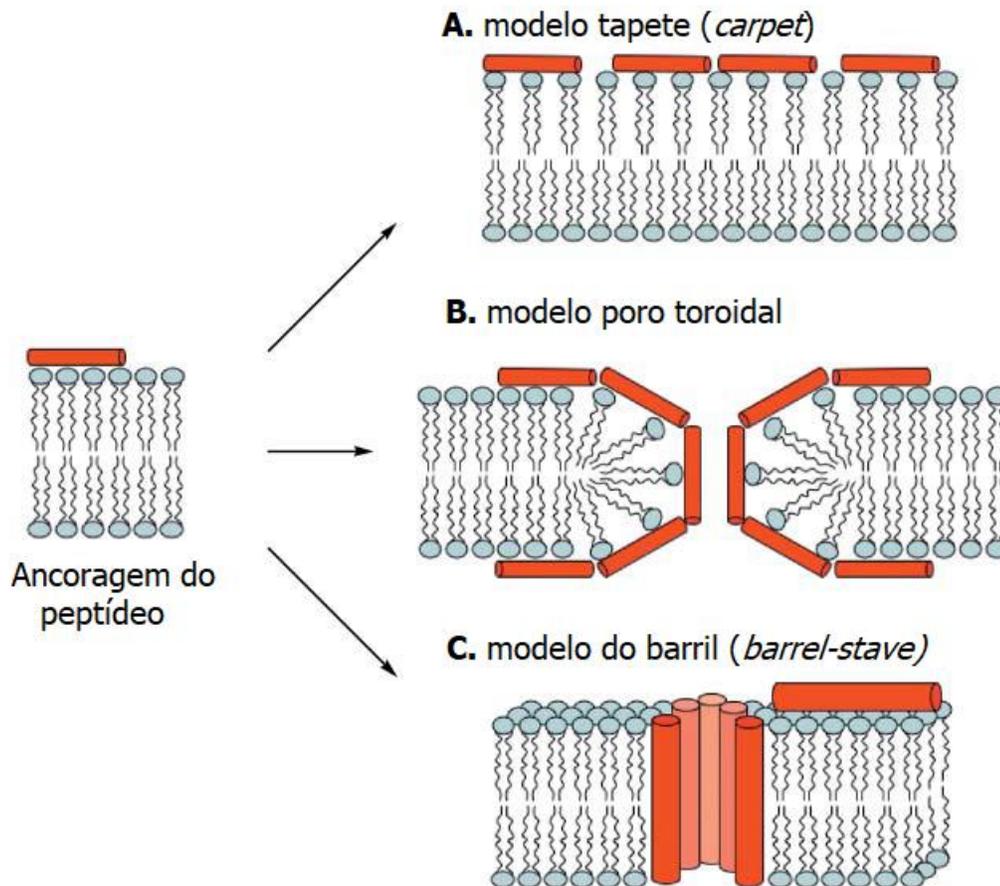


Figura 9: Mecanismos de formação de poros em bicamadas fosfolipídicas por meio da ação de bacteriocinas.

Fonte: Adaptado de Sanderson, 2005.

A nisina e a epidermina (essa última isolada de *Staphylococcus epidermidis*) atuam, primariamente, ancorando-se ao lipídeo II precursor do polissacarídeo peptidoglicano, garantindo direcionamento e precisão na formação de poros na membrana citoplasmática das células bacterianas e, simultaneamente, inibindo a biossíntese da parede celular dos procariontes (Ghosh *et al.*, 2019; Arauz *et al.*, 2009; Héchard; Sahl, 2002).

Similarmente, a lactococcina A, produzida por *Lactococcus lactis* subsp. *cremoris*, liga-se ao receptor de membrana manose-fosfotransferase e atua na formação de poros na membrana citoplasmática das células bacterianas sensíveis (Ogaki; Furlaneto; Maia, 2015).

Um dos mecanismos para formação de poros é chamado de modelo tapete, do inglês “*carpet*” (figura 9, A.). O modelo de tapete envolve a associação dos peptídeos paralelamente à membrana plasmática das células-alvo. A disposição desses peptídeos cobre a superfície da célula e acarreta uma reorganização dos lipídeos e afinamento da lamela externa da bicamada fosfolipídica, causando rompimento da estrutura com liberação de íons e alguns aminoácidos (Sanderson, 2005).

O segundo modelo é o poro toroidal (figura 9, B.). Segundo esse modelo, os peptídeos interagem paralelamente com a membrana plasmática. Essa interação gera uma penetração perpendicular superficial dos peptídeos na estrutura da membrana, interrompendo interações lipídio-lipídio e causando subsequente encurvadura da estrutura da membrana. Por fim, ao atingir uma proporção limite de peptídeo:lipídio, há completo rompimento da estrutura da membrana citoplasmática com alinhamento dos peptídeos em uma orientação transmembrana, formando um poro e permitindo livre passagem de íons e alguns aminoácidos (Bahar; Ren, 2013).

O que diferencia o mecanismo de tapete do mecanismo de poro toroidal é que, no primeiro, não há penetração na membrana plasmática, mas sim uma reorganização dos lipídeos, o que dificulta a manutenção da estrutura. No entanto, por vezes, é difícil distinguir entre esses dois mecanismos (Bertelsen *et al.*, 2012).

O terceiro modelo de formação de poros por peptídeos antimicrobianos é o do barril “*barrel-stave*” (figura 9, C.). Nesse modelo, peptídeos anfipáticos interagem eletrostática e paralelamente com a membrana plasmática, concentrando e agregando uma primeira camada das bacteriocinas na superfície da célula. Então, ao atingir um limiar de concentração, os peptídeos inserem-se verticalmente na membrana, de modo que domínios hidrofóbicos estejam voltados para a face externa e interajam com os lipídios da membrana, enquanto que os domínios hidrofílicos estejam voltados para a face interna, formando poros com estrutura de barris e preenchidos por água (Hécharde; Sahl, 2002).

Por outro lado, as bacteriocinas que atuam inibindo bactérias Gram negativas, precisam de ultrapassar a membrana externa da parede celular para, por fim, alcançar a membrana plasmática das células (Ogaki; Furlaneto; Maia, 2015). Sendo assim, algumas microcinas, produzidas por

Escherichia coli, são capazes de atravessar as barreiras celulares de bactérias Gram negativas e interferir no metabolismo de moléculas de DNA, RNA e proteínas (Ghosh *et al.*, 2019). O mecanismo de ação da microcina MccB17, por exemplo, é ligação com a enzima DNA-girase, produzindo quebras e degradação das moléculas de DNA; ao passo que a MccJ25 inibe a ação da RNA polimerase, uma vez que bloqueia o canal pelo qual os NTPs entram; e a MccC7-C51 é capaz de inibir a aspartil-tRNA sintetase, através da liberação de aspartil adenilato e consequente inibição competitiva pelo sítio ativo da enzima (Ogaki; Furlaneto; Maia, 2015).

1.4 BACTÉRIAS ÁCIDO-LÁCTICAS

No estudo de bacteriocinas, as bactérias ácido-lácticas recebem especial atenção e compõem um grupo de microrganismos que apresentam destaque na pesquisa de peptídeos antimicrobianos. Elas caracterizam-se pela produção de ácido láctico como um dos seus principais produtos de fermentação (Sieiro *et al.*, 2016). Essas bactérias destacam-se, também, pela produção de uma gama de metabólitos secundários de interesse biotecnológico e por apresentarem potencial uso como probióticos. Como resultado, a investigação de novos subprodutos desse grupo que possam trazer benefícios à saúde humana tem sido encorajada. Na indústria, essas bactérias têm um papel importante, tanto na produção de ácido láctico, quanto na produção de diacetil e outros ácidos orgânicos (Ortega *et al.*, 2019).

As bactérias ácido-lácticas formam um grupo de microrganismos Gram positivos e fermentativos, pertencentes ao filo Firmicutes. Os gêneros que compõem esse grupo são: *Aerococcus*, *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus* e *Weissella*. As bactérias produtoras de ácido láctico recebem atenção devido ao histórico de uso desses microrganismos em processos industriais de fermentação e também pelo status de “geralmente reconhecidas como seguras”, do inglês *Generally Recognized as Safe*, e pelo status de “presunção qualificada de segurança”, do inglês *Qualified Presumption of Safety* que algumas dessas bactérias possuem (Sieiro *et al.*, 2016). É desse grupo de bactérias que foram isoladas grande parte das bacteriocinas com aplicação na indústria alimentícia (Auffermann *et al.*, 2015; quadro 2).

Quadro 2: Exemplos de bactérias ácido-láticas produtoras de bacteriocinas usadas na indústria alimentícia.

| Produto | Bactéria produtora | Bacteriocina |
|---------------------|---|--|
| Cerveja | <i>Lactobacillus</i> spp. <i>Lactococcus lactis</i> <i>Lactobacillus sakei</i> <i>Lactococcus lactis</i> | B 7/3 Nisina Sakacina 5X; Sakacina 5T M30 |
| Vinho | <i>Lactobacillus plantarum</i> <i>Lactococcus lactis</i> | Plantaricina Nisina |
| Fermento | <i>Lactobacillus bavaricus</i> | Bavaricina A |
| Queijo | <i>Lactococcus lactis</i> <i>Enterococcus</i> spp. <i>Lactobacillus plantarum</i> <i>Lactococcus lactis</i> | Nisina Enterocina Pediocina AcH Lacticina 3147 |
| Queijo cremoso | <i>Lactobacillus plantarum</i> | Pediocina AcH |
| Iogurte | <i>Streptococcus thermophilus</i> <i>Lactobacillus acidophilus</i> | Thermophilina 347 Lactacina B10 |
| Kefir | <i>Lactococcus lactis</i> | Lactacina 3147 |
| Carne | <i>Pediococcus acidilactici</i> <i>Lactobacillus sake</i> <i>Carnobacterium piscicola</i> <i>Leuconostoc gelidum</i> | Pediocina AcH Sakacina A Piscicolina LV17 Leucocina A |
| Peixe | <i>Carnobacterium piscicola</i> <i>Carnobacterium divergens</i> <i>Lactobacillus sake</i> <i>Lactococcus lactis</i> | Piscicolina V1 Diversina V41 Sakacina P Nisina |
| Leites e laticínios | <i>Enterococcus</i> spp. | Enterocina AS-48 |
| Aves e suínos | <i>Lactobacillus sake</i> | Sakacina K |
| Carne bovina | <i>Lactobacillus casei</i> | Lactocina 705 |
| Salsichas bovinas | <i>Pediococcus acidilactici</i> | Pediocina AcH |

Fonte: Adaptado de Auffermann *et al.*, 2015.

1.4.1 O gênero *Streptococcus*

O gênero *Streptococcus*, pertencente ao grupo das bactérias ácido-láticas, é um gênero que passou por diversas revisões taxonômicas no passado e, atualmente, possui aproximadamente 72 espécies de bactérias (Richards *et al.*, 2014). Essas bactérias são Gram positivas, catalase negativa, em formato de cocos e organizadas em cadeias, possuindo representantes comensais,

patogênicos e patogênicos oportunistas, tanto para humanos, quanto para animais. Tradicionalmente, a taxonomia do gênero *Streptococcus* foca em características morfológicas, bioquímicas e sorológicas, baseando-se na classificação de Rebecca Lancefield com base nos anticorpos que reagem com os carboidratos da parede celular das diferentes espécies (Haenni; Lupo; Madec, 2018; quadro 3).

Quadro 3: Características e agrupamentos das principais bactérias pertencentes ao gênero *Streptococcus*, notórios causadores de infecções em seres humanos e animais.

| Espécie | Hemolisina | Grupo de Lancefield | Hospedeiros |
|---|---------------------|----------------------------|---|
| <i>S. agalactiae</i> | Beta | B | Humano, vaca, camelo, cavalo, golfinho, peixe |
| <i>S. dysgalactiae</i> subsp. <i>dysgalactiae</i> | Alfa | C (G/A/L) | Humano, vaca, peixe |
| <i>S. dysgalactiae</i> subsp. <i>equisimilis</i> | Beta | | Humano, cavalo, suíno, foca |
| <i>S. equi</i> subsp. <i>equi</i> | | | Cavalo |
| <i>S. equi</i> subsp. <i>zooepidemicus</i> | | | Humano, cavalo, ovelha |
| <i>S. phocae</i> | | | Peixe, foca |
| <i>S. gallolyticus</i> subsp. <i>gallolyticus</i> | Não hemolítico | D | Humano, ave, coala |
| <i>S. gallolyticus</i> subsp. <i>pasteurianus</i> | | | Humano, ave |
| <i>S. porcinus</i> | Beta | E (P/U/V) | Humano, suíno |
| <i>S. canis</i> | Beta | G | Humano, cão, gato |
| <i>S. suis</i> | Não hemolítico/Beta | R | Humano, suíno, javali, coelho |
| <i>S. iniae</i> | Beta | Indefinido | Humano, golfinho, peixe |
| <i>S. uberis/S. parauberis</i> | | | Vaca, cavalo |

Fonte: Adaptado de Haenni; Lupo; Madec, 2018.

No entanto, o agrupamento tradicional torna-se confuso, uma vez que diversos anticorpos reagem com isolados da mesma espécie. Sendo assim, aplicações mais recentes com abordagens químico-taxonômicas, hibridização DNA-DNA e sequenciamento do 16S rDNA trouxeram elucidações a respeito das relações entre os diversos representantes de streptococci. Essas revisões

revelaram um delineamento natural das espécies nos grupos: *anginosus*; *bovis*; *mitis*; *mutans*; *pyogenes*; *salivarius*; e *sanguinus*, acomodando a maior parte das espécies do gênero, embora algumas ainda permaneçam sem grupo (Haslam; Geme, 2018; Gao *et al.*, 2014).

Estudos de bacteriocinas em streptococci datam da década de 1960. O foco de estudos de bacteriocinas do gênero *Streptococcus* está, em sua maioria, nas estirpes patogênicas, especialmente *Streptococcus mutans*, cujas bacteriocinas recebem a nomenclatura de mutacinas. Streptococci apresentam um grande potencial de produção de bacteriocinas, com descrição de produção de bacteriocinas em *S. salivarius*, *S. mutans*, *S. pyogenes*, *S. macedonicus*, *S. bovis*, *S. uberis*, *S. rattus*, *S. thermophilus* e *S. agalactiae*. Algumas estirpes de *S. mutans* são capazes de produzir até 3 bacteriocinas diferentes: as mutacinas I, II, III, com mecanismos de ação variados, como formação de poros nas membranas, inibição da captação de aminoácidos e esgotamento das reservas de ATP intracelular. Grande parte das bacteriocinas isoladas de bactérias desse gênero são pertencentes à subclasse Ia dos lantibióticos (Nes; Diep; Holo, 2007).

Muitas bacteriocinas já foram descobertas/isoladas de espécies pertencentes ao gênero *Streptococcus*, dentre elas: a Bovicina HJ50 de *Streptococcus bovis* (Xiao *et al.*, 2004); Mutacinas B-Ny266, II e 1140/III de *Streptococcus mutans* (Meira *et al.*, 1997; Novák; Caufield; Miller, 1994; e Qi; Chen; Caufield, 1999); Streptococcinas A-FF22, A-M49 e A-M57, Estreptina, Salivaricina A, SpbMN, Blp1 e Blp2 de *Streptococcus pyogenes* (Vogel; Spellerberg, 2021); Salivaricinas A as, A3 e A9 de *Streptococcus salivarius* (Ross; Ronson; Tagg, 1993; Wescombe *et al.*, 2006; Wescombe *et al.*, 2011); Thermophilina 347 de *Streptococcus thermophilus* (Villani *et al.*, 1995); e Ubericina A e Uberolisina de *Streptococcus uberis* (Heng *et al.*, 2007; Wirawan *et al.*, 2007).

Com foco na espécie *Streptococcus agalactiae*, recentemente foi identificada a Agalactina na estirpe de *Streptococcus agalactiae* ATCC 13813 por meio de mineração com o banco de dados BAGEL3 (van Heel *et al.*, 2016). Além disso, foi relatada a produção de substância tipo-bacteriocina (*Bacteriocin-like substance* - BLS) em uma estirpe de *Streptococcus agalactiae* isolada de uma paciente com vaginite no Iraque (Razzaq; Sabri, 2005). Além desses casos, não há muita informação na literatura científica a respeito da produção de substâncias antimicrobianas e, mais especificamente, bacteriocinas na espécie de *Streptococcus agalactiae*.

1.4.1.1 *Streptococcus agalactiae*

Streptococcus agalactiae, ou *Streptococcus* grupo B, é um dos principais patógenos causadores da mastite bovina, acarretando grandes perdas econômicas no ramo da pecuária devido à conseqüente redução na quantidade e qualidade do leite produzido nas fazendas (Castro *et al.*, 2017). Em animais, frequentemente estão associadas a mastites do tipo mistas ou de necroses supurativas (Bianchi *et al.*, 2019). Essa bactéria pode ser beta ou gama hemolítica, além de possuir uma cápsula de polissacarídeos, a qual representa o primeiro fator de virulência bacteriano identificado a permitir evasão do sistema imune e invasão das células. Essa cápsula de polissacarídeos é tão importante, que a sorotipagem dessa bactéria é realizada com base no tipo capsular e, até hoje, foram identificados 10 diferentes tipos capsulares (Ia, Ib, II-IX) (Castro *et al.*, 2017).

Com os avanços das ciências ômicas, Metcalf e colaboradores (2017), do Centro de Controle e Prevenção de Doenças, realizaram um estudo verificando a eficiência da bioinformática em identificar os sorotipos de *Streptococcus agalactiae* por meio de marcadores dentro do lócus cps, o qual conclui que a metodologia é válida e ainda gera menos resultados de estirpes não-identificadas que os métodos tradicionais de sorotipagem. A sorotipagem com dados WGS proposta por Metcalf e colaboradores (2017) é uma análise baseada em *Blast*, na qual é comparado o lócus de polissacarídeo capsular (*cps*) das estirpes estudadas com regiões marcadoras para os 10 sorotipos de *Streptococcus* grupo B (tabela 4).

Tabela 4: Protocolo para sorotipagem WGS de *Streptococcus* grupo B com sequências marcadoras de referência e os requisitos de identidade necessários para os 10 sorotipos capsulares.

| Sorotipos GBS | Coordenadas GenBank | Requisitos | Comentários |
|---------------|----------------------|-----------------------|--|
| Ia | AB028896: 7731-7873 | 96% de identidade | |
| Ib | AB050723: 3010-3152 | 96% de identidade | |
| II | AY375362: 7412-7550 | 96% de identidade | |
| III | AF163833: 7631-7802 | 96% de identidade | Também presente em estirpes do sorotipo II |
| IV | AF355776: 7356-7577 | 96% de identidade | |
| V | AF349539: 9929-10099 | 96% de identidade | |
| VI | AF337958: 8045-8143 | 96% de identidade | |
| VII | AY376403: 5360-5519 | Correspondência exata | Correspondência imperfeita (~96% de identidade) com o sorotipo IX, observada em alguns casos |

Continua

| Conclusão | | | |
|---------------|---------------------|-----------------------|---|
| Sorotipos GBS | Coordenadas GenBank | Requisitos | Comentários |
| VIII | AY375363: 6020-6132 | 96% de identidade | |
| IX | GQ499301: 1-130 | Correspondência exata | Correspondência imperfeita (~96% de identidade) com o sorotipo VII, observada em alguns casos |

Fonte: Centro de Prevenção e Controle de Doenças; Metcalf e colaboradores (2017).

Streptococcus agalactiae é um patógeno capaz de infectar tanto humanos, quanto animais de forma independente (Haenni; Lupo; Madec, 2018). Em humanos, essas bactérias são encontradas colonizando a vagina e fazendo parte da microbiota intestinal. Elas podem causar pneumonia, síndrome do choque séptico e meningite (Castro *et al.*, 2017). *Streptococcus* grupo B é considerada um importante patógeno humano, uma vez que apresenta grande risco para mulheres em gestação e seus bebês, sendo uma das principais causas de morte por septicemia em recém-nascidos no Brasil e no mundo (Nascimento *et al.*, 2019; figura 10).

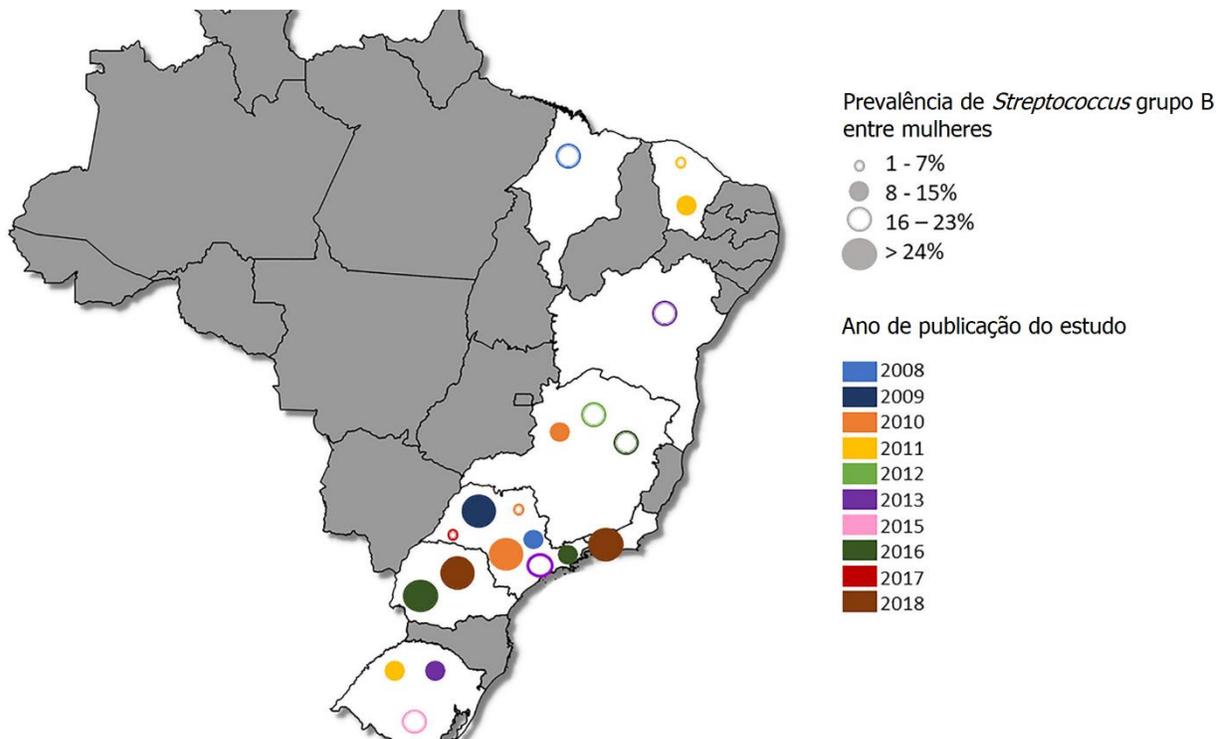


Figura 10: Distribuição geográfica da prevalência de *Streptococcus* grupo B em mulheres grávidas no Brasil.

Fonte: Adaptado de Nascimento *et al.*, 2019.

Os principais sorotipos detectados em infecções de mastite bovina são os sorotipos Ia, II, III e V; enquanto que, dentre os 10 sorotipos dessa espécie, os grandes responsáveis por doenças em seres humanos são os sorotipos Ia, Ib, II, III e V (Castro *et al.*, 2017).

Não somente responsáveis por infecções de grande importância clínica em humanos e animais, o relatório de 2019 de resistência antimicrobiana do CDC nos Estados Unidos mostrou que o número de infecções por *Streptococcus* grupo B resistentes à clindamicina têm aumentado e, embora a comparação entre os antigos dados de 2013 e os novos dados de 2019 não possa ser propriamente realizada, o patógeno ainda se mantém na categoria de “ameaças preocupantes” e precisa de ser monitorado nos próximos anos (tabela 5).

Tabela 5: Estimativa de resistência antimicrobiana para infecções por *Streptococcus agalactiae* no ano de 2019.

| Bactéria resistente | Estimativa de ameaça em 2019 | Contabilizado | Não contabilizado |
|--|--------------------------------|---------------------|---|
| <i>Streptococcus</i> grupo B resistente à clindamicina | 13.000 infecções 720 mortes | Infecções invasivas | Infecções não-invasivas e colonização intraparto assintomática dependente de profilaxia |

Fonte: Adaptado de CDC, 2019.

1.4.1.2 Histórico das estirpes de *Streptococcus agalactiae* do LGMAI

O Laboratório de Genética de Microrganismos Associados a Alimentos e a Indústria possui 54 estirpes de *Streptococcus agalactiae* isoladas de animais com mastite bovina de fazendas de Minas Gerais e São Paulo (quadro 4).

Quadro 4: Estirpes de *Streptococcus agalactiae* LGMAI.

| Estirpes | Origem |
|--|--------------------------|
| 3725, 3726, 3727, 3728, 3729, 3730, 3731, 3739, 3740, 3741, 3742, 3743, 3744, 3745, 3751, 3752 | Fazenda 1 – Minas Gerais |
| 3770, 3771, 3772, 3773, 3774, 3775, 3776 | Fazenda 2 – São Paulo |
| 3797, 3798, 3799, 3800, 3801, 3802, 3804 | Fazenda 3 – Minas Gerais |
| 3891, 3892, 3893, 3894, 3895, 3896, 3897, 3898, 3899 | Fazenda 4 – Minas Gerais |
| 129, 194, 835, 962, 1312, 1315, 3191, 3327, 3333, 3339, 3340, 3345, 3370, 3834, 3835 | N.A.* |

* A publicação do nome da fazenda e sua localização não foi autorizada.

Fonte: Silva, 2003.

Essas estirpes foram isoladas e fornecidas pela pesquisadora Maria Aparecida Vasconcelos Paiva Brito da Embrapa e fazem parte de um trabalho de colaboração com o projeto Embrapa Gado de Leite. A Embrapa Gado de Leite é uma das 42 Unidades Descentralizadas da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA) e tem como objetivo a realização de pesquisas para a pecuária leiteira de clima tropical (Embrapa).

Em seu trabalho no ano de 2003, intitulado “Caracterização de bacteriocinas produzidas por *Streptococcus agalactiae* isolados de mastite bovina”, o microbiologista Fabiano Murga da Silva realizou o *screening* por atividade antimicrobiana nas 54 estirpes de *Streptococcus agalactiae* LGMAI contra de 76 bactérias indicadoras (quadro 5).

Quadro 5: Pannel de bactérias indicadoras usadas por Silva (2003).

| Organismos | Estirpes | Origem |
|---------------------------------|---|---|
| <i>Corynebacterium fimi</i> * | N.A.** | Laboratório de Genética Molecular, UFRJ, Instituto de Microbiologia prof. Paulo de Góes |
| <i>Staphylococcus aureus</i> | RN451 | |
| <i>Streptococcus pyogenes</i> | St26 | Laboratório de Microbiologia de Alimentos, UFRJ, Instituto de Microbiologia prof. Paulo de Góes |
| <i>Listeria innocua</i> | N.A.** | |
| <i>Streptococcus agalactiae</i> | I14 | |
| <i>Lactococcus lactis</i> | N.A.** | |
| <i>Streptococcus porcinus</i> | 628, 662, 790, 857, 1058, 1124, 1217, 1451, 2378, 3123, 3176, 3658 e PR17 | CDC, Atlanta, Georgia, EUA, Dr. Richard Facklam |
| <i>Streptococcus pyogenes</i> | 1465, 1466, 1471, 1474, 1996, 2009, 2586, 2587, 2588, 2590, 2591, 2593, 2606, 2608, 2612, 2615, 2617, 2618, 2620 e 2625 | Laboratório de Apoio Biotecnológico, UFRJ, Instituto de Microbiologia prof. Paulo de Góes |
| <i>Streptococcus uberis</i> | 602, 752, 959, 2825, 3351, 3354, 3355, 3431, 3485, 3670, 3671 e 3724 | Maria Aparecida Vasconcelos Paiva Brito da Embrapa, Fazenda Gado de Leite, Juiz de Fora, Minas Gerais |
| <i>Streptococcus agalactiae</i> | 3731, 3739, 3725, 3726, 3751, 3752, 3743, 3728, 3742, 3727, 3729, 3730, 3740, 3744, 3745 e 3741 | Fazenda 1 – Minas Gerais |

* atualmente classificada como *Cellulomonas fimi*.

**Informação não disponível.

Fonte: Silva (2003).

1.5 O POTENCIAL DA GENÔMICA

Tradicionalmente, a descrição de novas bacteriocinas emerge de um trabalho de pesquisa que objetiva a identificação de estirpes produtoras de substâncias antimicrobianas capazes de inibir bactérias patogênicas de interesse, através de uma triagem por atividade antimicrobiana, seguida de etapas de purificação e caracterização bioquímica do peptídeo e de seus determinantes genéticos (Balciunas *et al.*, 2013). Estratégias de purificação adequadas para esses peptídeos são necessárias, tendo em vista que a caracterização dessas substâncias demanda um isolamento puro e homogêneo do peptídeo. No entanto, não existe um protocolo universal de purificação de bacteriocinas, em decorrência de uma ampla diversidade de naturezas bioquímicas dessas substâncias (Pingitore *et al.*, 2007). Algumas metodologias de purificação foram desenvolvidas e descritas na literatura, das quais, a mais convencional é baseada em precipitação proteica, seguida de cromatografia de troca-iônica ou gel filtração e HPLC de fase reversa (Vuyst; Leroy, 2007).

A descoberta de novos produtos naturais dá-se, como dito anteriormente, por meio de fracionamentos de extratos ativos de caldos ou por meio de triagens químicas com uso de técnicas de cromatografia e espectrometria. No entanto, as técnicas de sequenciamento de genomas em associação ao uso de ferramentas de bioinformática permitem, atualmente, a mineração *in silico* em massa dos genomas de bactérias e a subsequente descoberta de potenciais genes de novos produtos naturais com aplicação na saúde humana e animal, na proteção de cultivares na agricultura, além de diversas outras substâncias com aplicações biotecnológicas, como as bacteriocinas (Bachmann; Lanen; Baltz, 2014; Vélasquez; Donk, 2011).

As tecnologias de sequenciamentos de nova geração correspondem a plataformas que realizam sequenciamento massivo e em paralelo de milhões de pequenos fragmentos de DNA (Behjati; Tarpey, 2013). Nos últimos anos, essas tecnologias passaram por diversas modificações e melhorias, de modo que se tornaram o estopim da era genômica, caracterizada pela grande quantidade de dados e um vasto espectro de áreas de pesquisa abrangidas. Esse rápido progresso trouxe uma maior capacidade de processamento e acurácia, assim como menores custos, acarretando o sequenciamento de inúmeros organismos e uma inundação de dados genéticos disponíveis (Liu *et al.*, 2012).

Sequenciamento genômico da microbiota intestinal, por exemplo, permite a identificação dos microrganismos que a constituem a nível de espécie, às vezes chegando até em identificação de estirpes, além de contribuir na identificação de genes relacionados à resistência antimicrobiana, auxiliando no monitoramento da evolução de resistência a antibióticos nas diversas espécies, como também de gene para a produção de vitaminas, ácidos graxos de cadeia curta e de toda uma diversidade de peptídeos antimicrobianos bacterianos (bacteriocinas), com potencial aplicação no nicho dos antibióticos (O'Connor *et al.*, 2020; Chernov *et al.*, 2019).

O surgimento e melhoramento das ciências ômicas promete melhorias na forma como infecções e novos tipos de resistência antimicrobiana são diagnosticadas, abrindo um mar de possibilidades e facilidades, no qual se é possível a resistência antimicrobiana mais rapidamente e com maior eficiência (O'Neill, 2014).

Um dos grandes objetivos das ciências ômicas é o melhor entendimento da biologia por trás dos patógenos, almejando utilizar esse conhecimento no desenvolvimento de estratégias para o descobrimento de novas drogas. Nesse cenário, as ômicas integradas oferecem diversas ferramentas com esses objetivos (Chernov *et al.*, 2019; figura 11).

A era das ciências ômicas traz uma grande diversidade de informações com aplicações nas mais diversificadas áreas da biologia, como farmacêutica, biotecnológica e de ciência fundamental. No que tange ao estudo da resistência antimicrobiana e de novos antimicrobianos, as multi-ômicas potencializam o cenário de pesquisa, trazendo informações sobre anotação de genomas, genomas *core*, fatores responsáveis pela virulência bacteriana, genes essenciais à sobrevivência, potenciais novos antimicrobianos, entendimento das condições necessárias para expressão gênica de grupos de genes de interesse, perfis de transcrição de proteínas, modificações pós-traducionais, redes de interação proteína-proteína e reconstrução de modelos metabólicos e vias biossintéticas (Chernov *et al.*, 2019).

A descoberta de novas bacteriocinas é um ótimo exemplo do uso das ciências ômicas acopladas aos *pipelines* de descoberta de novas drogas. As características estruturais e funcionais desses compostos, assim como a organização de seus *operons* biossintéticos permitiram a identificação de bacteriocinas em diferentes organismos por meio de bancos de dados de peptídeos antimicrobianos e metabólitos secundários, como BAGEL, antiSMASH, BACTIBASE, e de ferramentas de análises ômicas, como ClusterFinder, BOA, entre outros (Chernov *et al.*, 2019).

Esses esforços já foram aplicados na mineração por tiopeptídeos (peptídeos antimicrobianos modificados pós-traducionalmente), os quais apresentam uma composição gênica bastante característica e conservada (Schwalen *et al.*, 2018); na descoberta do antibiótico avilamicina A, um potente antibiótico contra bactérias patogênicas Gram positivas, em *Streptomyces viridochromogenes* Tü57 (Grüning *et al.*, 2013). Trabalhos também têm aplicado a mineração por novos antimicrobianos em dados genômicos de microbiota humana, na esperança que substâncias lá identificadas apresentem menor toxicidade.

Zheng e colaboradores (2015) identificaram, em dados metagenômicos de microbiota intestinal humana, 4.875 bacteriocinas putativas, das quais 802 eram de classe I, 3.048 eram de classe II e 1.025 eram bacteriocinas grandes. Similarmente, Mousa e colaboradores (2017) usaram 6.009 genomas curados do projeto de *genome shotgun* do NCBI, para realizar um levantamento de potenciais genes para novos antimicrobianos encontrados em diferentes tecidos do corpo humano, identificando 3000 clusters de antimicrobianos já conhecidos nesses tecidos. Elucidando, assim, o grande poder das ciências ômicas em estudos de busca e desenvolvimento de novos antibióticos.

1.6 JUSTIFICATIVA

Os antibióticos representam um conjunto de substâncias com atividade antimicrobiana que revolucionaram a medicina e até hoje desempenham um papel muito importante no tratamento de doenças infecciosas e na profilaxia antimicrobiana de procedimentos diversos. No entanto, nos últimos anos, com a mesma velocidade em que antibióticos são introduzidos no mercado, estirpes resistentes a eles são identificadas, ameaçando sua eficácia. Ademais, é necessário ressaltar que as principais classes de antimicrobianos foram descobertas na era de ouro dos antibióticos e, desde então, novas classes não foram mais encontradas.

Sendo assim, o presente trabalho de mestrado almeja dar continuidade a pesquisa de bioprospecção de substâncias antimicrobianas em estirpes de *Streptococcus agalactiae* isolados de mastite bovina, iniciada pelo microbiologista Fabiano Murga da Silva, ex-aluno do Instituto de Microbiologia Paulo de Góes. Nessa conjuntura, tal pesquisa representa uma área com potencial de suprir a demanda por substâncias antimicrobianas alternativas que possam ocupar o papel desempenhado pelos antimicrobianos tradicionais, tais quais os antibióticos, com um foco nas áreas de conservação alimentícia e no tratamento de infecções em gado, como a própria mastite bovina. A identificação de peptídeos antimicrobianos em espécies de *Streptococcus* é ampla, principalmente para a espécie *Streptococcus mutans*; no entanto, pouco se sabe sobre bacteriocinas identificadas em estirpes de *Streptococcus agalactiae*. Ainda, a genômica é uma grande aliada de estudos de bioprospecção, já que possibilita a realização de varreduras em todo conteúdo gênico de um organismo, potencializando a descoberta de novas substâncias para aplicações biotecnológicas, dentro das quais se compreendem as bacteriocinas. A identificação dessas substâncias de potencial aplicação biotecnológica auxilia na diminuição do uso indiscriminado de antibióticos atualmente em risco, substituindo-o por substâncias mais específicas contra as espécies a serem combatidas e conseqüentemente diminuindo a pressão seletiva sobre espécies não-alvo.

Por fim, haja vista a importância clínica e o potencial patogênico de bactérias do gênero *Streptococcus*, é necessário monitorar a evolução de resistência antimicrobiana das mesmas, em se considerando principalmente, por exemplo, o nível de ameaça preocupante que apresentam *Streptococcus pyogenes* e *Streptococcus agalactiae* – sendo essa última a espécie estudada no presente trabalho de dissertação.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Analisar, por intermédio da genômica, a origem da atividade antimicrobiana de 4 estirpes de *Streptococcus agalactiae* isolados de mastite bovina, previamente identificados como produtores de substâncias antimicrobianas capazes de inibir *Streptococcus* spp..

2.2 OBJETIVOS ESPECÍFICOS

- Realizar testes de antibiose contra *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Cellulomonas fimi* e *Micrococcus* sp.;
- Confirmar a identificação taxonômica das estirpes de *Streptococcus agalactiae*, por meio de uma análise por sequências multilocus com genes *housekeeping* e com inferência de árvore filogenética pela metodologia da máxima verossimilhança;
- Verificar a presença de resistência antimicrobiana e potencial virulento nas estirpes com buscas específicas por genes que codificam fatores de virulência e que conferem resistência a antibióticos;
- Determinar por meio de buscas *in silico* a presença peptídeos antimicrobianos e metabólitos secundários com atividade antimicrobiana de interesse biotecnológico.

3 METODOLOGIA

3.1 ESTIRPES BACTERIANAS, MEIOS DE CULTURA E CONDIÇÕES DE CRESCIMENTO

3.1.1 Estirpes bacterianas

Como mencionado, as estirpes de *Streptococcus agalactiae* do Laboratório de Genética de Microrganismos associados a Alimentos e a Indústria já haviam sido previamente caracterizadas quanto a seu potencial de antibiose (Silva, 2003). Quatro estirpes produtoras de substâncias antimicrobianas (quadro 6) [LGMAI_St_03 (3774), LGMAI_St_08 (3772), LGMAI_St_11 (3800) e LGMAI_ST14 (3799)] foram selecionadas para dar continuidade a pesquisa e sequenciamento genômico, pois já havia caracterização prévia do caráter peptídico das suas substâncias e da resistência das mesmas ao aquecimento a 100°C. Essas estirpes inibem *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Streptococcus porcinus* e *Streptococcus uberis*, como descrito por Silva (2003) (anexo A e B). A estirpe de *Streptococcus mutans* UA140 (quadro 6), fornecida pelo pesquisador Page W. Caufield da Universidade do Alabama nos Estados Unidos da América, foi usada como controle positivo nos testes de antibiose, devido a sua produção de duas bacteriocinas: o lantibiótico mutacina I; e o não-lantibiótico mutacina IV (Qi; Chen; Caufield, 2001).

Quadro 6: Estirpes bacterianas produtoras.

| Estirpes | Organismo | Origem |
|---------------------|--------------------------------------|---|
| LGMAI_St_03 (3774) | <i>Streptococcus agalactiae</i> | Fazenda 2 – São Paulo |
| LGMAI_St_08 (3772) | <i>Streptococcus agalactiae</i> | |
| LGMAI_St_11 (3800) | <i>Streptococcus agalactiae</i> | Fazenda 3 – Minas Gerais |
| LGMAI_St_14 (3799) | <i>Streptococcus agalactiae</i> | |
| LGMAI_St_21 (UA140) | <i>Streptococcus mutans</i> UA140 | Departamento de Biologia Oral, Faculdade de Odontologia, Universidade do Alabama, EUA |

Com o objetivo de verificar que a atividade identificada por Silva (2003) ainda se mantinha após esses anos, foi realizado teste de antibiose contra a estirpe de *Streptococcus pyogenes* usada por Silva (2003) (quadro 7). Além disso, foram realizados testes de antibiose com novas bactérias indicadoras (quadro 7), compreendendo estirpes de *Staphylococcus aureus* e *Klebsiella pneumoniae*, com e sem resistência, devido a sua importância clínica e seus níveis elevados de resistência antimicrobiana relatados pela Anvisa no Brasil, assim como estirpes de *Cellulomonas fimi* e *Micrococcus* sp. por sua importância na indústria de alimentos, como degradação por atividade celulolítica de cultivares e deterioração do leite, respectivamente.

Quadro 7: Estirpes bacterianas indicadoras.

| Estirpes | Características | Origem |
|---|--|---|
| <i>Streptococcus pyogenes</i> | Bactéria patogênica e previamente testada por Silva (2003) | Laboratório de Microbiologia de Alimentos, UFRJ, Instituto de Microbiologia Paulo de Góes |
| <i>Cellulomonas fimi</i> | Suscetibilidade a antimicrobianos diversos | |
| <i>Micrococcus</i> sp. | Suscetibilidade a antimicrobianos diversos | |
| <i>Staphylococcus aureus</i> ATCC 6538 | Ausência de resistência a antimicrobianos | |
| <i>Klebsiella pneumoniae</i> ATCC 13883 | Ausência de resistência a antimicrobianos | |
| <i>Klebsiella pneumoniae</i> KPC | Resistência à carbapenema | Laboratório de Investigação em Microbiologia Médica, UFRJ, Instituto de Microbiologia Paulo de Góes |
| <i>Staphylococcus aureus</i> ATCC 25923 | Ausência de resistência a antimicrobianos | Laboratório de Infecções Hospitalares, UFRJ, Instituto de Microbiologia Paulo de Góes |
| <i>Staphylococcus aureus</i> ATCC 29213 | Ausência de resistência a antimicrobianos | |
| <i>Staphylococcus aureus</i> ATCC 33591 | Resistência à meticilina | |

3.1.2 Meios de cultura

As estirpes de *Streptococcus agalactiae* e *Streptococcus pyogenes* foram crescidas em meios de infusão de cérebro e coração (BHI) (Difco™). As estirpes de *Staphylococcus aureus* e *Klebsiella pneumoniae* foram crescidas em meio tripton de soja (TSB) (HiMedia™). As estirpes de *Micrococcus* sp. e *Cellulomonas fimi* foram crescidas em meio Luria-Bertani (LB), segundo a composição: peptona 1%, NaCl 1% e extrato de levedura 0,5%. Meios em placa foram acrescidos de 2% de ágar microbiológico Acumedia®, ao passo que meios semi-sólidos foram acrescidos de 1% de ágar microbiológico. Todos os testes de antibiose foram realizados em meio BHI.

3.1.3 Condições de crescimento

Todas as estirpes foram mantidas em duplicata em criotubos contendo meio Luria-Bertani (LB) com adição de 20% de glicerol em refrigeração a -80°C. As estirpes de *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Micrococcus* sp. e *Cellulomonas fimi* foram incubadas em estufa a 37°C, durante a noite, antes dos experimentos.

3.2 TESTES DE DETECÇÃO DE ATIVIDADE ANTIMICROBIANA

Para detecção de atividade antimicrobiana, foram realizados testes de antibiose pelo método de difusão em sobrecamada (*spots-on-lawn*) (Stoyancheva *et al.*, 2014) com algumas modificações. Nesses testes, as estirpes de *Streptococcus agalactiae* produtoras e o controle positivo de *Streptococcus mutans* UA140 (tabela 1) foram crescidos em 4 mL de caldo BHI a 37°C durante a noite. Desse crescimento, alíquotas de 10 µL foram coletadas e inoculadas na forma de pontos circulares sobre placas de BHI estéreis, as quais foram, em seguida, incubadas a 37°C durante a noite. Caldo BHI estéril foi usado como forma de controle negativo. Concomitantemente, as estirpes indicadoras foram crescidas em 4 mL de caldo TSB e LB, conforme descrito no tópico 3.1.2, e incubadas também a 37°C durante a noite. Depois da incubação, chumaços de algodão umedecidos com 1mL de clorofórmio foram colocados dentro das placas e lá permaneceram por 30 minutos, com o objetivo de que o vapor de clorofórmio eliminasse as células bacterianas que ali cresceram. Após o tratamento com vapor de clorofórmio, as placas foram abertas dentro da capela por 10 minutos, a fim de que todo resíduo de clorofórmio evaporasse. Por fim, 4 mL de semissólido BHI inoculado com 100 µL do crescimento das bactérias indicadoras foram vertidos sobre as placas recém tratadas com vapor de clorofórmio, e essas foram novamente incubadas a 37°C durante a noite para crescimento de tapete das bactérias indicadoras e posterior verificação da presença ou ausência de halos de inibição no dia seguinte.

3.3 EXTRAÇÃO DE DNA GENÔMICO, SEQUENCIAMENTO, MONTAGEM DE GENOMAS E ANÁLISES DE BIOINFORMÁTICA

3.3.1 Extração de DNA e sequenciamento de genomas

O DNA genômico das quatro estirpes foi extraído de acordo com o protocolo de extração por fenol-clorofórmio-álcool isoamílico (Chachaty; Saulnier, 2000), com algumas modificações (anexo C), e teve sua qualidade verificada por meio de gel de eletroforese e espectrofotometria de Nanodrop. Na espectrofotometria, foi verificada a absorção nas faixas de comprimento de onda de 260/280 nm e 260/230 nm, sendo o comprimento de onda de 260 nm, o ponto de maior absorbância para ácidos nucleicos, ao passo que os comprimentos de onda de 230 e 280 correspondem a máxima absorção para proteínas, fenol, EDTA e outros possíveis contaminantes residuais da extração de DNA. Os materiais genômicos foram, então, enviados para sequenciamento na Alemanha.

As etapas de fragmentação e construção das bibliotecas para realização do sequenciamento foram feitas pelo laboratório do professor e pesquisador Dr. Bertran Brening da Universidade de Göttingen, parceiros na Alemanha da Rede de Ciências Ômicas (RECOM), da qual faço parte. Para essa etapa, cinco $\mu\text{g}/\mu\text{L}$ de DNA genômico foram usados na construção de bibliotecas *paired-end* (2x150 pares de bases), em conformidade com as instruções do protocolo do fabricante do kit de fragmentação de DNA e preparação de bibliotecas NEBNext® *Fast DNA Fragmentation and Library Preparation Kit* (New England Biolabs Inc.). O sequenciamento foi realizado na plataforma Illumina Hi-Seq 2500, conforme o recomendado pelo fabricante.

3.3.2 Controle de qualidade e montagem dos genomas

A montagem dos genomas das estirpes de *S. agalactiae* LGMAI deste trabalho foi realizada pelos pesquisadores e professores Dr. Rommel Thiago Juca Ramos, Dr. Rodrigo Bentes Kato e Dr. Fabrício Almeida Araújo da Universidade Federal de Minas Gerais e colaboradores da Rede de Ciências Ômicas (RECOM). A primeira etapa desse processo foi a verificação da qualidade das leituras brutas saídas do sequenciador por meio dos softwares *FastQC* (Andrews, 2010) e *Adapter Removal* (Schubert; Lindgreen; Orlando, 2016.). Então, foi realizada a estimativa dos melhores valores de k-mer, que foram usados na montagem *de novo*, por meio do software *KmerStream* (Melsted; Halldórsson, 2014). Em seguida, eles realizaram a montagem *de novo* dos genomas, usando dois montadores: Edena (Hernandez *et al.*, 2008) e SPAdes (Bankevich *et al.*,

2012.). Por fim, os *contigs* oriundos de ambos montadores foram combinados, e o pacote PSI-CD-HIT do software CD-HIT (Fu *et al.*, 2012) foi usado para remover os *contigs* redundantes entre as duas montagens e produzir um arquivo final, contendo somente *contigs* únicos.

3.3.3 Controle de qualidade pós-montagens

A mensuração da qualidade das montagens também foi realizada pelos nossos colaboradores Dr. Rommel Thiago Juca Ramos, Dr. Rodrigo Bentes Kato e Dr. Fabrício Almeida Araújo da Universidade Federal de Minas Gerais. Nessa etapa, foi utilizada a ferramenta de avaliação de qualidade de montagens de genomas Quast (*Quality Assessment Tool*) (Mikheenko *et al.*, 2018). Essa ferramenta fornece as principais métricas de avaliação de montagens de genomas, como quantidade de *contigs*, tamanho total da montagem, conteúdo GC%, N50, entre outras. O N50 é uma métrica de avaliação da montagem dos genomas, que indica o tamanho do *contig* sobreposto ao ponto central do genoma, considerando que os *contigs* foram ordenados dos de maior tamanho para o de menor tamanho (Mäkinen; Salmela; Ylinen, 2012).

Após recebimento dos genomas montados, a fim de se completar as estratégias tradicionais de mensuração da qualidade das montagens, tais quais as métricas obtidas com a ferramenta Quast, e de adicionar informações de cunho biológico, analisei a completude dos genomas montados por meio da ferramenta BUSCO versão 5.1.2 (*Benchmarking Universal Single-Copy Orthologs*) (Seppey; Manni; Zdobnov, 2019). Essa ferramenta avalia a completude de uma montagem em termos de conteúdo gênico, realizando a busca por genes universais de cópia única evolutivamente conservados. Esses genes são obtidos do banco de dados OrthoDB (Kriventseva *et al.*, 2018), o qual contém genes ortólogos de cópia única, compartilhados por no mínimo 90% das espécies de um determinado grupo. Essa análise pode ser realizada para diversas linhagens de vertebrados, artrópodes, fungos, nemátodos, plantas, protistas e procariotos. Para essa análise, foi rodada a seguinte linha de comando, usando a linhagem das bactérias *bacteria_odb10*:

busco -i (arquivo .fasta de genoma) -l bacteria_odb10 -o (arquivo de resultado) -m genome

Depois disso, foram gerados vários documentos de saída com todas as informações da análise. Então, foi manualmente criada uma pasta “*summary*”, na qual foram adicionados os arquivos de resumo de cada uma das análises. Sendo assim, com o objetivo de gerar um gráfico contendo as principais informações das análises, foi rodado o script de plotagem de dados a seguir:

python3 scripts/generate_plot.py -wd (pasta “summary”)

3.3.4 Identificação filogenética e sorotipagem *in silico* com dados de sequenciamento genômico das estirpes produtoras de antimicrobianos

A identificação filogenética das 3 estirpes de *S. agalactiae* aqui sequenciados foi conduzida segundo uma sequência de análises, começando com uma análise mais abrangente e superficial e terminando com análises mais específicas e de alta confiabilidade científica. Essa metodologia foi escolhida, uma vez que análises filogenéticas mais específicas e de alta confiabilidade científica demoram muito para serem completadas; dessa forma, é necessário garantir que as análises estão corretamente encaminhadas desde o início a fim de evitar perdas de tempo desnecessárias. Para realização das análises descritas neste tópico, foram usados 39 genomas obtidos do NCBI, dos quais 37 correspondem a espécies válidas do gênero *Streptococcus* e 2 correspondem a espécies válidas do gênero *Lactococcus*, aqui usadas como grupos externos (quadro 8).

Quadro 8: Genomas usados nas análises filogenéticas e seus respectivos números de acesso no NCBI.

| Espécie | Número de acesso |
|---|------------------|
| <i>Streptococcus acidominimus</i> NCTC 12957 | GCA_900459045.1 |
| <i>Streptococcus agalactiae</i> NCTC 8181 | GCA_900458965.1 |
| <i>Streptococcus anginosus</i> subsp. <i>anginosus</i> NCTC 10713 | GCA_900636475.1 |
| <i>Streptococcus anginosus</i> subsp. <i>whileyi</i> CCUG 39159 | GCA_000257765.1 |
| <i>Streptococcus constellatus</i> subsp. <i>constellatus</i> NCTC 11325 | GCA_900459125.1 |
| <i>Streptococcus constellatus</i> subsp. <i>pharyngis</i> SK 1060 | GCA_000223295.2 |
| <i>Streptococcus criceti</i> HS 6 | GCA_000187975.3 |
| <i>Streptococcus downei</i> NCTC 11391 | GCA_900459175.1 |
| <i>Streptococcus dysgalactiae</i> subsp. <i>dysgalactiae</i> NCTC 13731 | GCA_900459225.1 |
| <i>Streptococcus dysgalactiae</i> subsp. <i>equisimilis</i> NCTC 13762 | GCA_900459095.1 |
| <i>Streptococcus equi</i> subsp. <i>equi</i> ATCC 33398 | GCA_900156215.1 |
| <i>Streptococcus equi</i> subsp. <i>ruminatorum</i> CECT 5772 | GCA_000706805.1 |
| <i>Streptococcus equi</i> subsp. <i>zooepidemicus</i> NCTC 4676 | GCA_900459475.1 |
| <i>Streptococcus equinus</i> NCTC 12969 | GCA_900459295.1 |
| <i>Streptococcus gallolyticus</i> subsp. <i>gallolyticus</i> DSM 16831 | GCA_002000985.1 |
| <i>Streptococcus gallolyticus</i> subsp. <i>macedonicus</i> NCTC 13767 | GCA_900459545.1 |
| <i>Streptococcus gallolyticus</i> subsp. <i>pasteurianus</i> NCTC 13784 | GCA_900478025.1 |
| <i>Streptococcus gordonii</i> NCTC 7865 | GCA_900475015.1 |
| <i>Streptococcus intermedius</i> NCTC 11324 | GCA_900475975.1 |
| <i>Streptococcus mitis</i> NCTC 12261 | GCA_000148585.3 |
| <i>Streptococcus mutans</i> NCTC 10449 | GCA_900475095.1 |
| <i>Streptococcus oralis</i> subsp. <i>dentisani</i> CECT 7747 | GCA_000382825.1 |
| <i>Streptococcus oralis</i> subsp. <i>oralis</i> NCTC 11427 | GCA_900637025.1 |
| Continua | |

| Conclusão | |
|---|------------------|
| Espécie | Número de acesso |
| <i>Streptococcus oralis</i> subsp. <i>tigurinus</i> AZ 3a | GCA_000344275.1 |
| <i>Streptococcus parasanguinis</i> ATCC 15912 | GCA_000164675.2 |
| <i>Streptococcus pneumoniae</i> NCTC 7465 | GCA_001457635.1 |
| <i>Streptococcus pseudopneumoniae</i> CCUG 49455 | GCA_002087075.1 |
| <i>Streptococcus pseudoporcinus</i> NCTC 13786 | GCA_900637075.1 |
| <i>Streptococcus pyogenes</i> NCTC 8198 | GCA_002055535.1 |
| <i>Streptococcus salivarius</i> subsp. <i>salivarius</i> NCTC 8618 | GCA_900636435.1 |
| <i>Streptococcus sanguinis</i> NCTC 7863 | GCA_900475505.1 |
| <i>Streptococcus sobrinus</i> NCTC 12279 | GCA_900475395.1 |
| <i>Streptococcus suis</i> S735 | GCA_000294495.1 |
| <i>Streptococcus thermophilus</i> NCTC 12958 | GCA_900474985.1 |
| <i>Streptococcus uberis</i> NCTC 3858 | GCA_900475595.1 |
| <i>Streptococcus vestibularis</i> ATCC 49124 | GCA_000188295.1 |
| <i>Lactococcus lactis</i> subsp. <i>cremoris</i> ATCC 19257 (grupo externo) | GCA_004354515.1 |
| <i>Lactococcus lactis</i> subsp. <i>lactis</i> ATCC 19435 (grupo externo) | GCA_900099625.1 |

A primeira análise realizada foi a de ANI (*Average Nucleotide Identity*). Essa análise calcula a identidade de nucleotídeos média de todos os genes ortólogos compartilhados por dois genomas microbianos. O programa usado nesta análise foi o fastANI (Jain *et al.*, 2018). Esse programa usa o algoritmo de mapeamento de sequências *Mashmap*, responsável por identificar todas as posições mapeadas entre a sequência de consulta e a sequência de referência, acima do limite de 80% de identidade. Esse mapeamento é realizado no sentido adiante e reverso, tanto para a sequência de consulta, quanto para a sequência de referência. Com isso, a identidade média das melhores correspondências recíprocas, identificadas nas análises em sentido adiante e reverso, culminam no valor final de ANI reportado na forma de porcentagem. Esse programa funciona na forma de linha de comando e permite a análise entre múltiplas entradas de sequências de consulta e de referência. O comando usado nesta análise foi o seguinte:

./fastANI --ql lista_consultas --rl lista_referencias -o resultados.ani

Nesse comando, --ql representa um arquivo contendo a lista das sequências de consulta, -rl representa um arquivo contendo a lista das sequências de referência e -o determina o arquivo de saída da análise.

Após a análise de identidade média de nucleotídeos, foi realizada uma análise filogenética dos 16S rRNAs do conjunto de genomas aqui usados. Para isso, a sequência desses rRNAs foi extraída com o programa Barrnap (Basic Rapid Ribosomal RNA Predictor) (<https://github.com/tseemann/barrnap>), com o seguinte comando:

```
barrnap -o rrna.fa < genoma.fasta > rrna.gff
```

A análise filogenética do 16S rRNA foi conduzida no software MEGA X (Kumar *et al.*, 2018). Esse programa permite a realização de todas as etapas necessárias para essa análise, contando com o alinhamento e edição das sequências, a predição dos melhores modelos de substituição de nucleotídeos e a inferência da árvore filogenética propriamente dita. O método usado foi o da máxima verossimilhança com 1000 replicações, usando o modelo de substituição de nucleotídeos de Kimura 2 parâmetros e distribuição gama com sítios invariáveis. Com o intuito de fortalecer o resultado da análise com o 16S rRNA e também de comparar o comportamento do delineamento das espécies nos 7 grandes subgrupos de *Streptococcus* (*anginosus*, *bovis*, *mitis*, *mutans*, *pyogenes*, *salivarius* e *sanguinis*), foi também realizada uma análise MLSA com outros 4 genes *housekeeping*. Os genes usados foram: *aroE* (chiquimato desidrogenase); *gki* (glicoquinase); *pheS* (fenilalanina-tRNA ligase subunidade alfa); e *recA* (recombinase A). Esses genes foram escolhidos com base no trabalho de taxonomia de *Streptococcus* de Thompson e colaboradores (2013). Para essa análise, foi usado o método de máxima verossimilhança com 1000 replicações, o modelo de substituição de nucleotídeos geral reversível no tempo e uma distribuição gama com sítios invariáveis. Árvores filogenéticas foram visualizadas com o pacote ggtree (Yu, 2020) para linguagem de programação R.

Por fim, foi realizada uma sorotipagem *in silico* de *Streptococcus* grupo B com dados de sequenciamento genômico, seguindo as diretrizes do protocolo do Centro de Controle e Prevenção de Doenças (Metcalf *et al.*, 2017).

3.3.5 Busca por genes e clusteres gênicos de peptídeos antimicrobianos

Com a finalidade de se buscar genes e *clusters* gênicos que codificam peptídeos antimicrobianos, alguns bancos de dados específicos para essas substâncias foram usados. O primeiro desses bancos de dados é o antiSMASH (*Antibiotics & Secondary Metabolites Analysis Shell*) (Medema *et al.*, 2011), no modo de busca rigorosa (“*strict*”) para identificação de clusteres bem definidos contendo todos os seus elementos. Ele tem o objetivo de fazer a anotação e identificação de clusteres gênicos de biossíntese de metabólitos secundários em genomas de bactérias e fungos. Ele foi usado, aqui, para buscar metabólitos secundários, uma vez que algumas dessas substâncias possuem importante papel na defesa dos organismos que as produzem e podem estar associadas a fenótipos de antibiose. Outro banco de dados usado foi o BAGEL4 (Van Heel *et al.*, 2018). Esse banco de dados conta com aproximadamente 820 sequências de antimicrobianos e tem como objetivo a anotação e identificação de clusteres gênicos de biossíntese de peptídeos sintetizados nos ribossomos e modificados pós-tradução (RiPPs), bacteriocinas grandes (>10 KDa) e bacteriocinas não modificadas. Por fim, foi usado o banco de dados BACTIBASE (Hammami *et al.*, 2007), o qual é uma plataforma especializada em bacteriocinas, contando com a função de anotação e identificação dos peptídeos estruturais de bacteriocinas, mas também portando informações estruturais e físico-químicas sobre essas substâncias. Nas buscas nos 3 bancos de dados, os arquivos fasta dos 3 genomas das estirpes de *Streptococcus agalactiae* isolados de mastite bovina foram submetidos para os respectivos servidores e a análise foi realizada seguindo os parâmetros padrões das ferramentas.

3.3.6 Análise dos genes e clusteres gênicos de peptídeos antimicrobianos

Após a identificação dos genes e clusteres gênicos de peptídeos antimicrobianos com auxílio dos bancos de dados descritos no item anterior, foi realizado o alinhamento das sequências encontradas com o software Clustal X (Larkin *et al.*, 2007), visando verificar discrepâncias, identidade e similaridades entre elas. Com o objetivo de identificar famílias e domínios proteicos, as sequências de peptídeos antimicrobianos encontradas foram submetidas a análise no banco de dados InterPro (Blum *et al.*, 2021). Essas sequências também foram submetidas a uma busca no banco de dados de proteínas UniProt (The UniProt Consortium, 2021). Na mesma plataforma, foi realizada uma busca pelo termo zoocina A, da seguinte forma: *name:“zoocin a”*. Em seguida, foi

escolhida a entrada O54308, pertencente ao organismo *Streptococcus equi* subsp. *zooepidemicus*, bactéria da qual a bacteriocina zoocina A foi originalmente isolada. A partir dessa entrada, foram então selecionadas proteínas similares com no mínimo 50% de identidade (tabela 6) e foi realizado um alinhamento múltiplo entre elas e as sequências encontradas para as estirpes LGMAI.

Tabela 6: Entradas das sequências proteicas usadas no alinhamento do UniProt.

| Entradas | Proteínas | Organismos | Tamanho |
|------------|---------------------|---|---------|
| A0A076Z6H4 | Peptidase (M23/M37) | <i>Streptococcus agalactiae</i> | 299 |
| A0A0H1Y377 | Peptidase (M23/M37) | <i>Streptococcus agalactiae</i> | 299 |
| Q8E7W5 | Não caracterizada | <i>Streptococcus agalactiae</i> sorotipo III (NEM316) | 299 |
| A0A380IVD1 | Peptidase (M23/M37) | <i>Streptococcus agalactiae</i> | 301 |
| A0A139QJJ1 | Peptidase (M23/M37) | <i>Streptococcus constellatus</i> | 285 |
| A0A2J9X4T6 | Zoocina A | <i>Streptococcus</i> sp. FDAARGOS_146 | 255 |
| A0A1E9XU90 | Zoocina A | <i>Streptococcus</i> sp. HMSC034B05 | 238 |
| A0A1F0CHF4 | Zoocina A | <i>Streptococcus</i> sp. HMSC069D09 | 299 |
| S8FKX0 | Zoocina A | <i>Streptococcus agalactiae</i> FSL S3-277 | 299 |
| A0A656FWC3 | Peptidase (M23/M37) | <i>Streptococcus agalactiae</i> (ATCC 13813/DSM 2134/JCM 5671/NCIMB 701348/NCTC 8181) | 299 |
| A0A656SN98 | Peptidase (M24/M37) | <i>Streptococcus equi</i> subsp. <i>zooepidemicus</i> SzAM60 | 285 |
| Q8E2F7 | Peptidase (M23/M37) | <i>Streptococcus agalactiae</i> sorotipo V (ATCC BAA-611/2603 V/R) | 299 |

Fonte: UniProt, disponível em <<https://www.uniprot.org>>.

Após, foi realizado um *blastx* dos três genomas contra a sequência da proteína Zif (código de acesso P74894 no banco de dados do UniProt). Essa sequência codifica uma aminoaciltransferase FemA de *Streptococcus equi* subsp. *zooepidemicus*, a qual atua como um fator de imunidade do tipo FemABX contra a ação da bacteriocina zoocina A. Esse fator de imunidade é de extrema importância para a sobrevivência das células após secreção da zoocina A. As sequências foram alinhadas e submetidas no Interpro para busca de domínios e famílias proteicas.

Os gráficos dos alinhamentos foram confeccionados com o pacote ggplot2, disponível na linguagem de programação R.

3.3.7 Avaliação de fatores de virulência, genes de resistência a antibióticos, ilhas genômicas de patogenicidade e resistência e sequências de fagos

Por se tratarem de estirpes potencialmente patogênicas, torna-se necessário realizar uma avaliação do potencial de virulência e resistência antimicrobiana das estirpes LGMAI de *S. agalactiae*. Para as análises de levantamento de fatores de virulência e genes de resistência, foi utilizado o software ABRicate (Seemann, 2014). Esse software tem como objetivo realizar uma varredura massiva em genomas por genes de resistência a antimicrobianos e fatores de virulência. Ele é composto por um conglomerado de 9 bancos de dados para resistência antimicrobiana e virulência bacteriana. Aqui, foram usados os bancos de dados ARG-ANNOT (Antibiotic Resistance Gene - Annotation) (Gupta *et al.*, 2014), CARD (Comprehensive Antibiotic Resistance Database) (Jia *et al.*, 2017), NCBI AMRFinderPlus tool (Feldgarden *et al.*, 2019) e Resfinder (Zankari *et al.*, 2012) para fazer o levantamento de genes que codificam resistência antimicrobiana. Para buscar por genes que codificam fatores de virulência, o banco de dados VFDB (Virulence Factors Database) (Chen *et al.*, 2016) foi usado. Os comandos a seguir foram usados nas buscas com o software ABRicate, nos quais fofn.txt representa um arquivo de texto com o caminho para os arquivos fasta dos genomas a serem analisados:

abricate --db argannot --fofn fofn.txt > argannot_results.tab

abricate --db card --fofn fofn.txt > card_results.tab

abricate --db ncbi --fofn fofn.txt > ncbi_results.tab

abricate --db resfinder --fofn fofn.txt > resfinder_results.tab

abricate --db vfdb --fofn fofn.txt > vfdb_results.tab

Para a predição de ilhas genômicas de patogenicidade e resistência, o software Gipsy (Soares *et al.*, 2015) foi usado. Esse software visa a predição de 4 tipos de ilhas genômicas (simbiose, metabólica, resistência antimicrobiana e patogenicidade) em genomas bacterianos, baseando-se nas seguintes características: desvios de assinatura genômicas, como conteúdo GC e uso de códons; presença de genes que codificam transposases, resistência a antibióticos, fatores de virulência, metabolismo e simbiose; e tRNAs flanqueadores. Ele requer o uso de um genoma de referência pertencente a um organismo não patogênico filogeneticamente próximo aos genomas estudados. Nessa análise, *Streptococcus thermophilus* JIM 8232, de acesso GCA_000253395.1, foi

escolhido e usado como o genoma de referência não patogênico requerido pelo programa. Foram buscadas ilhas genômicas de resistência e patogenicidade.

Por fim, uma vez que sequências de fagos permitem a aquisição de resistência antimicrobiana por bactérias, potencialmente tornando-as mais bem adaptadas a diferentes nichos e mais patogênicas, foi usada a ferramenta online Phaster (*Phage Search Tool Enhanced Release*) (Arndt *et al.*, 2016) para identificação e anotação dessas sequências.

3.3.8 Anotação funcional e genômica comparativa

A ferramenta *eggNOG-mapper* (Cepas *et al.*, 2017) foi usada para realizar a anotação funcional dos genomas. Essa ferramenta analisa grandes conjuntos de dados genômicos, baseando-se em clusters de ortólogos oriundos do banco de dados eggNOG v5.0 (Cepas *et al.*, 2019), e anota os genes em diferentes grupos COG (*Cluster of Orthologous Groups*) (quadro 9).

Quadro 9: Detalhamento das categorias funcionais COG usadas pela anotação da ferramenta *eggNOG-mapper*.

| Símbolos | Descrição |
|-----------------|--|
| A | Processamento e modificação de RNA |
| B | Estrutura e dinâmica da cromatina |
| C | Produção e conversão de energia |
| D | Controle do ciclo celular, divisão celular e partição de cromossomos |
| E | Transporte e metabolismo de amino ácidos |
| F | Transporte e metabolismo de nucleotídeos |
| G | Transporte e metabolismo de carboidratos |
| H | Transporte e metabolismo de coenzimas |
| I | Transporte e metabolismo de lipídios |
| J | Tradução, estrutura dos ribossomos e biogênese |
| K | Transcrição |
| L | Replicação, recombinação e reparo |
| M | Biogênese de parede celular/membrana/envelope |
| N | Motilidade celular |
| O | Modificação pós-traducional, metabolismo de proteínas e chaperonas |
| P | Transporte e metabolismo de íons inorgânicos |
| Q | Biossíntese, transporte e catabolismo de metabólitos secundários |
| R | Predição de função geral somente |
| Continua | |

| | | Conclusão |
|----------|---|-----------|
| Símbolos | Descrição | |
| S | Função desconhecida | |
| T | Mecanismos de transdução de sinal | |
| U | Tráfego intracelular, secreção e transporte vesicular | |
| V | Mecanismos de defesa | |
| W | Estruturas extracelulares | |
| X | Mobiloma: prófagos e transposões | |
| Y | Estrutura nuclear | |
| Z | Citoesqueleto | |

Fonte: Banco de dados de COGs (Clusteres de Genes Ortólogos) do NCBI, disponível em <<https://www.ncbi.nlm.nih.gov/research/cog/>>.

Ainda em relação à anotação funcional dos genomas, o servidor de anotação automática de genomas e reconstrução de vias metabólicas (KAAS) (Moriya *et al.*, 2007) do banco de dados KEGG (Kanehisa; Sato, 2020) foi usado para anotar e designar os ortólogos funcionais KEGG presentes nos genomas montados.

Por fim, a plataforma OrthoVenn2 (Xu *et al.*, 2019) foi usada para a genômica comparativa. Essa ferramenta realiza a identificação de grupos ortólogos no arquivo de proteínas enviado para análise, fazendo a anotação e o agrupamento em clusteres desses ortólogos.

Os gráficos com os resultados da anotação funcional foram confeccionados com o pacote ggplot2, disponível na linguagem R. Os resultados da genômica comparativa foram plotados com o pacote *treemap* do autor Martijn Tennekes, disponível na linguagem R, para visualização hierárquica dos grupos funcionais em retângulos aninhados

O software BRIG (Alikhan *et al.*, 2011) foi usado para comparar a identidade de nucleotídeos entre os genomas das estirpes LGMAI e 15 genomas completos de *Streptococcus agalactiae* (tabela 7). Além da comparação entre os genomas, também foram plotadas as ilhas genômicas preditas para verificar a conservação das mesmas. O genoma da estirpe LGMAI_St_08 foi escolhido como o genoma de referência por ser o maior em tamanho (pb).

Tabela 7: Genomas completos de *Streptococcus agalactiae* usados na comparação contra as estirpes LGMAI por meio do software BRIG.

| Genoma | Número de acesso no NCBI | Tamanho (pb) |
|--|---------------------------------|---------------------|
| <i>Streptococcus agalactiae</i> 2-22 | GCA_000967445.1 | 1.838.867 |
| <i>Streptococcus agalactiae</i> 09mas018883 | GCF_000427035.1 | 2.138.694 |
| <i>Streptococcus agalactiae</i> 138P | GCA_000599965.1 | 1.838.701 |
| <i>Streptococcus agalactiae</i> 138spar | GCA_000636115.1 | 1.838.126 |
| <i>Streptococcus agalactiae</i> 2603V/R | GCF_000007265.1 | 2.160.267 |
| <i>Streptococcus agalactiae</i> A909 | GCF_000012705.1 | 2.127.839 |
| <i>Streptococcus agalactiae</i> CNCTC 10/84 | GCF_000782855.1 | 2.013.842 |
| <i>Streptococcus agalactiae</i> COH1 | GCF_000689235.1 | 2.065.074 |
| <i>Streptococcus agalactiae</i> GD201008-001 | GCF_000299135.1 | 2.063.112 |
| <i>Streptococcus agalactiae</i> ILRI005 | GCF_000427075.1 | 2.109.759 |
| <i>Streptococcus agalactiae</i> ILRI112 | GCA_000427055.1 | 2.029.198 |
| <i>Streptococcus agalactiae</i> NEM316 | GCF_000196055.1 | 2.211.485 |
| <i>Streptococcus agalactiae</i> NGBS061 | GCF_000730215.1 | 2.221.207 |
| <i>Streptococcus agalactiae</i> NGBS572 | GCF_000730255.1 | 2.061.426 |
| <i>Streptococcus agalactiae</i> SA20 | GCA_000302475.3 | 1.841.952 |

Fonte: Informações coletadas do banco de dados do NCBI, disponível em < <https://www.ncbi.nlm.nih.gov/assembly/>>.

Por fim, o pacote *CGView Comparison Tool* (Grant; Arantes; Stothard, 2012) foi usado para confecção dos mapas gráficos dos três genomas montados e visualização das anotações realizadas para os mesmos.

4 RESULTADOS

4.1 TESTES DE DETECÇÃO DE ATIVIDADE ANTIMICROBIANA

Com os novos testes de detecção de atividade antimicrobiana, observou-se que todas as 4 estirpes de *Streptococcus agalactiae* do LGMAI foram capazes de inibir as estirpes de *Cellulomonas fimi* e *Micrococcus* sp. No entanto, elas não foram capazes de inibir as estirpes de *Klebsiella pneumoniae* nem *Staphylococcus aureus*, com ou sem resistência a antimicrobianos. Além disso, observou-se que a inibição descrita por Silva (2003) contra *Streptococcus pyogenes* se mantém (quadro 10).

Quadro 10: Espectro de ação das estirpes produtoras de substância antimicrobiana. O sinal de (+) representa a presença de halos de inibição acima de 0,2 cm; (++) com halos acima de 1,5 cm; enquanto que (-) indica ausência de inibição.

| Estirpes indicadoras | Estirpes produtoras | | | | |
|---|---------------------|-------------|-------------|-------------|-------------|
| | LGMAI_St_03 | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 | LGMAI_St_21 |
| <i>Streptococcus pyogenes</i> | + | + | + | + | - |
| <i>Cellulomonas fimi</i> | + | + | + | + | ++ |
| <i>Micrococcus</i> sp. | + | + | + | + | ++ |
| <i>Klebsiella pneumoniae</i> ATCC 13883 | - | - | - | - | - |
| <i>Klebsiella pneumoniae</i> KPC | - | - | - | - | - |
| <i>Staphylococcus aureus</i> ATCC 6538 | - | - | - | - | - |
| <i>Staphylococcus aureus</i> ATCC 29213 | - | - | - | - | - |
| <i>Staphylococcus aureus</i> ATCC 25923 | - | - | - | - | - |
| <i>Staphylococcus aureus</i> ATCC 33591 | - | - | - | - | - |

4.2 GENOMAS

O material genômico enviado para o sequenciamento da estirpe LGMAI_St_03 não pôde ser sequenciado, pois não apresentava qualidade suficientemente boa que permitisse o sequenciamento confiável do genoma da mesma. Desse modo, foram sequenciados com êxito os genomas das estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14, genomas com os quais sucederam as análises de bioinformática.

Por meio da ferramenta Quast de avaliação da qualidade de genomas, foi observado que: a estirpe LGMAI_St_08 obteve um genoma com 3.523.890 pb, conteúdo GC de 36,5% e um total de 2.915 proteínas anotadas; enquanto que a estirpe LGMAI_St_11 obteve um genoma com 2.683.705 pb, conteúdo GC de 35,8% e um total de 2.651 proteínas anotadas; e por fim, que a estirpe LGMAI_St_14 obteve um genoma de 2.690.254 pb, conteúdo GC de 35,7% e um total de 2.640 proteínas anotadas (tabela 8).

Tabela 8: Estatísticas dos três genomas montados.

| Estirpe | Atributo | Valor | Porcentagem ^b |
|---------------------------------------|---------------------------------------|-------------------|--------------------------|
| LGMAI_St_08 | Tamanho do genoma | 3.523.890 | 100% |
| | Conteúdo GC | 1.285.867 | 36,5% |
| | N50 | 47.478 | 1,4% |
| | Número de <i>contigs</i> ^a | 258 | 100% |
| | L50 | 23 | 8,9% |
| | Total de genes | 2.915 | 100% |
| | rRNA | 8 | 0,3% |
| | tRNA | 63 | 2,2% |
| | LGMAI_St_11 | Tamanho do genoma | 2.683.705 |
| Conteúdo GC | | 961.035 | 35,8% |
| N50 | | 47.509 | 1,8% |
| Número de <i>contigs</i> ^a | | 117 | 100% |
| L50 | | 17 | 14,5% |
| Total de genes | | 2.651 | 100% |
| rRNA | | 9 | 0,3% |
| tRNA | | 56 | 2,1% |
| LGMAI_St_14 | | Tamanho do genoma | 2.690.254 |
| | Conteúdo GC | 960.959 | 35,7% |
| | N50 | 47.121 | 1,8% |
| | Número de <i>contigs</i> ^a | 119 | 100% |
| | L50 | 20 | 16,8% |
| | Total de genes | 2.640 | 100% |
| | rRNA | 9 | 0,3% |
| | tRNA | 70 | 2,7% |

a: Leva-se em consideração somente *contigs* acima de 100pb.

b: As porcentagens são baseadas no tamanho do genoma (pb), quantidade de *contigs* ou de genes, dependendo do atributo em questão.

N50 representa o tamanho do *contig* sobreposto ao ponto central do genoma, onde *contigs* estão em ordem decrescente. L50 representa a menor quantidade de *contigs*, que somados compõem metade do genoma.

Para os três genomas, foram encontrados os 124 ortólogos universais evolutivamente conservados para a linhagem das bactérias, dos quais 2 encontram-se fragmentados (figura 11).

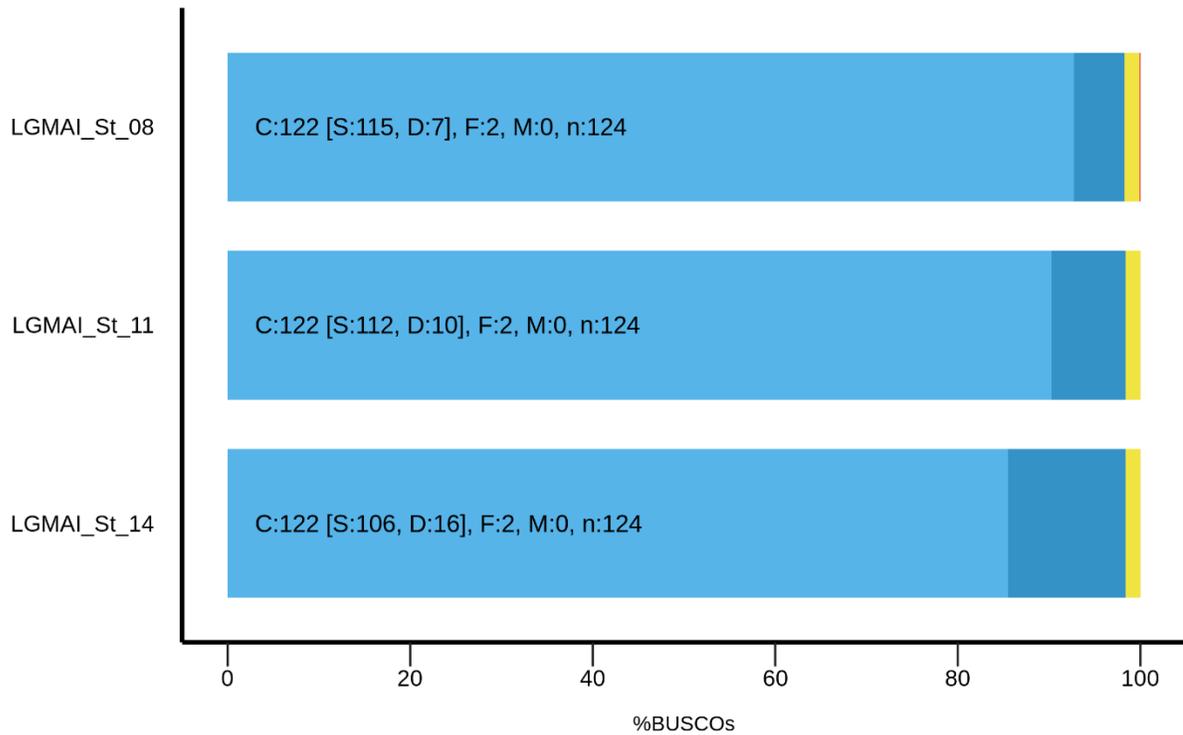


Figura 11: Resultados da análise de completude por busca de ortólogos universais realizada com a ferramenta Busco. No eixo y, encontram-se os três genomas montados. A cor azul representa genes completos (C) de cópia única (S) ou duplicados (D); a cor amarela representa genes fragmentados (F); e a cor vermelha representa genes não identificados (M). A letra n indica o total de genes ortólogos usados na análise.

Os genomas montados das estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14 estão submetidos no NCBI sob o mesmo *BioProject* de número PRJNA637496, *BioSamples* de números SAMN17072134, SAMN17072135 e SAMN17072136, respectivamente. Os números de acesso aos genomas no *GenBank* são JAIWPA000000000, JAIWPB000000000 e JAIWPC000000000, respectivamente.

4.3 IDENTIFICAÇÃO FILOGENÉTICA E SOROTIPAGEM *IN SILICO* COM DADOS DE SEQUENCIAMENTO GENÔMICO DAS ESTIRPES PRODUTORAS DE ANTIMICROBIANOS

A primeira análise do fluxo de trabalho da identificação filogenética foi a de ANI, a qual calcula a identidade média de nucleotídeos entre os genomas analisados, considerando somente genomas com identidade entre 80% a 100%. O algoritmo, no entanto, consegue ainda computar genomas um pouco mais distantes, isto é, com identidade abaixo de 80%. Observou-se que as três estirpes LGMAI tiveram maior identidade média de nucleotídeos com o genoma de referência *Streptococcus agalactiae* NCTC 8181, com as três estirpes apresentando identidade por volta de 98% (figura 12).

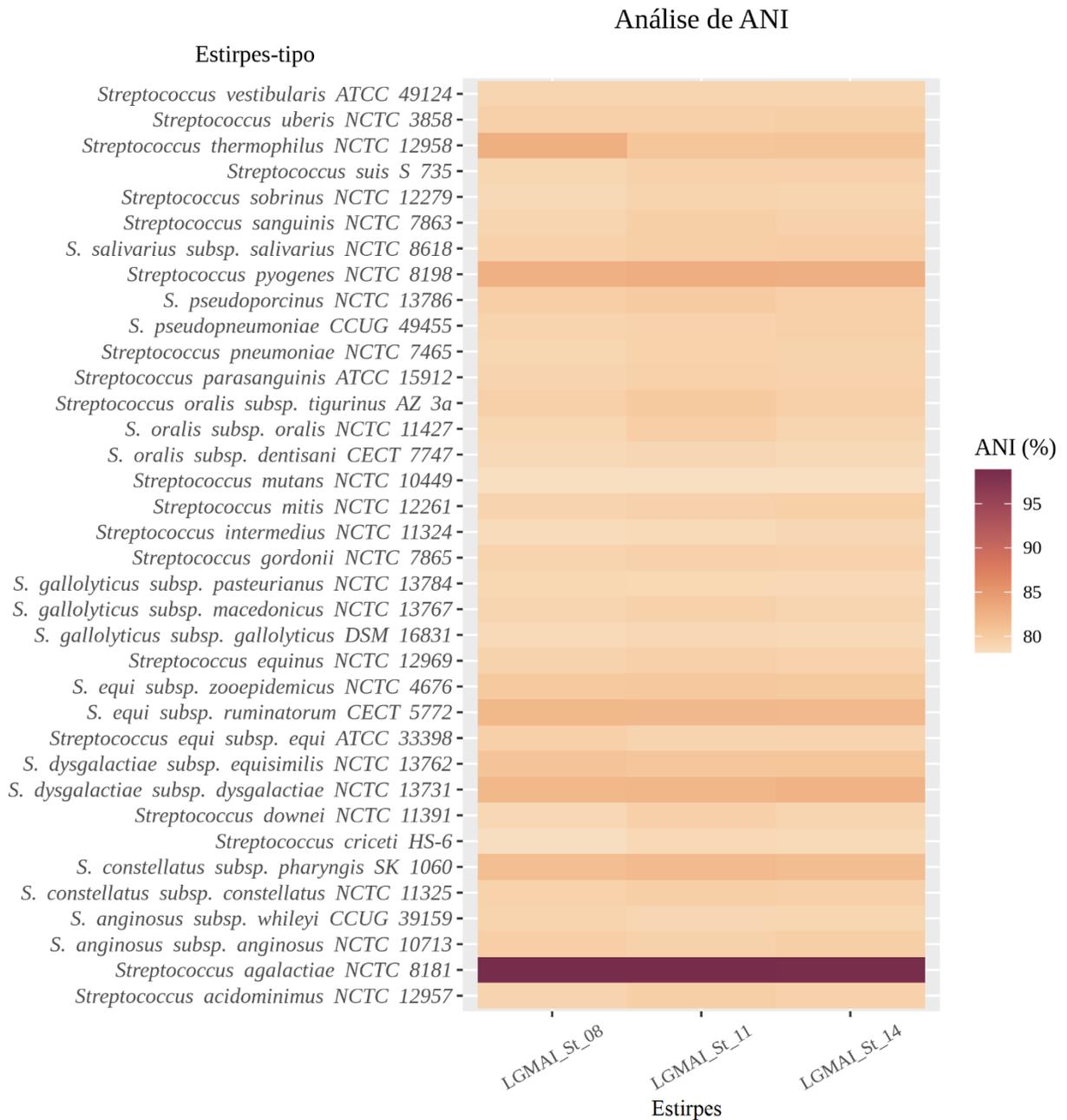


Figura 12: Mapa de calor dos resultados obtidos pela análise de ANI. Estirpes-tipo das principais espécies do gênero *Streptococcus* estão no eixo y, e as estirpes LGMAI estão no eixo x. À direita, legenda contendo o espectro de cores relacionado à faixa de variação para as porcentagens (~80% - 100%) obtidas na análise. Blocos de coloração vinho indicam alta identidade (95% - 100%), ao passo que os de coloração bege indicam baixa identidade (80% - 85%).

As estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14 apresentaram identidade de 98,87%, 98,81% e 98,77%, respectivamente, com *Streptococcus agalactiae* NCTC 8181; e os segundo e terceiro melhores resultados obtidos foram na faixa de 80% com as espécies *Streptococcus thermophilus* e *Streptococcus pyogenes* (tabela 9) e, portanto, bem inferiores. Os resultados da análise de ANI mostraram que o gênero trabalhado está correto e que as estirpes potencialmente pertencem à espécie *Streptococcus agalactiae*. A tabela completa com os resultados da análise de ANI encontra-se disponível no anexo D.

Tabela 9: Recorte dos resultados da análise de ANI, contendo as 5 melhores identidades.

| Espécies de referência | Estirpes | | |
|--|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| <i>S. agalactiae</i> NCTC 8181 | 98,87% | 98,81% | 98,77% |
| <i>S. thermophilus</i> NCTC 12958 | 82,90% | 80,61% | 80,69% |
| <i>S. pyogenes</i> NCTC 8198 | 82,69% | 82,95% | 82,78% |
| <i>S. equi</i> subsp. <i>ruminatorum</i> CECT 5772 | 82,03% | 81,98% | 81,87% |
| <i>S. dysgalactiae</i> subsp. <i>dysgalactiae</i> NCTC 13731 | 81,99% | 82,10% | 82,51% |

A partir da inferência filogenética realizada com os genes que codificam o 16S rRNA, pode-se observar um delineamento das espécies do gênero *Streptococcus* em 6 grandes grupos: “*anginosus*”, “*bovis*”, “*mitis*”, “*mutans*”, “*pyogenes*” e “*salivarius*” (figura 13). As 3 estirpes LGMAI ficaram posicionadas dentro do grupo “*pyogenes*” e agruparam-se com o genoma de referência de *Streptococcus agalactiae* NCTC 8181 com 100% na estatística de *bootstrap* (figura 13). Esse resultado corrobora com o resultado obtido na análise de ANI e por si só já é muito forte, haja vista a clara taxonomia das principais espécies do gênero *Streptococcus*.

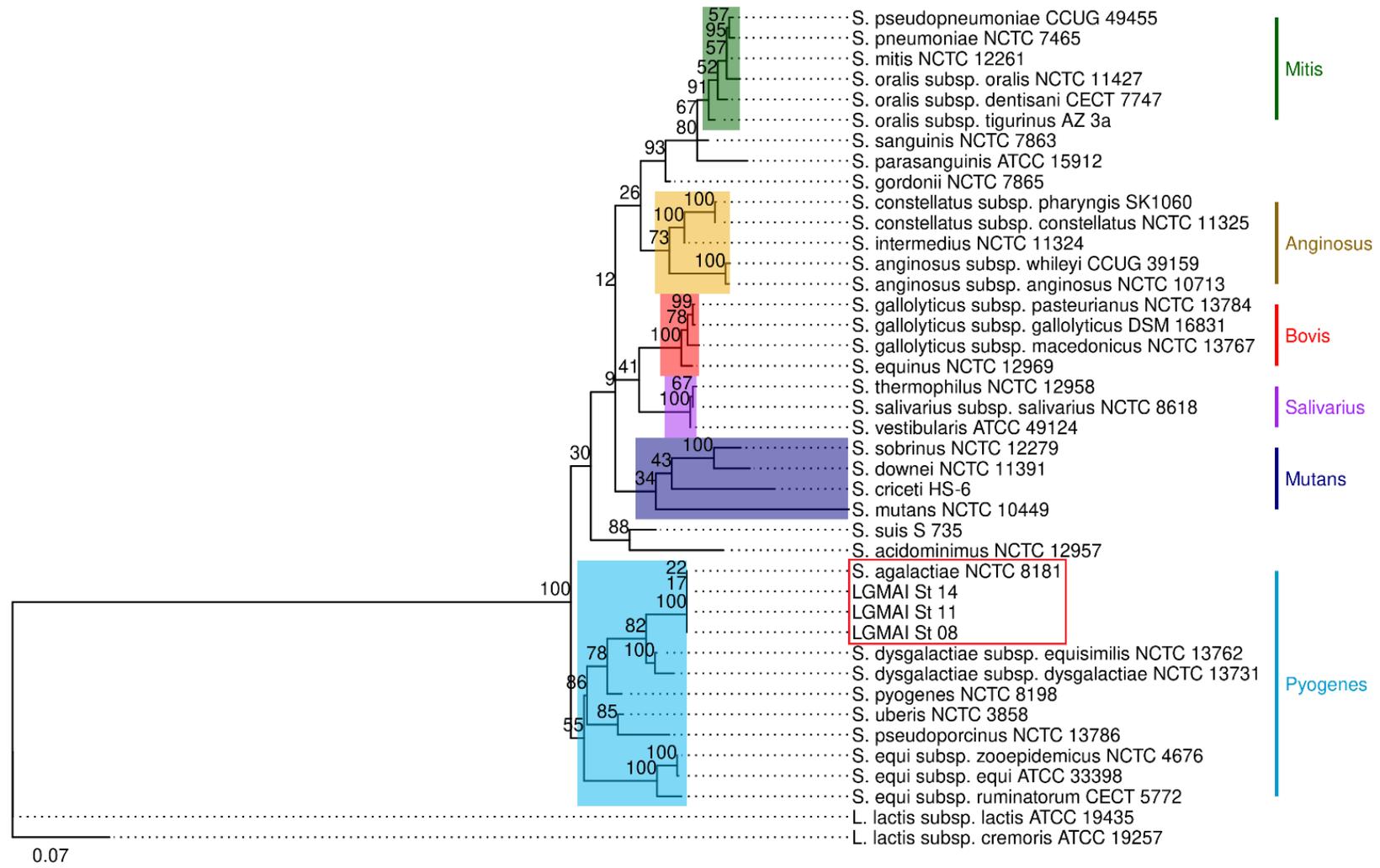


Figura 13: Árvore filogenética com o 16S rRNA, inferida por máxima verossimilhança com 1000 replicações, usando o modelo de substituição de nucleotídeos de Kimura 2 parâmetros e distribuição gama com sítios invariáveis. Destacado com retângulo de bordas vermelhas, encontra-se o grupo das estirpes LGMAI com a referência de *Streptococcus agalactiae* NCTC 8181. Retângulos preenchidos representam delineamentos em grupos de *Streptococcus*.

A análise MLSA foi realizada, visando trazer maior confiabilidade à identificação filogenética das estirpes LGMAI. Nela, além do gene que codifica o 16S rRNA, foram usados os genes *housekeeping aroE* (chiquimato desidrogenase), *gki* (glicoquinase), *pheS* (fenilalanina-tRNA ligase subunidade alfa) e *recA* (recombinase A). A partir da árvore filogenética inferida nessa análise, observou-se que novamente as 3 estirpes LGMAI formaram um grupo com o genoma de referência *Streptococcus agalactiae* NCTC 8181 com 100% de estatística *bootstrap* (figura 14). Além disso, diferentemente da análise que continha somente o 16S rRNA, pode-se ver o delineamento das espécies analisadas em 7 grandes grupos: “*anginosus*”, “*bovis*”, “*mitis*”, “*mutans*”, “*pyogenes*”, “*salivarius*” e “*sanguinis*” (figura 14). Agora, pode-se observar o início de um delineamento do grupo “*sanguinis*”. No entanto, teoricamente, a espécie *S. parasanguinis* faz parte do grupo “*sanguinis*”, mas, na análise aqui realizada, isso não pôde ser observado. Tal fato é indicativo de que, para um melhor delineamento dos mesmos, é necessário o uso de mais genes *housekeeping* e talvez mais representantes das espécies na análise MLSA.

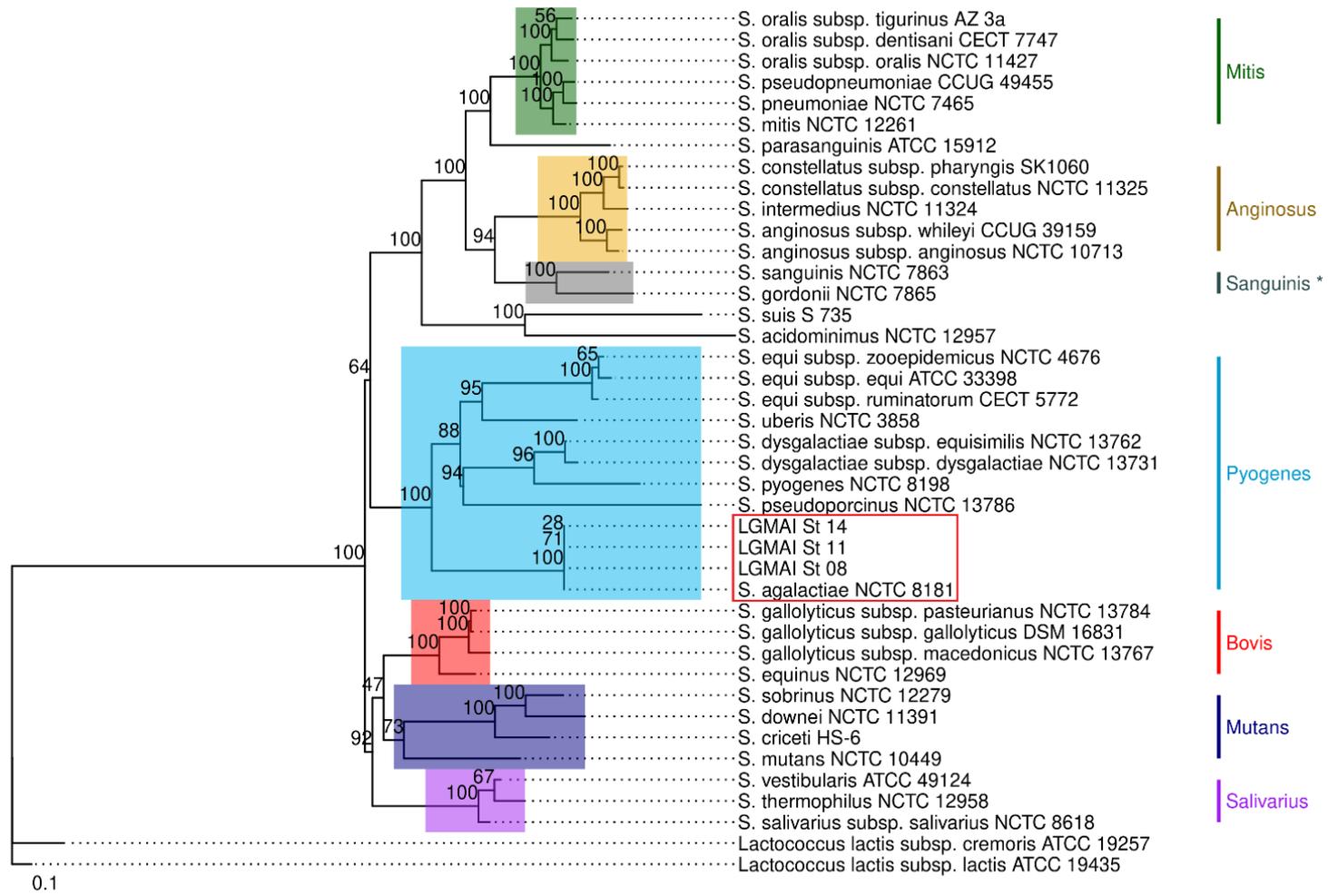


Figura 14: Árvore filogenética resultante da MLSA com os genes *aroE*, *gki*, *pheS* e *recA* e o gene para o 16S rRNA, inferida pelo método da máxima verossimilhança com 1000 replicações, usando o modelo de substituição de nucleotídeos geral reversível no tempo e uma distribuição gama com sítios invariáveis. Destacado com retângulo de bordas vermelhas, encontra-se o grupo das estirpes LGMAI com a referência de *Streptococcus agalactiae* NCTC 8181. Retângulos preenchidos representam delineamentos em grupos de *Streptococcus*.

Já com a sorotipagem *in silico* realizada para *Streptococcus* grupo B com dados de sequenciamento genômico, seguindo o protocolo do Centro de Controle e Prevenção de Doenças, observou-se que os *loci cps* extraídos das estirpes LGMAI tiveram correspondência de 99% de identidade com o marcador para o sorotipo III (figura 15, A.), com uma única mutação na posição 76 (figura 15, B.).

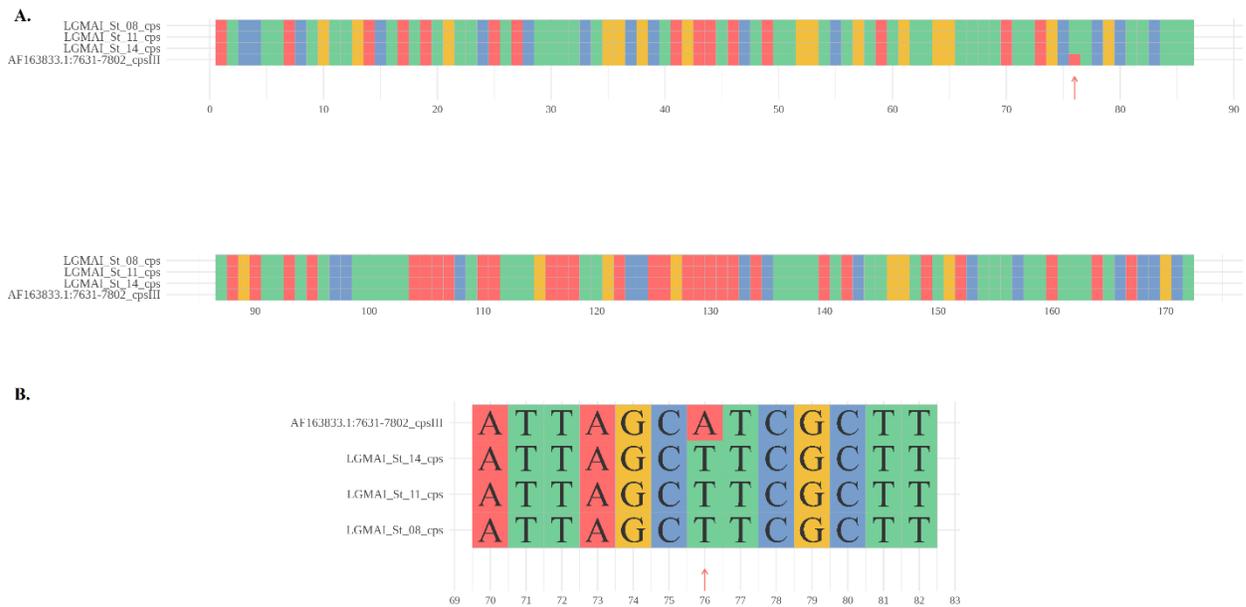


Figura 15: Resultados da sorotipagem *in silico* com dados de sequenciamento genômico para *Streptococcus* grupo B. Em **A.**, alinhamento dos 3 *loci cps* extraídos das estirpes LGMAI com o marcador para o sorotipo III, no qual a seta indica a única posição variável entre as sequências. Em **B.**, recorte do alinhamento do mesmo alinhamento, mostrando em mais detalhes a posição 76 e a mutação ali presente.

Segundo as diretrizes deste protocolo, o marcador para o sorotipo III também é encontrado em bactérias do sorotipo II (tabela 5, página 53-54). Dessa forma, ao encontrar correspondência com o marcador do sorotipo III, as estirpes avaliadas podem ser tanto do sorotipo II, quanto do III. No entanto, uma vez que não foi encontrada nenhuma correspondência com o marcador do sorotipo II na análise, as estirpes LGMAI são, dessa forma, classificadas como *Streptococcus agalactiae* pertencentes ao sorotipo III.

4.4 AVALIAÇÃO DE FATORES DE VIRULÊNCIA, GENES DE RESISTÊNCIA A ANTIBIÓTICOS, ILHAS GENÔMICAS DE PATOGENICIDADE E RESISTÊNCIA E SEQUÊNCIAS DE FAGOS

Na busca por genes que codificam resistência antimicrobiana com a ferramenta ABRicate (Seemann, 2014), não foi encontrado nenhuma correspondência com os genes presentes nos bancos de dados do NCBI AMRFinderPlus tool (Feldgarden *et al.*, 2019) e ARG-ANNOT (Gupta *et al.*, 2014). No entanto, foram encontrados os genes *mprF* e *mreA* nos bancos de dados CARD (Jia *et al.*, 2017) e Resfinder (Zankari *et al.*, 2012), respectivamente (tabela 10).

Tabela 10: Genes de resistência antimicrobiana identificados pela ferramenta ABRicate nos bancos de dados CARD e Resfinder.

| Banco de dados | | Estirpe | | |
|------------------|------------|----------------------|----------------------|-------------|
| | | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| CARD | Sequência | L2_16 | L2_25 | L2_21 |
| | Região | 8935-11478 | 8590-11133 | 24890-27433 |
| | Gene | <i>mprF</i> | <i>mprF</i> | <i>mprF</i> |
| | Cobertura | 100% | 100% | 100% |
| | Identidade | 99,49% | 99,49% | 99,49% |
| Resfinder | Sequência | NODE_11_length_50934 | NODE_1_length_258160 | L2_1 |
| | Região | 17023-17958 | 211039-211974 | 46788-47723 |
| | Gene | <i>mreA</i> | <i>mreA</i> | <i>mreA</i> |
| | Cobertura | 100% | 100% | 100% |
| | Identidade | 99,79% | 99,79% | 99,79% |

Na busca por genes que codificam fatores de virulência com a ferramenta ABRicate (Seemann, 2014), usando o banco de dados VFDB (Chen *et al.*, 2016), foram encontrados 24 genes para as estirpes LGMAI_St_11 e LGMAI_St_14 e 25 genes para a estirpe LGMAI_St_08 (Tabela 11).

Tabela 11: Recorte dos resultados da análise de busca por fatores de virulência no banco de dados VFDB. As porcentagens correspondem às respectivas identidades dos genes encontrados nas estirpes LGMAI em relação às sequências presentes no banco de dados VFDB.

| Genes | | Estirpe | | | |
|------------------|-------------------|-------------|-------------|-------------|------|
| | | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 | |
| Locus <i>cps</i> | <i>cpsA</i> | 99,45% | 99,45% | 99,45% | |
| | <i>cpsB</i> | 99,86% | 99,86% | 99,86% | |
| | <i>cpsC</i> | 99,71% | 99,71% | 99,71% | |
| | <i>cpsD</i> | 98,43% | 98,43% | 98,43% | |
| | <i>cpsE</i> | 99,35% | 99,35% | 99,35% | |
| | <i>cpsF</i> | 99,56% | 99,56% | 99,56% | |
| | <i>cpsK</i> | 98,33% | 98,33% | 98,33% | |
| | <i>cpsL</i> | 97,50% | 97,50% | 97,50% | |
| | <i>neuA</i> | 99,60% | 99,60% | 99,60% | |
| | <i>neuB</i> | 99,71% | 99,71% | 99,71% | |
| | <i>neuC</i> | 98,96% | 98,96% | 98,96% | |
| | <i>neuD</i> | 99,68% | 99,68% | 99,68% | |
| | Operon <i>cyl</i> | <i>cylX</i> | 100% | 100% | 100% |
| | | <i>cylD</i> | 100% | 100% | 100% |
| <i>cylG</i> | | 100% | 100% | 100% | |
| <i>acpC</i> | | 100% | 100% | 100% | |
| <i>cylZ</i> | | 100% | 100% | 100% | |
| <i>cylA</i> | | 100% | 100% | 100% | |
| <i>cylB</i> | | 99,43% | 99,43% | 99,43% | |
| <i>cylF</i> | | 99,06% | 99,06% | 99,06% | |
| <i>cylI</i> | | 99,18% | 99,18% | 99,18% | |
| <i>cylJ</i> | | 99,84% | 98,84% | 98,84% | |
| <i>cylK</i> | | 99,13% | 99,13% | 99,13% | |
| <i>hylB</i> | 98,94% | 98,94% | 98,94% | | |
| <i>cfa/cfb</i> | 98,96% | NA | NA | | |

Os resultados completos desta análise, contendo informações de *contigs* e posição onde os genes foram identificados, assim como cobertura e descrição de função dos genes, encontram-se disponíveis no anexo F.

A análise de predição de ilhas genômicas putativas de resistência e patogenicidade, realizada pela ferramenta Gipsy (Soares *et al.*, 2015), resultou na identificação de 12 ilhas genômicas com escores forte e normal para a estirpe LGMAI_St_08, 11 ilhas genômicas com escores forte e normal para a estirpe LGMAI_St_11 e 13 ilhas genômicas com escores forte e normal para a LGMAI_St_14 (tabela 12).

Tabela 12: Recorte dos resultados obtidos com o programa Gipsy. Aqui são mostradas somente as ilhas genômicas putativas com escores forte e normal. Res: Ilhas de resistência; Pat: Ilhas de patogenicidade; Misc: Ilhas miscelâneas.

| Estirpe | Ilha genômica | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | Posição | Escore |
|-------------|---------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| LGMAI_St_08 | Res 1 | 27% | 36% | 18% | 45% | 1128892-1164378 | Normal |
| | Res 2 | 14% | 0% | 42% | 71% | 1259251-1268895 | Normal |
| | Res 3 | 13% | 71% | 8% | 86% | 1545086-1602299 | Forte |
| | Pat 1 | 30% | 0% | 76% | 38% | 1106326-1118347 | Normal |
| | Pat 2 | 9% | 17% | 57% | 55% | 1652611-1843228 | Normal |
| | Pat 3 | 8% | 11% | 61% | 58% | 1906774-1954836 | Normal |
| | Pat 4 | 11% | 7% | 76% | 80% | 2607886-2641064 | Normal |
| | Pat 5 | 13% | 0% | 79% | 93% | 3196181-3234263 | Normal |
| | Pat 6 | 26% | 41% | 47% | 78% | 3237548-3371673 | Normal |
| | Pat 7 | 25% | 35% | 57% | 85% | 3403129-3461159 | Normal |
| | Misc 1 | 21% | 73% | 57% | 36% | 174580-199440 | Normal |
| | Misc 2 | 28% | 28% | 64% | 68% | 939710-987780 | Normal |

Continua

| Estirpe | Ilha genômica | Desvio de GC | Continuação | | | Posição | Escore |
|-------------|---------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|---------------|
| | | | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_11 | Res 1 | 14% | 0% | 42% | 71% | 47614-57258 | Normal |
| | Res 2 | 21% | 27% | 14% | 63% | 1204262-1284068 | Normal |
| | Pat 1 | 15% | 28% | 58% | 69% | 116993-145639 | Normal |
| | Pat 2 | 30% | 0% | 76% | 38% | 773514-785536 | Normal |
| | Pat 3 | 0% | 90% | 45% | 72% | 1167870-1179508 | Normal |
| | Pat 4 | 17% | 29% | 52% | 41% | 2109442-2135980 | Normal |
| | Pat 5 | 61% | 23% | 46% | 76% | 2295771-2305125 | Normal |
| | Pat 6 | 17% | 5% | 43% | 76% | 2319728-2367935 | Normal |
| | Pat 7 | 34% | 46% | 74% | 85% | 2414896-2447492 | Forte |
| | Misc 1 | 31% | 68% | 63% | 47% | 1135404-1153950 | Normal |
| | Misc 2 | 33% | 100% | 80% | 53% | 1760686-1785548 | Forte |
| | LGMAI_St_14 | Res 1 | 31% | 53% | 11% | 69% | 124909-188502 |
| Res 2 | | 14% | 0% | 42% | 71% | 825250-834894 | Normal |
| Res 3 | | 56% | 62% | 12% | 68% | 1878275-1895276 | Forte |
| Res 4 | | 77% | 33% | 11% | 55% | 2470033-2477686 | Normal |
| Pat 1 | | 21% | 0% | 78% | 47% | 458303-476783 | Normal |

Continua

| Estirpe | Ilha genômica | Desvio de GC | Conclusão | | | Posição | Escore |
|-------------|---------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_14 | Pat 2 | 38% | 21% | 51% | 65% | 690135-739229 | Normal |
| | Pat 3 | 30% | 0% | 76% | 46% | 1130321-1142343 | Normal |
| | Pat 4 | 0% | 100% | 45% | 81% | 1616117-1627755 | Normal |
| | Pat 5 | 8% | 14% | 79% | 35% | 1641675-1737549 | Normal |
| | Pat 6 | 3% | 0% | 73% | 70% | 1960979-1988147 | Normal |
| | Pat 7 | 61% | 0% | 46% | 76% | 2367253-2376665 | Normal |
| | Pat 8 | 23% | 26% | 80% | 83% | 2511034-2531428 | Normal |
| | Misc 1 | 33% | 100% | 80% | 53% | 1405406-1430268 | Forte |

Os resultados completos da análise de predição de ilhas genômicas, contendo todas as ilhas putativas identificadas, encontram-se disponíveis no anexo G.

A identificação e anotação de sequências de fagos com o servidor Phaster (Arndt *et al.*, 2016) não identificou nenhum fago completo nos 3 genomas analisados, mas identificou 5 sequências incompletas e 2 sequências questionáveis para a estirpe LGMAI_St_08, 6 sequências incompletas para a estirpe LGMAI_St_11 e 5 sequências incompletas para a estirpe LGMAI_St_14 (tabela 13).

Tabela 13: Resultados da análise de busca por sequências de fagos realizada com o servidor Phaster.

| Estirpe | Região (posição) | Tamanho | Escore | Palavras-chave | Total de proteínas | Proteínas de fago |
|----------------|---------------------------------------|----------------|---------------|--|---------------------------|--------------------------|
| LGMAI_St_08 | L2__13 (894-23399) | 22,5 kb | 20 | Portal | 12 | 7 |
| | L2__39 (6914-30874) | 23,9 kb | 40 | Cabeça Portal Capsídeo | 32 | 23 |
| | L2__42 (11768-24647) | 12,8 kb | 20 | Integrase | 21 | 16 |
| | L2__53 (1-21632) | 21,6 kb | 40 | Protease Cauda | 23 | 21 |
| | NODE_4_length_70496 (13705-70304) | 56,6 kb | 80 | Lisina Capsídeo Portal Cabeça | 45 | 32 |
| | NODE_6_length_66474 (52956-66473) | 13,5 kb | 70 | Transposase Integrase Cabeça | 19 | 11 |
| | NODE_18_length_47478 (22850-47478) | 24,6 kb | 40 | Cabeça Portal Capsídeo | 37 | 23 |
| LGMAI_St_11 | L2__19 (895-23400) | 22,5 kb | 20 | Portal | 12 | 7 |
| | L2__34 (1523-25483) | 23,9 kb | 40 | Capsídeo Portal Cabeça | 33 | 23 |
| | L2__39 (19190-30383) | 11,1 kb | 30 | Lisina Capsídeo | 11 | 9 |
| | L2__44 (1-21632) | 21,6 kb | 40 | Protease Cauda | 23 | 21 |
| | NODE_21_length_38741 (14316-38069) | 23,7 kb | 40 | Cabeça Portal Capsídeo | 32 | 22 |

Continua

| Estirpe | Região (posição) | Conclusão | | Palavras-chave | Total de proteínas | Proteínas de fago |
|-------------|---------------------------------------|-----------|--------|------------------------------|--------------------|-------------------|
| | | Tamanho | Escore | | | |
| LGMAL_St_11 | NODE_12_length_47836 (22934-46894) | 23,9 kb | 40 | Cabeça Portal Capsídeo | 35 | 22 |
| LGMAL_St_14 | L2__20 (895-23400) | 22,5 kb | 20 | Portal | 12 | 7 |
| | L2__35 (1523-25276) | 23,7 kb | 40 | Capsídeo Portal Cabeça | 32 | 22 |
| | L2__40 (2-11195) | 11,1 kb | 30 | Capsídeo Lisina | 10 | 9 |
| | L2__44 (1-21632) | 21,6 kb | 40 | Protease Cauda | 23 | 21 |
| | NODE_13_length_47932 (941-24901) | 23,9 kb | 40 | Capsídeo Portal Cabeça | 34 | 23 |

OBS: As regiões de sequências de fagos identificadas são caracterizadas de acordo com os seguintes critérios: intacta (>90); questionável (70-90); e incompleta (<70).

4.5 BUSCA POR GENES E CLUSTERES GÊNICOS ASSOCIADOS A PEPTÍDEOS ANTIMICROBIANOS

Na busca por metabólitos secundários com o banco de dados antiSMASH, foram identificadas 4 regiões relacionadas a metabólitos secundários na estirpe LGMAI_St_08 e 5 regiões nas estirpes LGMAI_St_11 e LGMAI_St_14 (tabela 14).

Tabela 14: Resultados dos clusters de metabólitos secundários obtidos na análise com a ferramenta antiSMASH.

| Estirpe | Tipo (Cluster) | Região | Localização | Similaridade |
|---------------|------------------------|----------------------|---------------|--------------|
| LGMAI_St_08 | T3PKS | L2_4 | 42.489-77.407 | NA |
| | NRPS | NODE_2_length_78597 | 33.066-78.597 | 28% |
| | (Equibactina) | | | |
| | Lantipeptídeo classe V | NODE_4_length_70496 | 46.265-70.496 | NA |
| LGMAI_St_11 | Arilpoliene | NODE_12_length_51642 | 1-31.620 | NA |
| | T3PKS | L2_2 | 45.75080.611 | NA |
| | NRPS | L2_5 | 46.815-97.343 | 21% |
| | (Equibactina) | | | |
| | Lantipeptídeo classe V | NODE_21_length_38741 | 1-24.542 | NA |
| | Arilpoliene | NODE_9_length_51476 | 20.026-51.476 | NA |
| | NRPS | NODE_2_length_133314 | 36.525-87.053 | 21% |
| (Equibactina) | | | | |
| LGMAI_St_14 | T3PKS | L2_2 | 45.750-80.145 | NA |
| | Arilpoliene | L2_9 | 1-31.818 | NA |
| | Lantipeptídeo classe V | L2_35 | 15.050-39.667 | NA |
| | NRPS | NODE_1_length_132751 | 35.487-86.015 | 21% |
| | (Equibactina) | | | |
| | NRPS | NODE_1_length_132274 | 35.487-86.015 | 21% |
| (Equibactina) | | | | |

Já na busca por clusteres gênicos de peptídeos antimicrobianos na ferramenta BAGEL4, foi identificada 1 área de interesse em cada uma das 3 estirpes, compreendendo a bacteriocina zoocina A (tabela 15).

Tabela 15: Resultados gerais da busca por clusteres de peptídeos antimicrobianos, realizada com a ferramenta BAGEL4.

| Informações gerais | Estirpes | | |
|-----------------------|----------------------|---------------|---------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| Área de interesse | NODE_1_length_142099 | L2__8 | L2__7 |
| Localização | 86054-106873 | 3236-24055 | 35273-56092 |
| Classe | 93.3;Zoocin_A | 93.3;Zoocin_A | 93.3;Zoocin_A |

Nas 3 áreas de interesse identificadas pela ferramenta BAGEL4, foi anotado o peptídeo *core* do antimicrobiano zoocina A (tabela 16), originalmente identificado em *Streptococcus equi* subsp. *zoepidemicus*. Todas as ORFs e miniORFs identificadas pela ferramenta BAGEL4 encontram-se disponíveis no anexo E.

Tabela 16: Recorte dos resultados da anotação de clusteres gênicos de peptídeos antimicrobianos, realizada na ferramenta BAGEL4.

| Estirpe | Nome da ORF | Localização | Fita | Identidade | E-value |
|-------------|---------------|-------------|------|------------|---------|
| LGMAI_St_08 | 93.3;Zoocin_A | 95981-96880 | + | 58,76% | 3e-117 |
| LGMAI_St_11 | 93.3;Zoocin_A | 13162-14061 | + | 58,76% | 3e-117 |
| LGMAI_St_14 | 93.3;Zoocin_A | 45198-46097 | + | 58,76% | 3e-117 |

A análise no banco de dados BACTIBASE gerou resultados que corroboram com os resultados obtidos pela ferramenta BAGEL4 e sugerem que os peptídeos identificados são muito similares ou idênticos entre si (tabela 17).

Tabela 17: Resultados da análise com o banco de dados de bacteriocinas BACTIBASE.

| Estirpe | Alinhamentos significativos | Tamanho | Pontuação (bits) | Cobertura | E-value | Identidade |
|-------------|--------------------------------|---------|---------------------|-----------|--------------|------------|
| LGMAI_St_08 | Zoocina A (BAC198) | 299 | 338,689 | 89% | 3,44045e-101 | 60% |
| LGMAI_St_11 | Zoocina A (BAC198) | 299 | 338,689 | 89% | 3,44045e-101 | 60% |
| LGMAI_St_14 | Zoocina A (BAC198) | 299 | 338,689 | 89% | 3,44045e-101 | 60% |

4.6 ANÁLISE DOS GENES E CLUSTERES GÊNICOS ASSOCIADOS A PEPTÍDEOS ANTIMICROBIANOS

O alinhamento das sequências identificadas pelo banco de dados BAGEL4 evidenciou que os peptídeos zoocina A, anotados nos três genomas, são 100% idênticos entre si (figura 16). A partir dessa etapa, a sequência é, então, tratada como uma só.

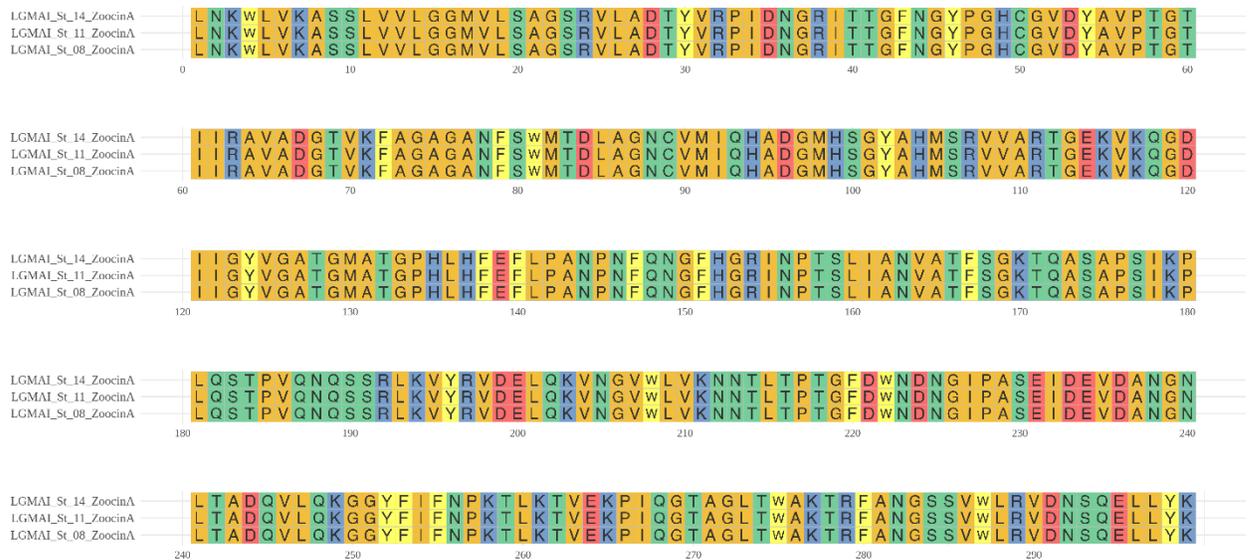


Figura 16: Alinhamento com Clustal X das sequências de aminoácidos do peptídeo antimicrobiano LGMAI zoocina A (dividida em cinco partes consecutivas por falta de espaço). No eixo y, encontram-se as estirpes LGMAI_St_14, LGMAI_St_11 e LGMAI_St_08 de cima para baixo. Os aminoácidos são coloridos de acordo com a química de suas cadeias laterais – amarelo: aromáticos; vermelho: ácidos; azul: básicos; laranja: não polares; verde: polares.

A análise no banco de dados Interpro identificou 1 região de peptídeo sinal e dois domínios na sequência da bacteriocina zoocina A anotada: o domínio catalítico de peptidase M23; e o domínio de reconhecimento do alvo da exoenzima lítica (tabela 18, figura 17 e figura 18).

Tabela 18: Resultados da análise de busca por domínios e região de peptídeo sinal identificados com o Interpro.

| | Regiões | | |
|-----------------------|----------------|---------------|--|
| | Peptídeo sinal | Peptidase M23 | Reconhecimento do alvo da exoenzima lítica |
| Banco de dados | Phobius | Pfam | Pfam |
| Acesso | SIGNAL_PEPTIDE | PF01551 | PF16775 |
| Localização | 1-27 | 48-141 | 194-298 |
| e-value | NA | 6,4e-30 | 5,6e-50 |

Tabela 19: Resultados do *blastx* entre os genomas e a sequência proteica do fator de imunidade *zif* de *Streptococcus equi* subsp. *zooepidemicus*.

| | Estirpe | | |
|-----------------------|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| Região | L2_68 | L2-56 | L2-55 |
| Posição | 8343-9572 | 8343-9572 | 1916-3145 |
| Fita | + | + | - |
| Cobertura | 100% | 100% | 100% |
| Identidade | 50% | 50% | 50% |
| Similaridade | 65% | 65% | 65% |
| <i>e-value</i> | 5e-123 | 5e-123 | 5e-123 |

O alinhamento das sequências de Zif mostrou que elas apresentam 410 resíduos de aminoácidos e são 100% idênticas entre si (figura 19).

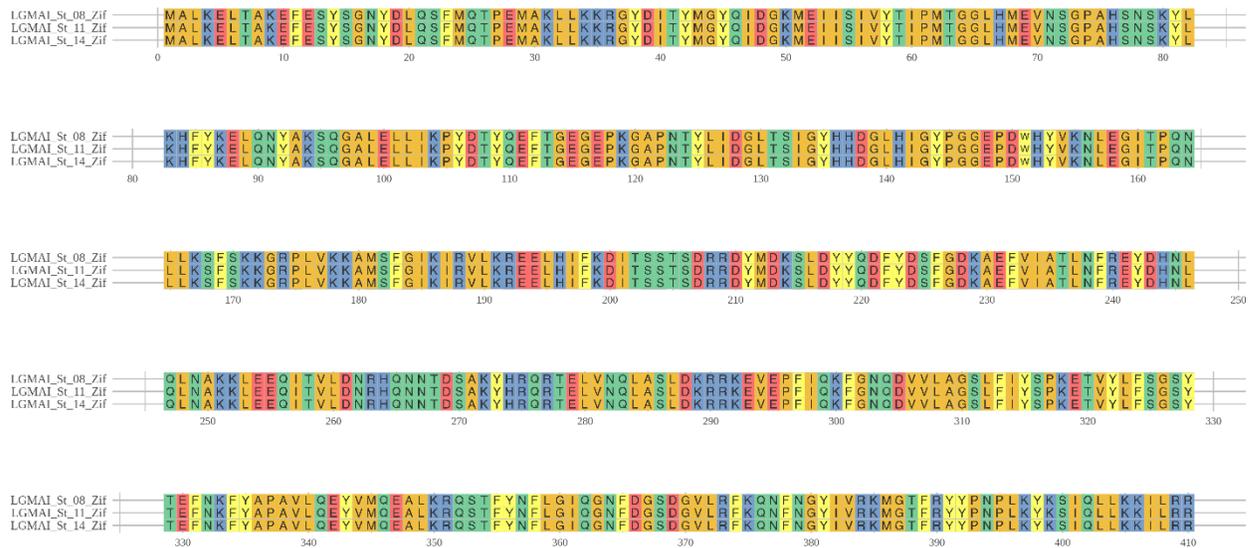


Figura 19: Alinhamento com Clustal X das sequências de aminoácidos do fator de imunidade *zif* (dividida em cinco partes consecutivas por questão de espaço). No eixo y, encontram-se as estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14 de cima para baixo. Os aminoácidos são coloridos de acordo com a química de suas cadeias laterais – amarelo: aromáticos; vermelho: ácidos; azul: básicos; laranja: não polares; verde: polares.

A análise da sequência do fator de imunidade zif no banco de dados do Interpro resultou na identificação de regiões correspondentes a superfamília e família proteicas, sítio de ligação com braço de tRNA e estrutura de α -hélice (tabela 20).

Tabela 20: Regiões de superfamília e família proteicas, sítio de ligação com braço de tRNA e estrutura de α -hélice identificadas na sequência do fator de imunidade putativo zif pelo banco de dados Interpro.

| Regiões | Banco de dados | Acesso | Localização | e-value |
|--|----------------|----------|-------------|-----------|
| Família FemABX de peptidil-transferase | Pfam | PF02388 | 6-406 | 1,1 e-147 |
| Superfamília acil-CoA N-aciltransferase | Superfamily | SSF55729 | 1-159 | 5,3 e-43 |
| Superfamília acil-CoA N-aciltransferase | | | 163-403 | 2,7 e-28 |
| Classes I e II de aminoacil-tRNA sintetase (braço de ligação de tRNA) | | SSF46589 | 239-303 | 7,3 e-7 |
| α -hélice | Coils | Coil | 243-263 | NA |

4.7 ANOTAÇÃO FUNCIONAL E GENÔMICA COMPARATIVA

A anotação de clusteres de grupos ortólogos com a ferramenta *eggNOG-mapper* (Cepas *et al.*, 2017) identificou 2.518 grupos de ortólogos para a estirpe LGMAI_St_08, 2.418 para a estirpe LGMAI_St_11 e 2.423 para a estirpe LGMAI_St_14. Considerando os grupos identificados, a categoria mais abundante nos três genomas foi a de metabolismo (tabela 21).

Tabela 21: Ocupação dos grupos de ortólogos identificados pelo *eggNOG-mapper* nas 4 grandes categorias.

| Categorias | Estirpe | | |
|---|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| Processos e sinalização celular | 556 | 511 | 505 |
| Processamento e armazenamento de informação | 601 | 589 | 586 |
| Metabolismo | 828 | 826 | 852 |
| Pouco caracterizado | 533 | 492 | 480 |

Já as categorias funcionais mais expressivas foram as de função desconhecida (S), replicação, recombinação e reparo (L) e transcrição (K) (figura 20). Além disso, não foram designadas proteínas para as categorias de predição de função geral (R), estrutura e dinâmica de cromatina (B), citoesqueleto (Z) e estrutura nuclear (Y) em nenhum dos três genomas analisados (figura 20).

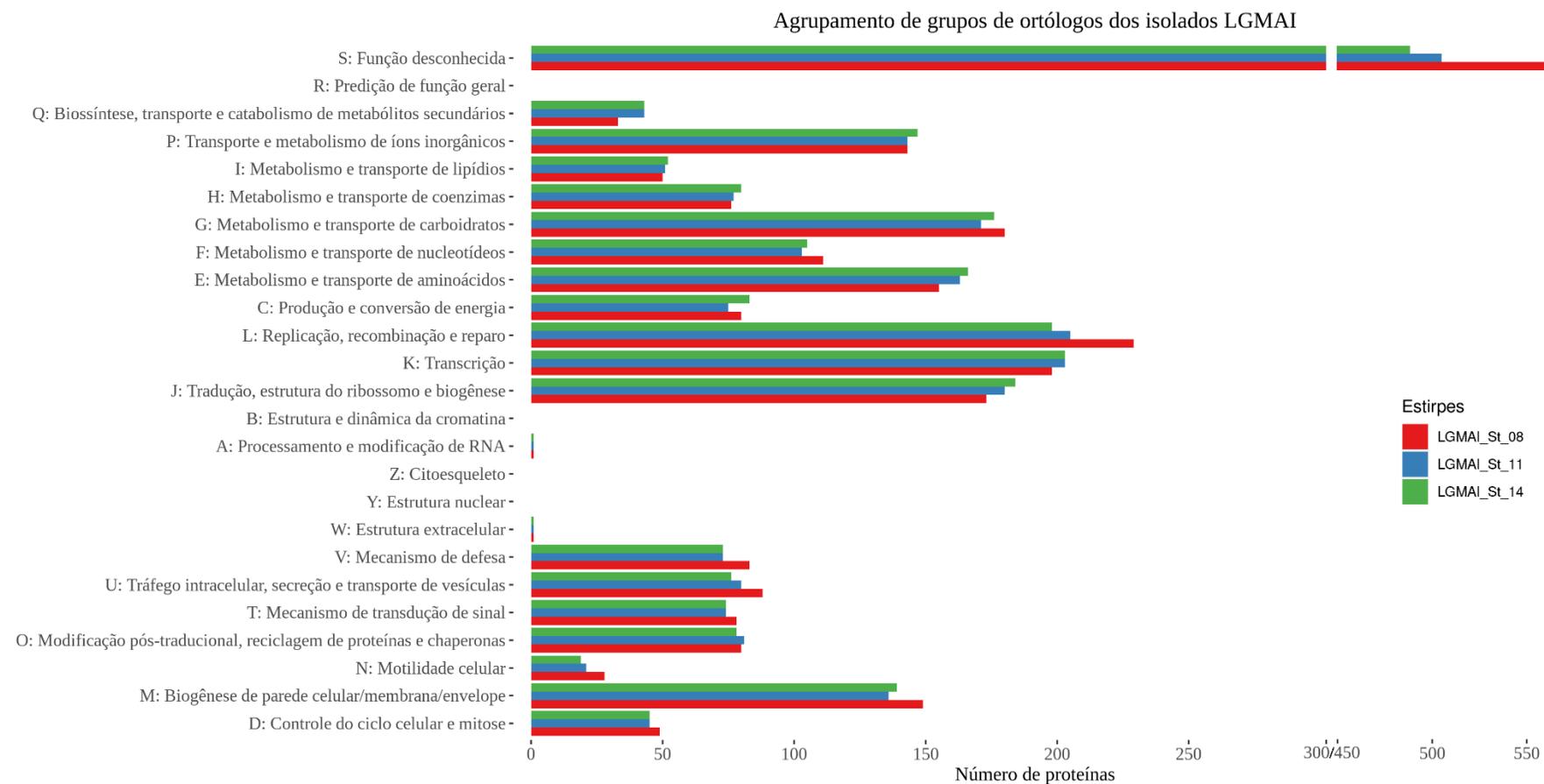


Figura 20: Anotação funcional de grupos ortólogos para as estirpes LGMAI. No eixo X, encontram-se as diferentes categorias de grupos de ortólogos. No eixo Y, encontra-se o quantitativo de proteínas identificadas para cada categoria. As barras possuem três colorações distintas que correspondem às três estirpes LGMAI: LGMAI_St_08 (em vermelho); LGMAI_St_11 (em azul); e LGMAI_St_14 (em verde). O gráfico de barras passou por um corte na porção de 300 a 450 para melhor visualização de categorias menos expressivas quantitativamente.

A anotação de clusteres de grupos ortólogos funcionais KEGG com o servidor de anotação automática de genomas e reconstrução de vias metabólicas (KAAS) (Moriya et al., 2007) identificou 657 grupos de ortólogos KEGG para a estirpe LGMAI_St_08, 133 para a estirpe LGMAI_St_11 e 122 para a estirpe LGMAI_St_14. Os grupos de ortólogos KEGG identificados distribuem-se em distintas categorias de vias metabólicas (tabela 22).

Tabela 22: Ocupação dos grupos de ortólogos KEGG identificados pelo servidor KAAS nas 6 grandes categorias.

| Categorias | Estirpe | | |
|---------------------------------------|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| Metabolismo | 823 | 200 | 195 |
| Processamento de informação genética | 63 | 28 | 30 |
| Processamento de informação ambiental | 73 | 7 | 10 |
| Processos celulares | 59 | 11 | 5 |
| Sistemas orgânicos | 12 | 4 | 2 |
| Doenças humanas | 38 | 9 | 5 |

Já as categorias funcionais KEGG mais expressivas foram as de mapas gerais e globais e metabolismo de carboidrato (figura 21). Além disso, a estirpe LGMAI_St_14 teve um total de 26 grupos funcionais KEGG identificados, enquanto que a estirpe LGMAI_St_11 teve um total de 30 grupos funcionais KEGG identificados, apresentando exclusivamente o grupo funcional de sistema nervoso. E, por fim, a estirpe LGMAI_St_08 obteve um total de 37 grupos funcionais KEGG identificados, possuindo exclusivamente os grupos funcionais de sistema imune, adaptação ao ambiente, câncer – tipos específicos, doenças cardiovasculares, doenças endócrinas e metabólicas, doenças infecciosas virais e resistência a antimicrobianos (figura 21).

Classificação de ortologia KEGG das estirpes LGMAI

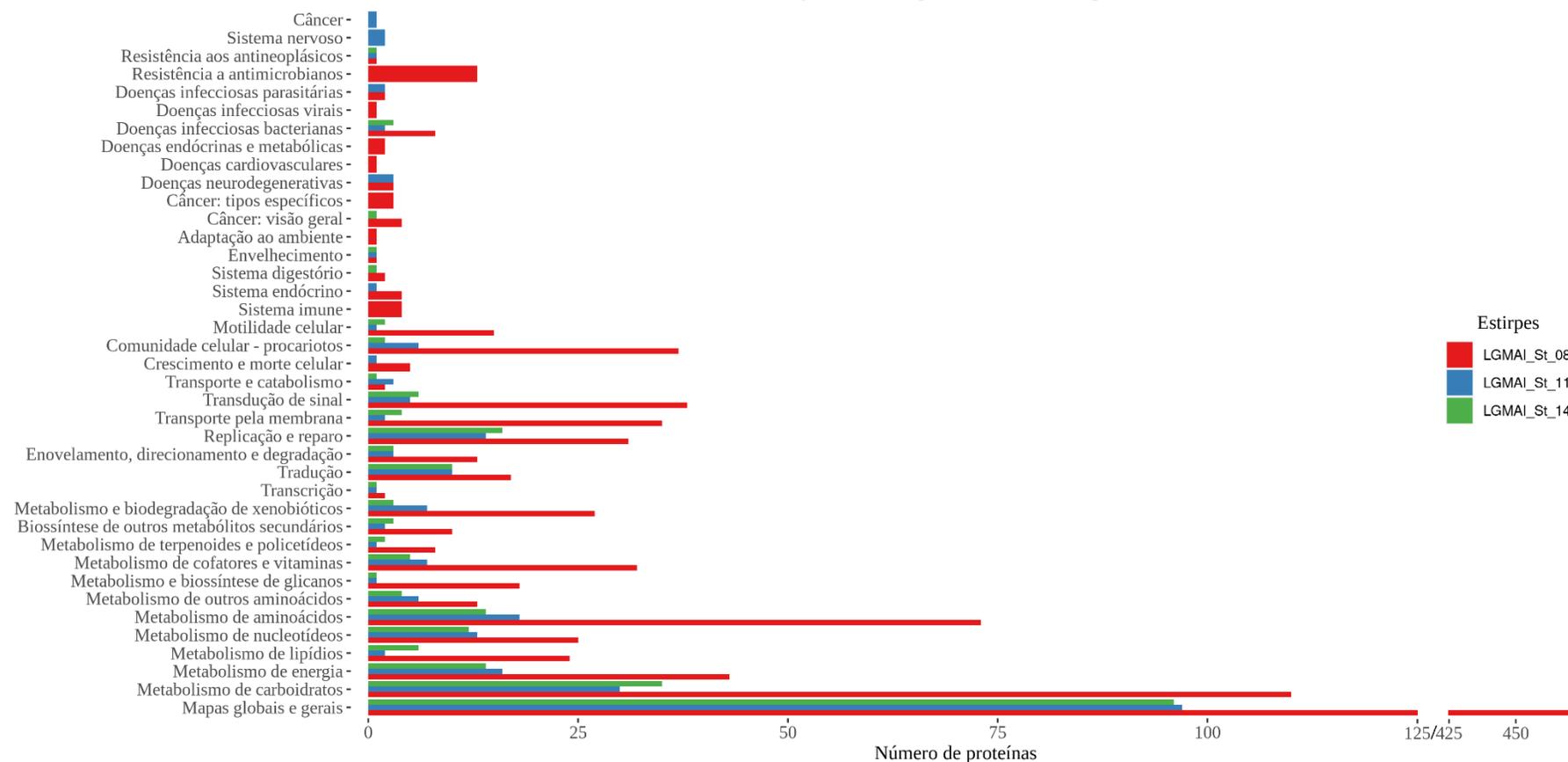


Figura 21: Anotação funcional de grupos ortólogos KEGG para as estirpes LGMAI. No eixo y, estão indicados os grupos funcionais identificados. No eixo X, encontra-se a quantidade de proteínas identificadas para cada grupo funcional. As barras possuem três colorações distintas que correspondem às três estirpes LGMAI: LGMAI_St_08 (em vermelho); LGMAI_St_11 (em azul); e LGMAI_St_14 (em verde). O gráfico de barras passou por um corte na porção de 125 a 425 para melhor visualização de categorias menos expressivas quantitativamente

A comparação entre os três genomas das estirpes LGMAI, realizada com a ferramenta OrthoVenn2 (Xu *et al.*, 2019), mostrou que, no geral, eles apresentam grande similaridade uns com os outros, compartilhando entre si 2248 grupos ortólogos. A respeito do compartilhamento entre pares de genomas, foi observado que as estirpes LGMAI_St_11 e LGMAI_St_14 são o par de genomas que compartilham, relativamente, o maior número de grupos ortólogos (25) um com o outro. Além disso, nem a estirpe LGMAI_St_11 nem a LGMAI_St_14 possuem grupos exclusivos; enquanto que a estirpe LGMAI_St_08 apresenta 69 grupos exclusivos (figura 22).

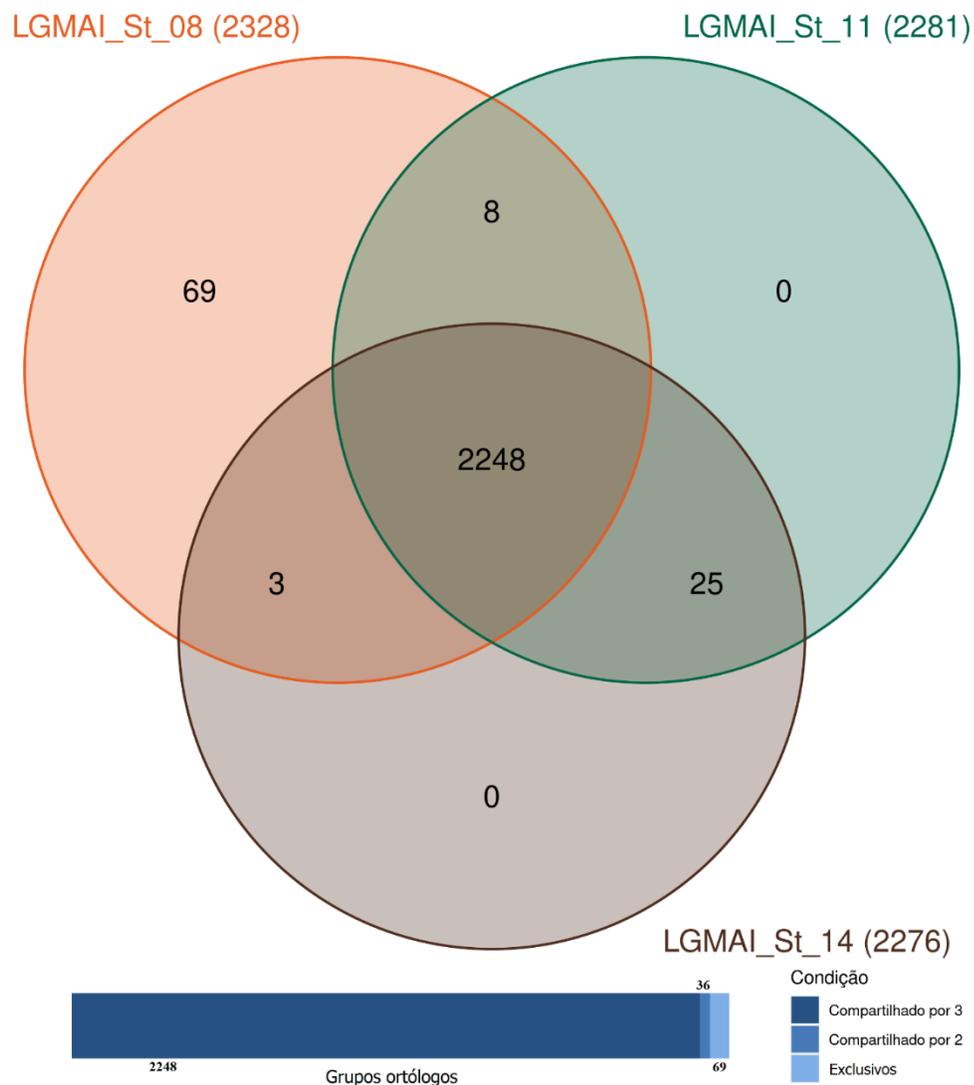


Figura 22: Diagrama de Venn, mostrando os resultados gerais da análise de genômica funcional comparativa das estirpes LGMAI. As cores vermelha, verde e marrom indicam os conjuntos dos genomas das estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14, respectivamente. O gráfico de barras mostra a distribuição e a quantidade dos grupos ortólogos que são compartilhados por 3 e 2 genomas e os genes que são específicos.

Dos 2248 grupos de ortólogos compartilhados pelas 3 estirpes LGMAI, foram identificadas um total de 7809 proteínas, das quais são 2597 (33,26%) de LGMAI_St_14, 2606 (33,37%) de LGMAI_St_11 e 2606 (33,37%) de LGMAI_St_08. Dos 8 grupos compartilhados pelas estirpes LGMAI_St_11 e LGMAI_St_08, 16 proteínas foram identificadas, das quais são 8 (50%) oriundas de cada estirpe. Dos 3 grupos compartilhados pelas estirpes LGMAI_St_14 e LGMAI_St_08, foram identificadas 6 proteínas, das quais são 3 (50%) oriundas de cada estirpe. Ainda, dos 25 grupos compartilhados pelas estirpes LGMAI_St_14 e LGMAI_St_11, foram anotadas 55 proteínas, das quais são 26 (47,27%) da estirpe LGMAI_St_14 e 29 (52,73%) da estirpe LGMAI_St_11. Por fim, dos 69 grupos identificados exclusivamente na estirpe LGMAI_St_08, foram contabilizadas 139 proteínas (figura 23).

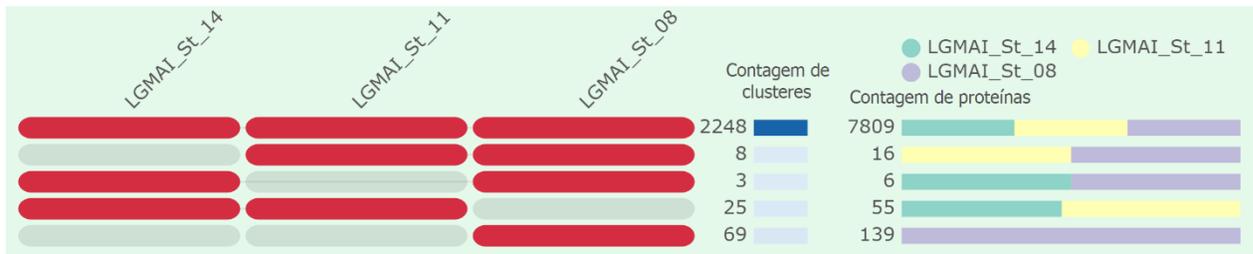


Figura 23: Distribuição da contagem de clusters e de proteínas identificados pela ferramenta de comparação e anotação de genomas OrthoVenn2.

Mais especificamente em relação aos grupos funcionais identificados, observou-se que a estirpe LGMAI_St_08 não apresenta nenhum gene exclusivo de função molecular nem compartilha, par a par, genes de função molecular com as outras [figuras 24 e 25 (A.)]. A interseção dos conjuntos de ortólogos identificados para as estirpes LGMAI_St_08 e LGMAI_St_14 é a menor de todas e apresenta somente genes para processos biológicos com função de transposição [figura 25 (A.)]. O grupo de função molecular foi o mais conservado de todos, sendo identificado apenas na interseção entre os três genomas (figura 26) e entre os genomas das estirpes LGMAI_St_11 e LGMAI_St_14 na categoria do subgrupo funcional de atividade de transferase [figura 25 (B.)]. Para o grupo funcional de componente celular, também foi observada conservação de genes do subgrupo funcional de parte de célula e membrana (figura 26).

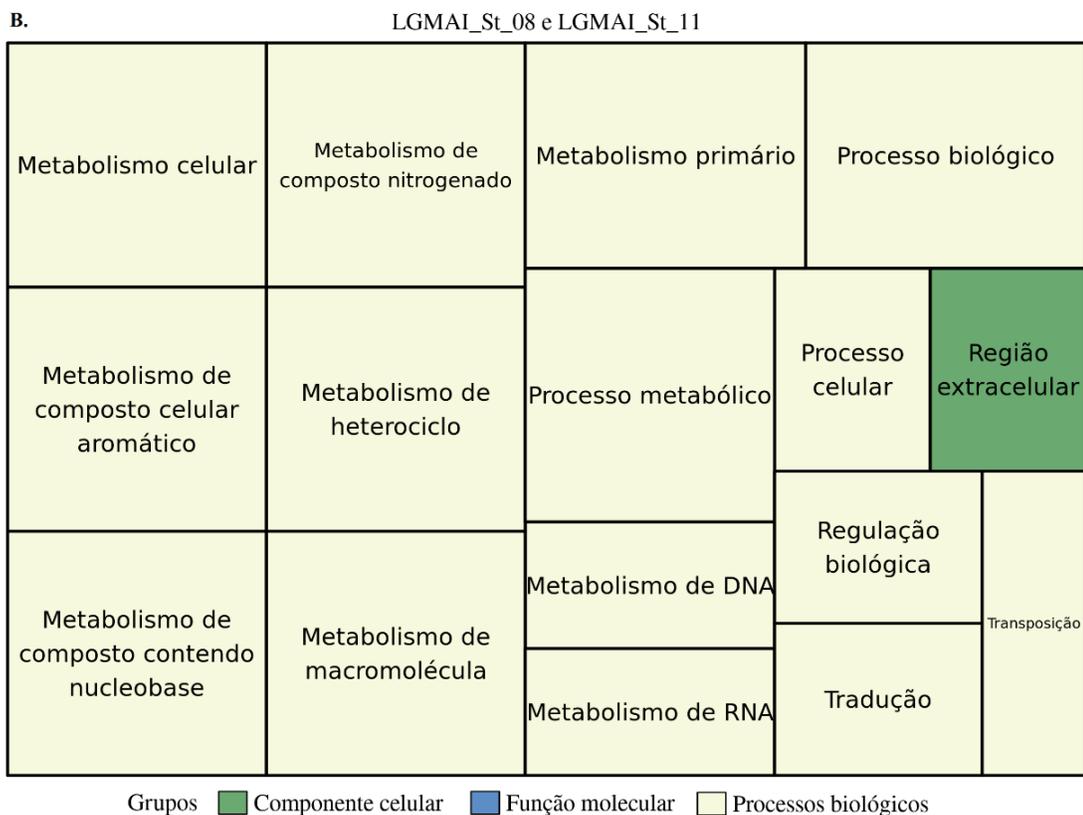
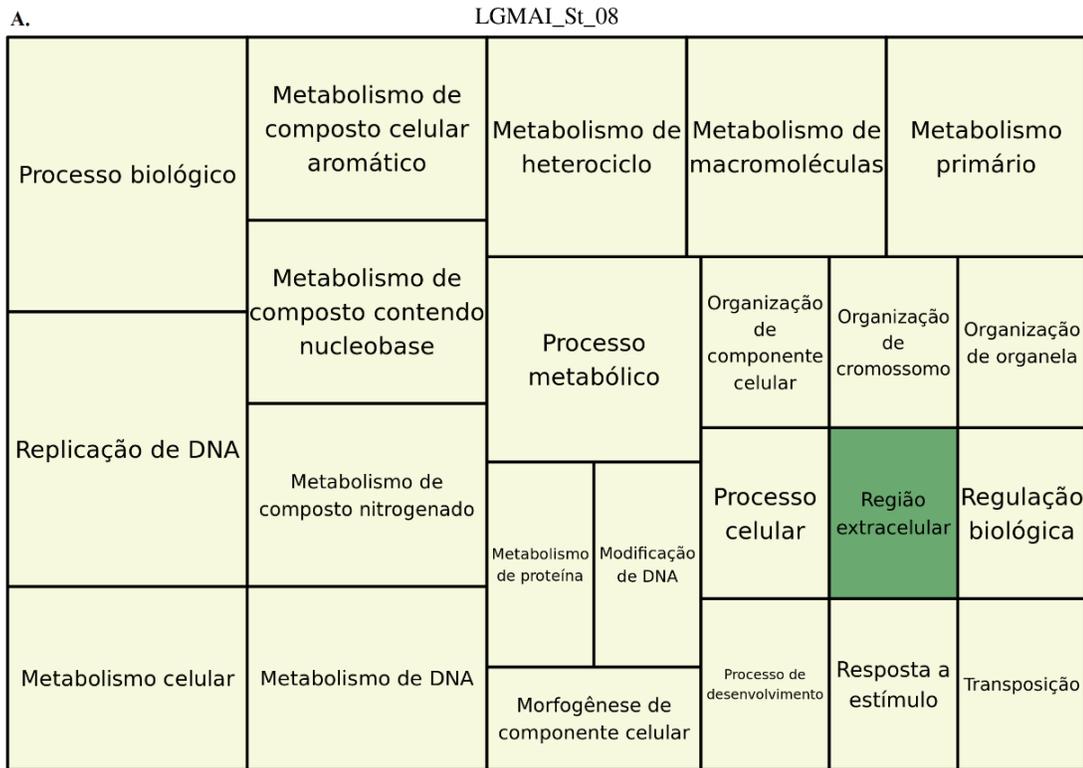


Figura 24: Mapa hierárquico dos grupos e subgrupos funcionais anotados (A) exclusivamente para a estirpe LGMAI_St_08 e (B) compartilhados entre as estirpes LGMAI_St_08 e LGMAI_St_11. As cores representam grupos funcionais de componente celular (verde), função molecular (azul) e processos biológicos (bege).

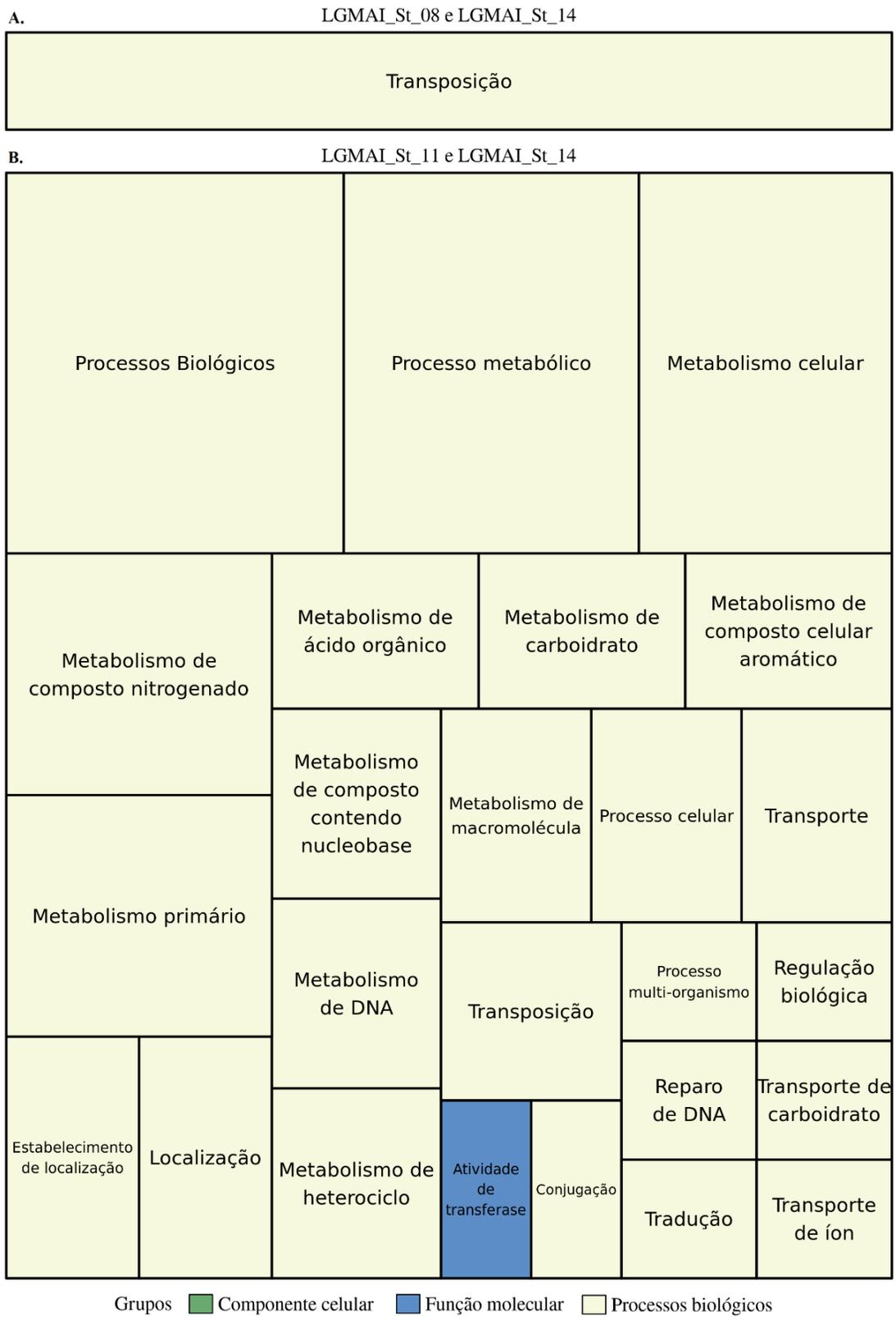
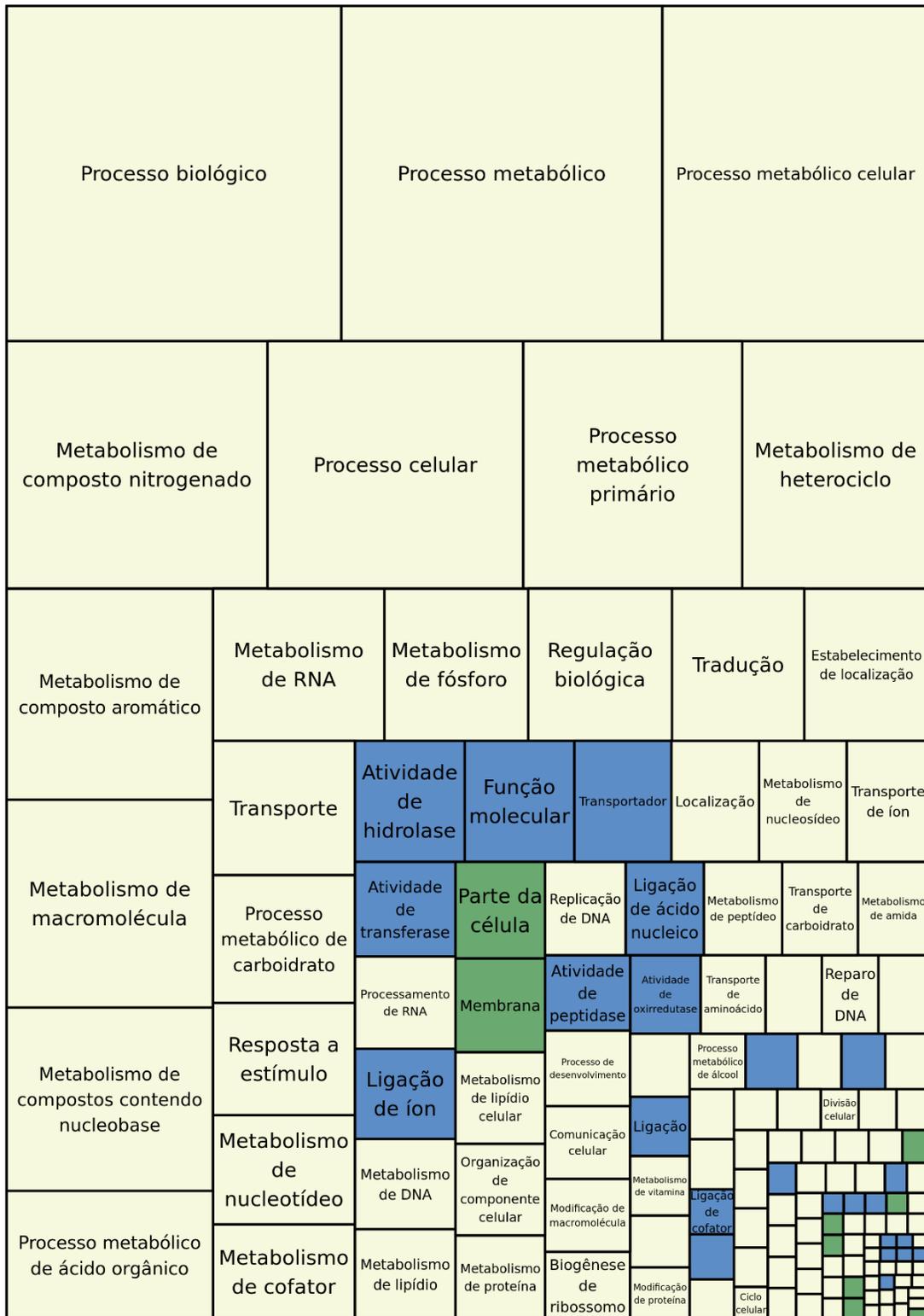


Figura 25: Mapa hierárquico dos grupos e subgrupos funcionais anotados (A) compartilhados entre as estirpes LGMAI_St_08 e LGMAI_St_14 e (B) compartilhados entre as estirpes LGMAI_St_11 e LGMAI_St_14. As cores representam grupos funcionais de componente celular (verde), função molecular (azul) e processos biológicos (bege).

LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14



Grupos ■ Componente celular ■ Função molecular ■ Processos biológicos

Figura 26: Mapa hierárquico dos grupos e subgrupos funcionais anotados que são compartilhados pelas três estirpes LGMAI. As cores representam grupos funcionais de componente celular (verde), função molecular (azul) e processos biológicos (bege). O tamanho dos retângulos varia conforme a quantidade de genes por subgrupo.

O mapa circular de identidade de nucleotídeos gerado pela ferramenta Brig mostrou que os genomas das três estirpes LGMAI apresentam alta similaridade entre si e possuem regiões que não estão presentes nos outros 15 genomas de *Streptococcus agalactiae* usados na análise. Algumas dessas regiões foram identificadas como ilhas de resistência e patogenicidade pelo software de predição de ilhas genômicas Gipsy (figura 27).

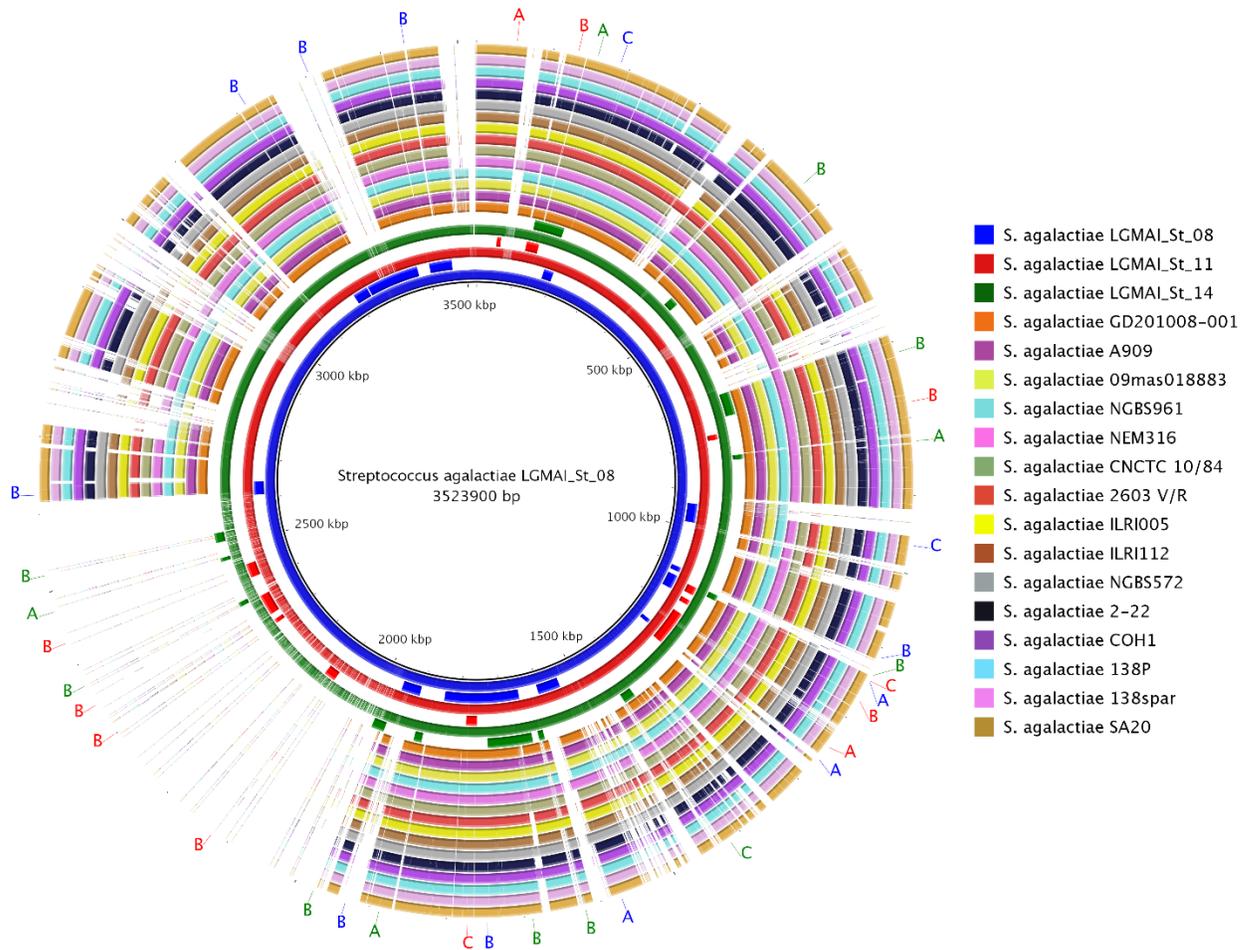


Figura 27: Mapa genômico circular contendo os 3 genomas das estirpes LGMAI e outros 15 genomas de *Streptococcus agalactiae*. De dentro para fora, os genomas usados na análise seguem a ordem presente na legenda à direita. Fora dos 3 grandes círculos mais internos (dos genomas LGMAI), estão apresentadas as ilhas genômicas preditas pelo software Gipsy. As ilhas possuem a mesma cor do genoma a que se referem. As letras ao redor do mapa circular representam (A) ilhas de resistência, (B) ilhas de patogenicidade e (C) ilhas miscelâneas, isto é, ilhas que contêm partes de resistência e partes de patogenicidade. A coloração das letras também está relacionada aos genomas a que se referem: azul – LGMAI_St_08; vermelho – LGMAI_St_11; e verde – LGMAI_St_14. Quanto mais intensa a coloração dos círculos, maior o grau de identidade em relação ao genoma de referência LGMAI_St_08.

Nos mapas circulares feitos com o *software CGView Comparison Tool* (Grant; Arantes; Stothard, 2012) são mostrados os resultados da anotação funcional de grupos COG e destacadas as anotações dos fatores de virulência e resistência, assim como os genes relacionados ao peptídeo antimicrobiano putativo zoocina A (figuras 28, 29 e 30).

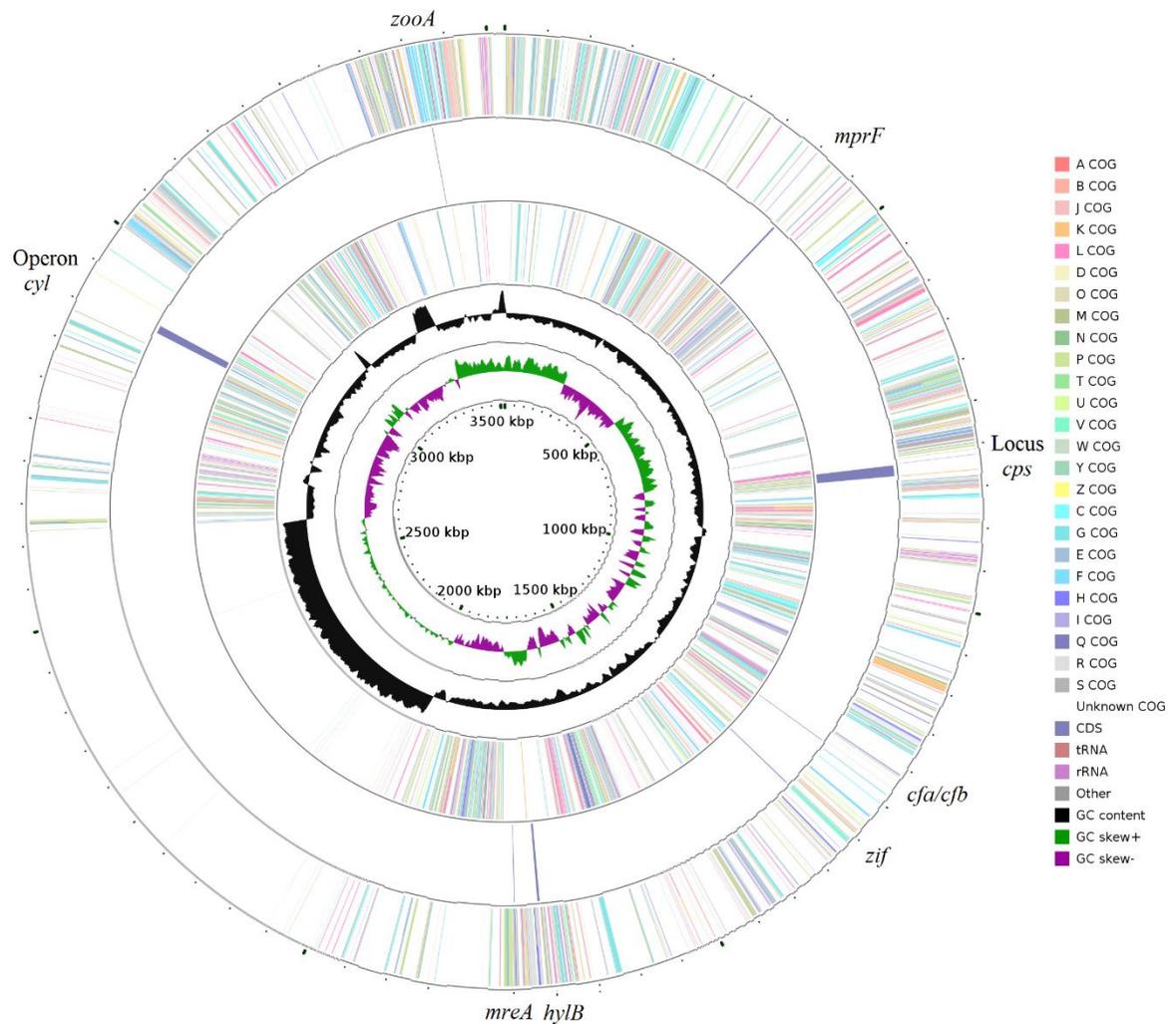


Figura 28: Mapa circular do genoma da estirpe LGMAI_St_08. De dentro para fora, encontram-se: régua com tamanho do genoma, conteúdo GC skew, conteúdo GC, CDS no sentido reverso, genes destacados e CDS no sentido pra frente. Os termos ao redor do mapa correspondem aos nomes dos genes destacados. Os genes destacados correspondem àqueles anotados nas ferramentas de busca por fatores de virulência, genes de resistência e genes de peptídeos antimicrobianos. As CDS apresentam coloração baseada no grupo funcional COG para o qual foram designadas (legenda à direita).

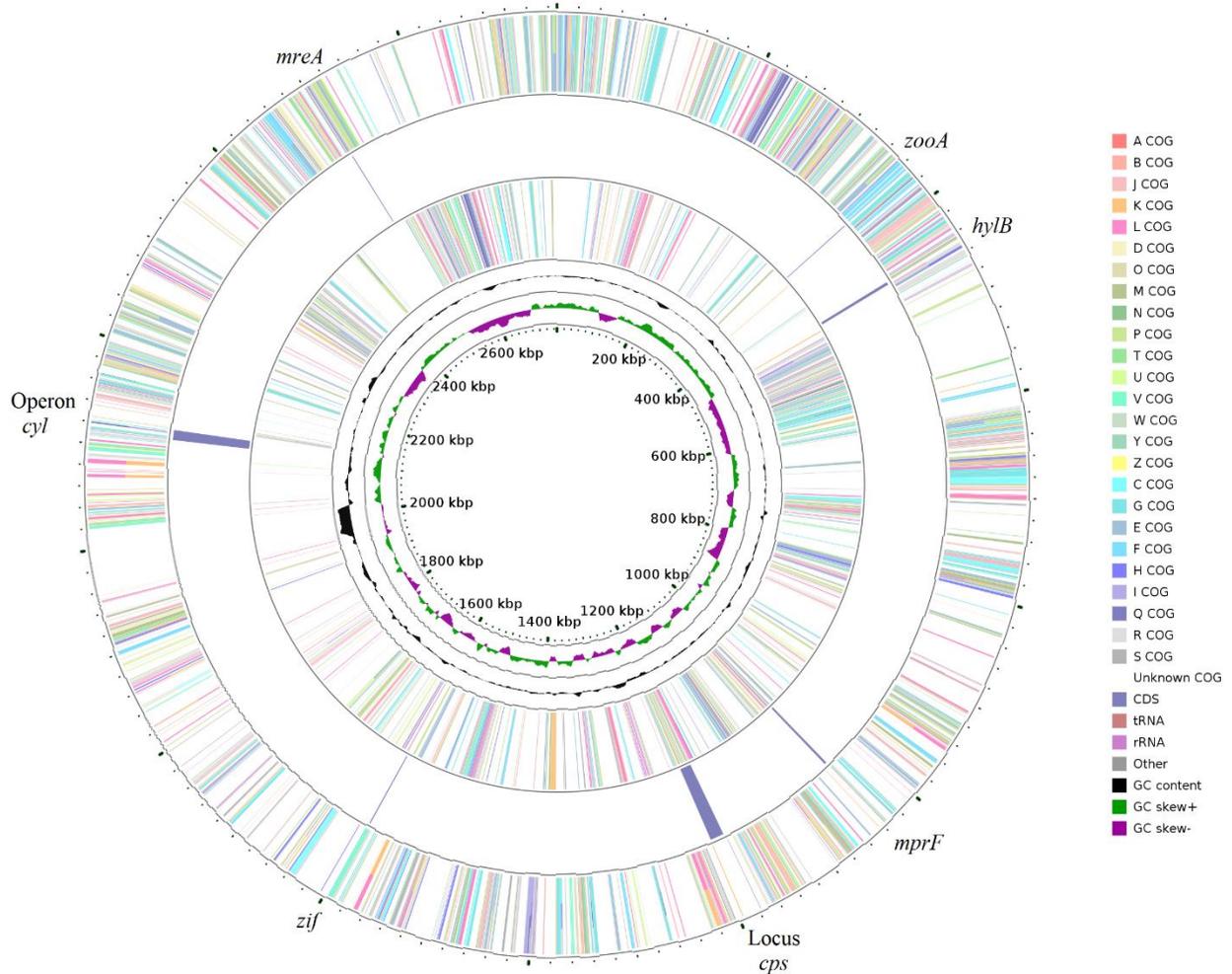


Figura 29: Mapa circular do genoma da estirpe LGMAI_St_11. De dentro para fora, encontram-se: régua com tamanho do genoma, conteúdo GC *skew*, conteúdo GC, CDS no sentido reverso, genes destacados e CDS no sentido pra frente. Os termos ao redor do mapa correspondem aos nomes dos genes destacados. Os genes destacados correspondem àqueles anotados nas ferramentas de busca por fatores de virulência, genes de resistência e genes de peptídeos antimicrobianos. As CDS apresentam coloração baseada no grupo funcional COG para o qual foram designadas (legenda à direita).

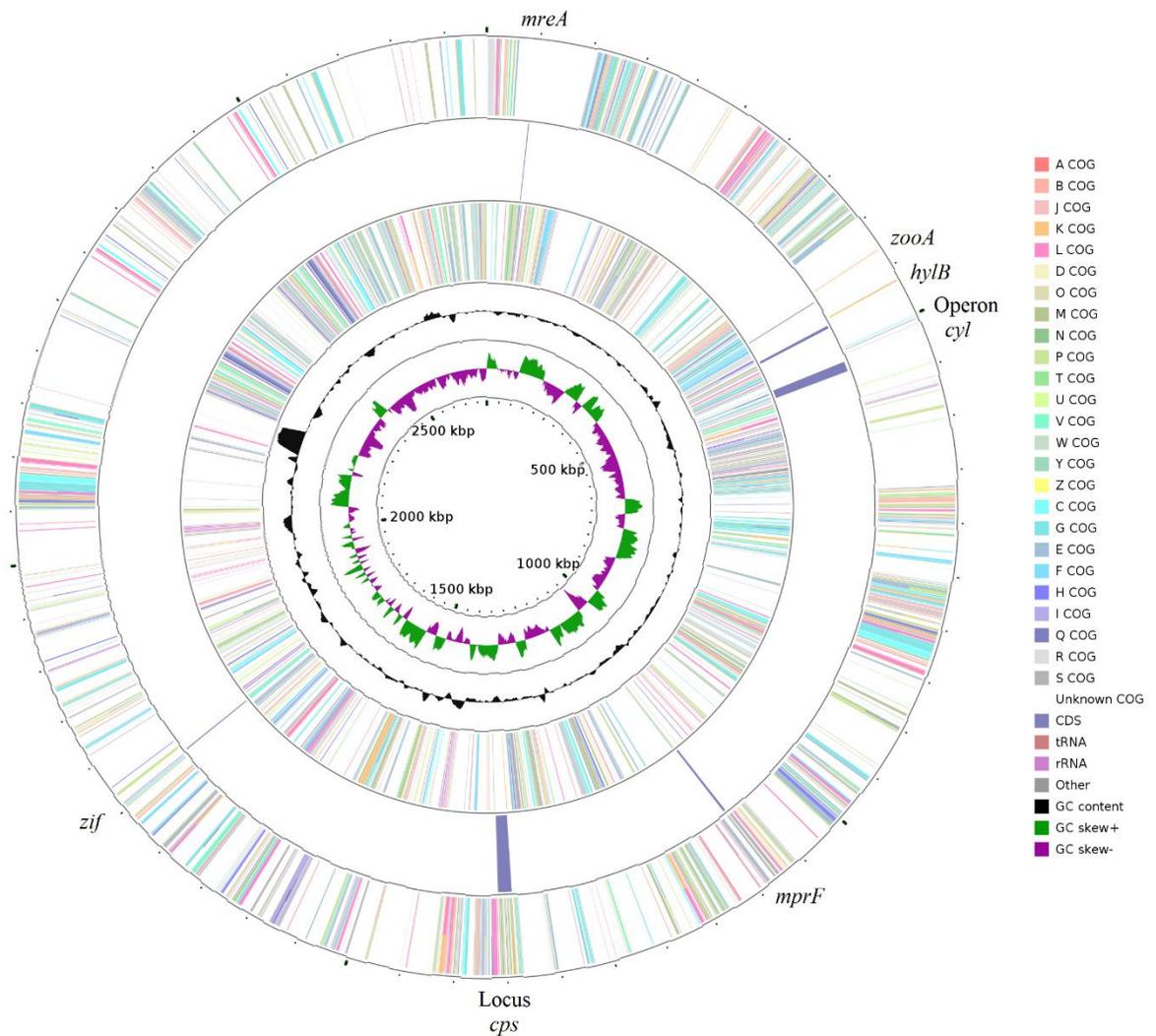


Figura 30: Mapa circular do genoma da estirpe LGMAI_St_14. De dentro para fora, encontram-se: régua com tamanho do genoma, conteúdo GC skew, conteúdo GC, CDS no sentido reverso, genes destacados e CDS no sentido pra frente. Os termos ao redor do mapa correspondem aos nomes dos genes destacados. Os genes destacados correspondem àqueles anotados nas ferramentas de busca por fatores de virulência, genes de resistência e genes de peptídeos antimicrobianos. As CDS apresentam coloração baseada no grupo funcional COG para o qual foram designadas (legenda à direita).

5 DISCUSSÃO

A atividade inibitória das estirpes de *Streptococcus agalactiae* do LGMAI contra *Streptococcus pyogenes* foi reavaliada e observou-se que ela se mantém após esses anos. Os testes adicionais de antibiose mostraram que as quatro estirpes de *Streptococcus agalactiae* LGMAI foram capazes de inibir *Cellulomonas fimi* e *Micrococcus* sp.. O gênero *Micrococcus* possui representantes responsáveis pela deterioração do leite, como *M. luteus*, *M. varians*, *M. freudenreichii* e *M. aloeverae*. Essas bactérias são termodúricas, isto é, são capazes de resistir à temperatura de pasteurização do leite e permanecer no produto, diminuindo assim seu tempo de vida de prateleira. Além disso, mesmo leites com alta qualidade (baixa contagem de células bacterianas) ainda há a persistência de alguns representantes, tornando necessária uma forma de controle adicional à pasteurização (Júnior *et al.*, 2018).

Embora menos frequentemente relatado, o gênero *Cellulomonas* também possui representantes que causam danos na indústria alimentícia, como é o caso de *Cellulomonas flavigena* causando deterioração de tecidos de plantações de oliveiras. Devido a sua atividade celulolítica, bactérias do gênero *Cellulomonas* são capazes de degradar materiais que contêm celulose, como papéis, géis de carboximetil celulose e diversos tecidos vegetais (Patel; Vaughn, 1973). A associação da atividade celulolítica dessas bactérias com o fato de elas serem frequentemente encontradas no solo, torna-as potenciais causadoras de danos em plantações, acarretando perdas econômicas para esse setor. Sendo assim, os resultados de inibição de *Micrococcus* sp. e *Cellulomonas fimi* são interessantes devido a sua aplicação potencial na conservação alimentícia e prevenção de perdas em massa de cultivares.

Além disso, elas não foram capazes de inibir as estirpes de *Klebsiella pneumoniae* e *Staphylococcus aureus* resistentes ou não a antimicrobianos. Tais resultados mostraram, então, que as substâncias antimicrobianas das estirpes de *Streptococcus agalactiae* isolados de mastite bovina não apresentam potencial aplicação como substâncias alternativas no controle desses microrganismos.

No entanto, em Silva (2003), foi observado que as estirpes de *Streptococcus agalactiae* do LGMAI tiveram resultados bastante promissores de antibiose, uma vez que se mostraram capazes de inibirem *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus bovis* e *Streptococcus uberis*, as quais são espécies causadoras de doenças em humanos e animais. Esse

espectro de inibição com foco em espécies do gênero *Streptococcus* é muito interessante, haja vista a grande importância clínica de diversas espécies desse gênero (Peters; Price; Llewelyn, 2017).

Streptococcus uberis é uma bactéria de colonização intramamária em bovinos. A espécie possui sorotipos de aspecto ambiental (D, E, F1, F2, G e H), os quais causam infecção intramamária em gado, mas estão associados a casos transitórios com cura espontânea. Por outro lado, a espécie também possui sorotipos contagiosos (A1, A2, B, C1 e C2), responsáveis por infecções persistentes que se desenvolvem em quadros clínicos de mastite bovina. O sucesso de infecção e desenvolvimento da doença está inteiramente associado à quantidade de microrganismos, à eficiência da defesa dos úberes e a fatores de risco ambientais (Leelahapongsathon *et al.*, 2020).

Streptococcus equinus, por sua vez, faz parte do complexo *Streptococcus bovis* e *Streptococcus equinus* (SBSEC), o qual acomoda 7 espécies não β -hemolíticas atualmente, tais quais *Streptococcus equinus*, *Streptococcus infantarius* subsp. *infantarius*, *Streptococcus lutetiensis*, *Streptococcus alactolyticus*, *Streptococcus gallolyticus* subsp. *gallolyticus*, *Streptococcus gallolyticus* subsp. *macedonicus* e *Streptococcus gallolyticus* subsp. *pasteurianus*. Essa bactéria é comensal e é um patógeno oportunista, encontrado tanto em humanos quanto em animais (Jans *et al.*, 2015). Os mecanismos que permeiam a transição dessas bactérias de comensais para patogênicas ainda são pouco elucidados. Elas são encontradas colonizando rúmen, bolsas sublinguais e cloacas em animais, ao passo que são encontradas no cólon de seres humanos. São responsáveis principalmente por casos de mastite em animais (Pompilio; Bonaventura; Gherardi, 2019). No entanto, ganharam relevância clínica por casos de infecção em humanos nos últimos anos, envolvidas em endocardite infecciosa e lesões de cólon, e apresentando forte associação com câncer colorretal (Jans *et al.*, 2015).

Uma característica muito preocupante de bactérias pertencentes ao complexo SBSEC são suas elevadas taxas de resistência antimicrobiana. Dentro do gênero *Streptococcus*, caracterizam-se por apresentarem resistência antimicrobiana reportada contra diversos antibióticos, como clindamicina, eritromicina, tetraciclina, ceftriaxona, linezolida, combinação de quinupristina-dalfopristina, penicilina e vancomicina. Além disso, atuam como reservatórios de genes de resistência antimicrobiana adquiridos (Pompilio; Bonaventura; Gherardi, 2019). Desse modo, a capacidade de inibição de *Streptococcus bovis* pelas estirpes LGMAI apresenta grande atrativo na contenção de doenças causadas pelas bactérias pertencentes ao complexo SESBC.

A terapia com antibióticos ainda é uma das principais formas de tratamento para casos de mastite bovina. Tanto *Streptococcus uberis*, como *Streptococcus equinus* e *Streptococcus agalactiae* são responsáveis por infecções de mastite em gado (Leelahapongsathon *et al.*, 2020; Pompilio; Bonaventura; Gherardi, 2019; Raabe; Shane, 2019). Em decorrência do aumento da conscientização a respeito da resistência antimicrobiana, torna-se necessário o uso de terapias alternativas aos antibióticos (Angelopoulou *et al.*, 2019). É interessante especular sobre a aplicação tópica de bacteriocinas para proteger os úberes do gado, uma vez que essa é a aplicação mais viável para peptídeos antimicrobianos (Koo; Seo, 2019). Testes com soluções desinfetantes contendo nisina e lenços umedecidos foram realizados e obtiveram resultados satisfatórios na limpeza das superfícies do úbere e consequente prevenção de infecções por bactérias causadoras de mastite em bovinos (Angelopoulou *et al.*, 2019).

A espécie *Streptococcus pyogenes*, ou *Streptococcus* grupo A (GAS), acomoda bactérias patogênicas de humanos. Essa é a única espécie do grupo *pyogenes* de *Streptococcus* a ter o perfil de infecção restrita a seres humanos. Elas colonizam a mucosa da nasofaringe, a qual representa um sítio com baixa disponibilidade de carboidratos livres, dificultando a aquisição de nutrientes pelas bactérias. Por isso, *Streptococcus pyogenes* apresenta uma ilha genômica de patogenicidade composta por 35 genes, com destaque especial aos genes *speB* e *ropB*, que codificam uma exotoxina e seu regulador, respectivamente, e que permitem sua infecção restrita aos humanos e também sua adaptação ao ambiente desafiante da nasofaringe (Ijaz *et al.*, 2020).

Essas bactérias são uma das maiores causadoras de morte por doenças infecciosas, sendo responsáveis por aproximadamente meio milhão de mortes por ano. A infecção por *Streptococcus* grupo A pode causar quadros leves, como laringite, mas também evoluir para quadros mais severos, como septicemia, síndrome de choque tóxico, febre reumática e sua seqüela, a doença reumática cardíaca (LaRock *et al.*, 2020). *Streptococcus pyogenes* apresenta um nível preocupante de ameaça, devido ao aumento da incidência de resistência a eritromicina e clindamicina, limitando e dificultando o tratamento das infecções causadas por GAS (CDC, 2019), e aumentando o apelo pela descoberta de novos antimicrobianos capazes de inibir GAS.

Por fim, *Streptococcus agalactiae*, ou *Streptococcus* grupo B, são patógenos de animais e de seres humanos. Originalmente isolados de bovinos com casos clínicos de mastite, tiveram sua infecção em humanos notificada no ano de 1938 em casos fatais pós-parto. Apresentam 10 sorotipos, dos quais, os mais envolvidos em doenças infecciosas são Ia, Ib, II, III e V (Raabe;

Shane, 2019). São encontrados colonizando o trato gastrointestinal e o trato genital feminino, por meio da formação de estruturas similares ao biofilme e da aderência pilial nas superfícies mucosas (Peters; Price; Llewelyn, 2017).

São responsáveis por bacteremia, podendo evoluir para casos de endocardite, e por infecções de tecidos moles, como celulite, abscesso e úlcera de decúbito. *Streptococcus* grupo B também está ligado a casos de fasciíte necrosante, pneumonia, meningite e infecções do trato urinário. Mulheres grávidas estão no grupo de maior risco de infecções invasivas por GBS, principalmente pelos sorotipos Ia e III. Infecções invasivas em recém-nascidos apresentam alta mortalidade. Indivíduos adultos portadores de comorbidades, como doenças cardíacas, renais e hepáticas, distúrbios neurológicos, alcoolismo, tabagismo, obesidade, comprometimento do sistema imune e velhice também estão em risco de infecção (Raabe; Shane, 2019).

O antibiótico penicilina G é a primeira linha de tratamento contra infecção por GBS. *Streptococcus agalactiae* apresenta também suscetibilidade a outros β -lactâmicos, como ampicilina, cefalosporinas e carbapenemas. No entanto, em casos de alergia aos antibióticos β -lactâmicos, são usados clindamicina, eritromicina, fluoroquinolonas e vancomicina (Raabe; Shane, 2019). *Streptococcus agalactiae*, assim como *Streptococcus pyogenes*, apresenta um nível preocupante de ameaça, devido ao aumento da incidência de resistência a eritromicina e clindamicina, usados como segunda linha de tratamento contra infecção por GBS, aumentando apelo por maior consciência na forma como antibióticos são prescritos e na descoberta e desenvolvimento de novas estratégias de combate a essas infecções (CDC, 2019).

O genoma da estirpe LGMAI_St_03 não pôde ser sequenciado com sucesso, já que não apresentou qualidade suficiente para gerar um genoma confiável. Isso pode ter decorrido por conta de contaminantes residuais dos materiais usados na extração que não foram completamente removidos nas duas etapas de separação das fases aquosa e proteica por solução de clorofórmio:álcool isoamílico. A presença desses resíduos prejudica a qualidade do DNA para o sequenciamento.

As três estirpes restantes tiveram seus genomas sequenciados com êxito. Os três genomas apresentaram boas métricas de montagem, de acordo com ferramentas de avaliação de montagem e boa quantidade de proteínas preditas. Os genomas das estirpes LGMAI_St_08, LGMAI_St_11 e LGMAI_St_14 apresentaram, respectivamente, conteúdos GC de 36,5%, 35,8% e 35,7%, valores estes próximos aos 35,4% - conteúdo GC médio dos genomas disponíveis para *Streptococcus*

agalactiae no NCBI. Ademais, os genomas montados apresentam ótima completude, sem nenhum gene ortólogo universal evolutivamente conservado de procariontes faltante.

Streptococcus agalactiae ou *Streptococcus* grupo B é um dos principais patógenos causadores da mastite bovina (Castro *et al.*, 2017). Além disso é um patógeno também capaz de infectar humanos (Haenni; Lupo; Madec, 2018), caracterizando-se como um importante patógeno humano, uma vez que apresenta grande risco para mulheres em gestação e seus bebês, sendo uma das principais causas de morte por septicemia em recém-nascidos no Brasil e no mundo (Nascimento *et al.*, 2019). Considerando-se a importância clínica de bactérias dessa espécie, a contribuição desses genomas é importante, uma vez que eles podem ser aplicados em estudos de pangenômica para auxiliar no entendimento da dinâmica do fluxo gênico, especialmente dos aspectos da virulência e patogenicidade desses microrganismos, desvendando potenciais novos alvos para vacinas (Azarian; Huang; Hanage, 2020).

No entanto, os três genomas apresentam regiões de grande desvio de conteúdo GC, como pôde ser observado nos mapas genômicos circulares construídos para cada uma das estirpes. O genoma da estirpe LGMAI_St_08 foi o que apresentou a maior faixa com grande desvio de conteúdo GC observada, localizada entre 1.968.713-2.616.399, constituindo uma região de aproximadamente 647.686 pares de base com *contigs* variando entre 70 pb e 1.200 pb. Já no genoma da estirpe LGMAI_St_11, a faixa com grande desvio de conteúdo GC foi menor, localizada entre 1.950.334-2.016.069, totalizando 65.735 pares de base com *contigs* variando em tamanho entre 74 pb e 1.400 pb. Por fim, no genoma da estirpe LGMAI_St_14, o desvio de conteúdo GC está localizado entre 2.172.030-2.221.147, constituindo uma região de 49.117 pares de base com *contigs* variando entre 72 pb e 1.312 pb. Essas regiões são provavelmente artefatos de técnica resultantes do sequenciamento e da metodologia usada na montagem dos genomas e foram, assim, pobremente anotadas pelas ferramentas de anotação genômica. Ao se remover os 647.686 pares de base do genoma da estirpe LGMAI_St_08, obtém-se um tamanho mais próximo da média das estirpes LGMAI_St_11 e LGMAI_St_14.

A análise por sequência multilocus (MLSA) é uma das principais técnicas usadas em estudos de identificação taxonômica nos dias de hoje. O uso dos genes *housekeeping aroE* (chiquimato desidrogenase), *gki* (glicoquinase), *pheS* (subunidade alfa da fenilalanina tRNA ligase), *ddl* (D-alanina-D-alanina ligase) e *recA* (recombinase A), além do 16S rRNA, já haviam sido usados em trabalhos de taxonomia por *neighbor joining* do gênero *Streptococcus* (Thompson

et al., 2013). No presente trabalho de dissertação de mestrado, os mesmos genes, excetuando-se *ddl*, foram usados para realizar a identificação taxonômica das estirpes LGMAI com a metodologia de máxima verossimilhança.

Essa análise foi capaz de realizar a identificação taxonômica das estirpes LGMAI como pertencentes à espécie *Streptococcus agalactiae* com bastante clareza, além de demonstrar o delineamento natural das espécies de *Streptococcus* nos 7 grandes grupos atuais (*anginosus*, *bovis*, *mitis*, *mutans*, *pyogênico*, *salivarius* e *sanguinis*) (Haslam; Geme, 2018; Gao *et al.*, 2014), apresentando somente a espécie *Streptococcus parasanguinis* ATCC 15912 fora do grupo *sanguinis*. O uso de um alinhamento concatenado com mais proteínas ortólogas na análise MLSA, como feito no trabalho de Gao e colaboradores (2014), pode trazer mais clareza nos delineamentos do gênero *Streptococcus*. A análise taxonômica baseada em dados de sequenciamento é muito útil e traz grandes esclarecimentos, especialmente, a respeito das relações entre as espécies do grupo Viridans de Lancefield, o qual é confuso e problemático, uma vez que agrupa espécies que não compartilham muitas características em comum entre si (Haslam; Geme, 2018).

O uso do protocolo de sorotipagem de *Streptococcus* grupo B com dados de sequenciamento WGS do Centro de Controle e Prevenção de Doenças dos Estados Unidos da América permitiu a identificação das estirpes LGMAI como *Streptococcus agalactiae* sorotipo III. O trabalho de Castro e colaboradores (2017) sobre a epidemiologia molecular de mastite bovina por infecção de *Streptococcus agalactiae* no Brasil mostrou que o sorotipo III é o mais frequente, independentemente de localização geográfica (Minas Gerais, São Paulo, Paraná e Pernambuco) e severidade da doença (clínica e subclínica). Nos casos clínicos de mastite, 50%, no entanto, era causada por *Streptococcus agalactiae* sorotipo III. No Estado de Minas Gerais, origem das estirpes LGMAI_St_11 e LGMAI_St_14, o sorotipo mais frequentemente identificado é o III, correspondendo a 74,04% dos isolados. No caso de São Paulo, origem da LGMSI_St_08, o sorotipo mais frequente é o II (54,54%), seguido do sorotipo III (18,18%). A estirpe LGMAI_St_08, diferentemente das LGMAI_St_11 e LGMAI_St_14, corresponde a um sorotipo menos frequentemente identificado em mastite bovina no seu Estado de origem.

A análise de genes que conferem resistência antimicrobiana revelou a presença de dois genes nos genomas das estirpes LGMAI: *mreA* e *mprF*. O gene *mreA* codifica uma flavoquinase, responsável pela biossíntese de riboflavina, coenzima com função de oxirredução no metabolismo de carboidratos e proteínas (Clarebout; Villers; Leclercq, 2001). Originalmente, esse gene foi

descrito por sua associação com resistência a macrolídeos por meio de bombas de efluxo. Ele é responsável por conferir resistência aos antibióticos eritromicina, azitromicina e espiramicina (Clancy *et al.*, 1997). Embora as concentrações inibitórias mínimas (MIC) para eritromicina, espiramicina e clindamicina tenham seu valor duplicado e até quadruplicado, esse gene também induz aumento de MIC contra outros antibióticos, como algumas cefalosporinas, cloranfenicol, doxiciclina, ácido nalidíxico, novobiocina e rifampicina (Clarebout; Villers; Leclercq, 2001).

O gene *mprF*, por outro lado, codifica um fator de resistência a múltiplos peptídeos. Esse fator é responsável por realizar a transferência do grupo lisil do L-lisil-tRNA para o fosfatidilglicerol, componente das membranas plasmáticas, produzindo lisilfosfatidilglicerol (Oku *et al.*, 2004). Sendo assim, esse fator contribui com a virulência bacteriana, uma vez que a síntese de lisilfosfatidilglicerol está envolvida na resistência contra peptídeos antimicrobianos catiônicos (CAMP), tais quais defensinas e catelicidinas do sistema imune dos hospedeiros e até mesmo bacteriocinas de microrganismos competidores (Staubitz *et al.*, 2004). Além disso, mutações conferindo superexpressão de *mprF* em *Staphylococcus aureus* resistente à meticilina foram associadas à resistência cruzada contra daptomicina e vancomicina (Thitianapakorn *et al.*, 2020).

O gene *mreA*, responsável por conferir resistência antimicrobiana contra macrolídeos, como a eritromicina e clindamicina, tem grande importância clínica em *Streptococcus agalactiae*, porque essa é uma bactéria classificada como preocupante em termos de resistência antimicrobiana. Embora a resistência de *Streptococcus* grupo B contra clindamicina seja mais vital, uma vez que esse é o antibiótico usado em casos de alergia a penicilina (primeira linha de combate contra infecções por GBS), a resistência à clindamicina é encontrada em cerca de 40% das infecções por GBS, ao passo que resistência à eritromicina é identificada em mais da metade das infecções por BS e, portanto, precisam ser monitoradas (CDC, 2019).

A análise dos genes de virulência identificou resultados muito similares entre os 3 genomas das estirpes LGMAI, apresentando o locus *cps*, o operon *cyl* e o gene *hylB*. Além disso, no genoma da estirpe LGMAI_St_08 foi encontrado o gene *cfa/cfb*.

O locus *cps* corresponde a uma região de conglomerado de genes, responsável pela confecção do polissacarídeo capsular que obrigatoriamente reveste bactérias de *Streptococcus agalactiae*. Essa cápsula é rica em ácido siálico, um monossacarídeo de 9 carbonos também presente nas estruturas de glicano das células dos hospedeiros, e é um dos principais fatores responsáveis pela virulência deste microrganismo (Armistead *et al.*, 2019). Por esse motivo, essa

cápsula de polissacarídeos corresponde ao primeiro fator de virulência bacteriano identificado a permitir evasão do sistema imune e invasão das células, devido a suas propriedades antifagocitose (Castro *et al.*, 2017). Nas estirpes LGMAI, foram encontrados os genes *cpsA*, *cpsB*, *cpsC*, *cpsD*, *cpsE*, *cpsF*, *cpsK*, *cpsL*, *neuA*, *neuB*, *neuC* e *neuD*, compartilhados por todos os 10 sorotipos de *Streptococcus agalactiae* (Kapatai *et al.*, 2017).

O óperon *cyl* é a região com o conjunto de genes, responsável pela síntese do pigmento ramnolipídeo ornitina, isto é, o pigmento toxina que garante a capacidade hemolítica de bactérias de *Streptococcus agalactiae* (Vornhagen; Waldorf; Rajagopal, 2017). Os genes do operon *cyl* encontrados foram *cylX*, *cylD*, *cylG*, *acpC*, *cylZ*, *cylA*, *cylB*, *cylF*, *cylI*, *cylJ* e *cylK*. Tal pigmento é composto por uma cadeia de poliene, consistindo de 12 ligações duplas, com forte atividade citotóxica e antioxidante, que auxiliam na invasão e resiliência contra as defesas inatas do organismo hospedeiro (Armistead *et al.*, 2019). O pigmento hemolítico é capaz de causar o rompimento da barreira de células epiteliais amnióticas e, conseqüentemente, penetrar a placenta humana. Sendo assim, é um fator de virulência muito importante na colonização do trato genital feminino, por facilitar a infecção e também por diminuir a suscetibilidade das células bacterianas à eliminação por neutrófilos (Vornhagen; Waldorf; Rajagopal, 2017).

É um dos principais fatores de virulência de *Streptococcus agalactiae*, principalmente por sua associação com o sucesso das infecções e com o desenvolvimento de complicações tanto em mulheres grávidas, quanto em bebês, estando relacionado a partos prematuros em decorrência de contrações e respostas inflamatórias (Armistead *et al.*, 2019). Recém-nascidos estão em condição muito suscetível a infecção por contato com tecidos contaminados durante o parto normal e também pela aspiração de fluido amniótico ainda na condição de feto (Nascimento *et al.*, 2019). Além disso, o pigmento hemolítico também é capaz de auxiliar na penetração da barreira hematoencefálica e, com isso, desencadear respostas inflamatórias em células endoteliais microvasculares, evoluindo a infecção para um quadro clínico de meningite (Armistead *et al.*, 2019).

O gene *hylB* codifica a exoenzima hialuronato liase, responsável por clivar as ligações glicosídicas do ácido hialurônico, degradando-o em seus componentes: a N-acetilglicosamina e o ácido glicurônico. Esse fator de virulência é um dos mais importantes, conferindo hipervirulência até mesmo em bactérias que não produzem pigmento hemolítico, mas que superexpressam o gene *hylB* (Armistead *et al.*, 2019). Sua grande importância no processo de infecção vem da sua capacidade de modulação do sistema imune, uma vez que os componentes do ácido hialurônico

ligam-se aos receptores *toll-like* TLR2 e TLR4 das vias de reconhecimento e sinalização do sistema imune, bloqueando tais vias e atuando, assim, como imunossupressores (Vornhagen; Waldorf; Rajagopal, 2017). Assim, é um fator importante na colonização do trato genital feminino e na evolução da infecção em quadros clínicos mais invasivos. Na esfera da pecuária, o gene *hylB*, juntamente com os genes *fbsB* e *cfa/cfb* e a ilha de patogenicidade PI-2a, correspondem a genes cruciais nas infecções por GBS, sendo detectados em altas frequências em casos de mastite bovina (Castro *et al.*, 2017).

O gene *cfa/cfb* foi encontrado exclusivamente no genoma da estirpe de São Paulo LGMAI_St_08. Esse gene codifica um fator CAMP (Christie, Atkins e Munch-Petersen), o qual se caracteriza como uma toxina responsável pela formação de poros nas membranas das células do hospedeiro, por meio de um processo de oligomerização do mesmo e, conseqüentemente, causando a lise de células, em especial as células vermelhas (Lang; Palmer, 2003). Além disso, esse fator é capaz de se ligar à porção Fc das imunoglobulinas IgM e IgG por meio de uma reação não imune de modo similar à reação descrita para a proteína A de *Staphylococcus aureus* – responsável por evasão do sistema imune ao se ligar a imunoglobulinas (Jürgens; Sterzik; Fehrenbach, 1987). Assim como o gene *hylB*, o gene *cfa/cfb* também é um dos genes causadores de danos em tecidos e é crucial nos processos de infecção por GBS em gado, apresentando altas frequências de detecção em casos clínicos de mastite bovina (Castro *et al.*, 2017).

Nem todos os fatores de virulência de *Streptococcus agalactiae* foram encontrados, como os fatores de aderência FbsA (proteína A de superfície de ligação de fibrinogênio) (Schubert *et al.*, 2002), FbsB (proteína B de superfície de ligação de fibrinogênio) (Gutekunst; Eikmanns; Reinscheid, 2004) e Lmb (proteína de ligação das glicoproteínas lamininas) (Spellerberg *et al.*, 1999) e o gene de antifagocitose *cba* (antígeno β da proteína C) (Jerlström *et al.*, 1996). No entanto, ao analisar os arquivos de anotação dos 3 genomas, foi possível identificar os genes: *bca*, o qual codifica o antígeno α da proteína C, responsável por interagir com glicosaminoglicanos das células dos hospedeiros e mediar a internalização bacteriana por meio de rearranjos de actina (Baron *et al.*, 2004); e os genes *scpAB*, que codificam a C5a peptidase, responsável por evasão do sistema imune, uma vez que clivam o agente quimiotático C5a entre os resíduos de histidina₆₇ e lisina₆₈, agente esse essencial na imunidade inata humana (Brown *et al.*, 2005).

A análise de detecção de ilhas genômicas de resistência e patogenicidade foi capaz de identificar diversas ilhas contidas nos genomas das estirpes LGMAI. A análise comparativa, por

meio de identidade média de nucleotídeos, realizada pela ferramenta BRIG, mostrou que a maior parte das ilhas genômicas identificadas são conservadas. No entanto, algumas ilhas genômicas identificadas não são conservadas de acordo com a análise. Para a estirpe LGMAI_St_08, três ilhas encontram-se em regiões não conservadas (1.106.326-1.118.347, 1.906.774-1.954.486 e 3.403.129-3.461.159). Para a estirpe LGMAI_St_11, quatro ilhas encontram-se em regiões não conservadas (2.109.442-2.135.980, 2.295.771-2.305.125, 2.319.728-2.367.935 e 2.414.896-2.447.492). Por fim, para a estirpe LGMAI_St_14, quatro ilhas encontram-se em regiões não conservadas (1.960.979-1.988.147, 2.367.253-2.376.665, 2.470.033-2.477.686 e 2.511.034-2.531.428). Todas essas ilhas, com exceção de uma ilha de resistência da estirpe LGMAI_St_14 (2.470.033-2.477.686), foram preditas como ilhas genômicas de patogenicidade.

Dos genes de resistência, o gene *mreA* identificado no genoma da estirpe LGMAI_St_08 está localizado na faixa 1.746.557-1.747.492, a qual se encontra dentro de uma ilha genômica putativa de resistência de escore fraco, identificada com o software Gypsy (não foi plotada no mapa circular de identidade com o programa BRIG pelo escore fraco).

Já dos genes de virulência, duas cópias do gene *hylB* foram identificadas no genoma da estirpe LGMAI_St_08 e estão localizadas nas faixas de 177.577-180.795 e 1.710.797-1.714.015, as quais se encontram, respectivamente, dentro de uma ilha genômica mista e escore normal e uma ilha genômica de patogenicidade de escore normal. Para o genoma da estirpe LGMAI_St_11, o operon *cyl* encontra-se localizado dentro da faixa de 2.111.749-2.122.527, a qual está contida em uma ilha genômica de patogenicidade com escore normal. Por fim, para o genoma da estirpe LGMAI_St_14, o gene *hylB* detectado encontra-se na faixa de 471.540-474.758, a qual está contida em uma ilha genômica de patogenicidade com escore normal.

Bacteriófagos são vírus que infectam bactérias e, portanto, apresentam, atualmente, grande potencial no combate a bactérias patogênicas multirresistentes por meio da terapia de fagos. No entanto, bacteriófagos naturalmente introduzem seu material genético no processo de infecção em bactérias, o que desencadeia consequências inesperadas, tal qual a disseminação de genes que garantem resistência antimicrobiana (Barceló, 2018), além de alguns fagos carregarem genes que codificam toxinas, afetando também a virulência da bactéria hospedeira (Lichvariková *et al.*, 2020). Torna-se, portanto, necessária a avaliação de genomas bacterianos em busca de sequências de fagos a fim de se avaliar o potencial de disseminação de genes que induzem as bactérias a se tornarem

mais resistentes ou virulentas (Arndt *et al.*, 2016). Nenhuma sequência de fago completo foi identificada nos genomas das 3 estirpes LGMAI.

Lichvariková e colaboradores (2020) realizaram um estudo na Eslováquia a respeito da contribuição de variabilidade em isolados de *Streptococcus agalactiae* com base na presença de sequências de bacteriófagos nos seus genomas. Amostras foram coletadas de pacientes infectados por GBS de 2016-2018, além de amostras de testes de gravidez no ano de 2018. Após identificação de clonalidade e tipagem das amostras, 20 isolados representativos de *Streptococcus agalactiae* tiveram seus genomas sequenciados e analisados, e foram encontradas de 1 a 3 sequências de bacteriófagos por genoma. Essa elevada frequência de bacteriófagos nos genomas de *Streptococcus agalactiae* clínicos da Eslováquia pode estar associada a benefícios para a bactéria hospedeira. Diferentemente do encontrado por Lichvariková e colaboradores (2020), nenhum bacteriófago completo foi identificado nas estirpes LGMAI. No entanto, 2 sequências de bacteriófagos questionáveis foram identificadas na estirpe LGMAI_St_08, ao passo que, nas outras duas, apenas sequências de bacteriófagos incompletas foram reveladas. Tendo em vista a importância dos bacteriófagos na variabilidade e evolução de *Streptococcus agalactiae*, como relatada por Lichvariková e colaboradores (2020), a pouca capacidade de identificação de tais sequências, no presente trabalho, pode ser decorrente da metodologia usada não ser tão elaborada para identificação dessas sequências, quanto a usada por Lichvariková e colaboradores (2020).

A busca por antibióticos e metabólitos secundários nos genomas das estirpes LGMAI identificou três sítios putativos para produção de policetídeo sintase do tipo III, aril polienos e lantipeptídeos de classe V. Policetídeos representam uma classe de compostos bastante diversificada que são biologicamente ativos, atuando com características antineoplásicas, anticolesterol e antimicrobianas (Yu *et al.*, 2012). Aril polienos, por sua vez, são pigmentos bacterianos de estrutura similar a dos carotenóides, que, assim como os carotenóides, atuam na proteção bacteriana contra espécies reativas de oxigênio (Schöner *et al.*, 2016). Já os lantipeptídeos são peptídeos que contêm o aminoácido lantionina em sua composição. Eles são sintetizados nos ribossomos e sofrem modificações pós-traducionais, podendo apresentar uma gama de atividades biológicas, incluindo atividade antimicrobiana (Repka *et al.*, 2017). No entanto, nenhum desses sítios identificados apresentou gene estrutural para os peptídeos.

A região do cluster biossintético de peptídeo não ribossômico foi a que apresentou maior completude nas análises com o antiSMASH. Essa região sintetiza uma equibactina, um dos

primeiros sistemas de seu tipo identificado no gênero *Streptococcus*. Originalmente identificado em *Streptococcus equi* subsp. *equi*, esse sistema está associado à aquisição de ferro e, com a deleção do gene *eqbA* desse locus, observa-se um fenótipo de colônias menores como decorrência de toxicidade por altos níveis de importação de íons de ferro (Heather *et al.*, 2008). Assim como o sistema análogo, a yersiniabactina, presente na ilha de alta patogenicidade de *Yersinia* sp., a equibactina pode estar associada ao potencial de virulência de *Streptococcus agalactiae*. A aquisição de ferro é um processo importante para bactérias patogênicas, haja vista a baixa quantidade de ferro livre em células de hospedeiros mamíferos (Miethke; Marahiel, 2007).

Em relação à busca por genes que codificam peptídeos antimicrobianos com a ferramenta BAGEL4, foi identificado, nos três genomas, o gene que codifica o peptídeo estrutural da bacteriocina zoocina A com identidade de 58,76%. A zoocina A é uma exoenzima lítica originalmente identificada de *Streptococcus equi* subsp. *zooepidemicus* 4881, que atua na hidrólise da estrutura de peptidoglicano de microrganismos suscetíveis. Os primeiros trabalhos publicados relataram queda na eficiência lítica da enzima quando na presença de β -lactâmicos, indicando que se ligam a β -lactâmicos e possuem uma atividade na forma de uma endopeptidase de D-alanina (Heath *et al.*, 2004). Além disso, a bacteriocina zoocina A apresenta uma atividade inibitória com forte predileção para o gênero *Streptococcus*, sendo relatada inibição contra *Streptococcus equi* subsp. *zooepidemicus*, *Streptococcus pyogenes*, *Streptococcus gordonii* e *Streptococcus mutans* (Akesson *et al.*, 2007). Essa característica vai de encontro aos resultados de Silva (2003), nos quais foi observada a inibição de diversos representantes de *Streptococcus*, corroborando com o potencial uso dessa substância como um antibiótico no controle de *Streptococcus* patogênicas, em especial *Streptococcus pyogenes*, que apresenta extrema sensibilidade ao peptídeo (Akesson *et al.*, 2007).

A análise no banco de dados BACTIBASE das sequências do peptídeo putativo zoocina A identificada nos genomas LGMAI corroborou com os resultados obtidos na plataforma do BAGEL4, identificando correspondência com a zoocina A de *Streptococcus equi* subsp. *zooepidemicus*, com uma identidade de 60% e uma similaridade de 66%. Esse peptídeo antimicrobiano é composto de 2 domínios: um domínio catalítico N-terminal (peptidase M23) e um domínio de reconhecimento de alvo C-terminal (TRD) (Chen; Simmonds; Timkovich, 2013). Buscas por domínios com o servidor do Interpro foram capazes de identificar os dois domínios, catalítico e de reconhecimento, nas sequências extraídas dos genomas das estirpes LGMAI. Além disso, o alinhamento das zoocinas A putativas das estirpes LGMAI com outras zoocinas A,

disponíveis no banco de dados do Uniprot, mostrou que as regiões dos dois domínios e do peptídeo sinal são os sítios de maior conservação dessas sequências. Conservação similar já havia sido notificada, principalmente do sítio catalítico, o qual apresenta, também, grande similaridade com outras endopeptidases bacteriolíticas (Gargis *et al.*, 2009).

Segundo o banco de dados BACTIBASE, trabalhos anteriores com a zoocina A não apresentaram atividade inibitória contra *Micrococcus luteus* e *Streptococcus uberis* (Simmonds; Simpson; Tagg, 1997). No entanto, Silva (2003) observou atividade contra *Streptococcus uberis* e, no presente trabalho, foi observada atividade contra *Micrococcus* sp., a qual precisa de identificação a nível de espécie para maiores conclusões.

Além do peptídeo estrutural da zoocina A, o fator que garante imunidade às células produtoras é outra parte essencial do esquema de atuação desse peptídeo antimicrobiano. O fator de imunidade à zoocina A é chamado de Zif e caracteriza-se como uma proteína de imunidade do tipo Fem-ABX. Esse fator está associado à adição de uma l-alanina adicional em ligações cruzadas durante a biossíntese da estrutura de peptidoglicano, o que inibe tanto o reconhecimento realizado pelo sítio TRD, quanto a atividade catalítica do peptídeo (Gargis *et al.*, 2009). A busca realizada por meio de BLAST para identificação de regiões semelhantes ao fator de imunidade Zif identificou sítios, nos três genomas das estirpes LGMAI, com 50% de identidade e 65% de similaridade, os quais, quando comparados aos resultados da anotação genômica realizada com o programa Prokka, codificam aminoaciltransferases FemB. A busca por domínios e famílias para essas sequências obteve correspondência com a família FemABX de peptidiltransferases, solidificando a identificação dessas enzimas.

A identificação de grupos ortólogos pela ferramenta eggNOG-mapper mostrou uma distribuição bastante similar entre os três genomas das estirpes LGMAI. Não obstante, a identificação de grupos ortólogos KEGG pela ferramenta KAAS, por outro lado, revelou uma distribuição similar entre as estirpes mineiras LGMAI_St_11 e LGMAI_St_14, mas bastante discrepante para a estirpe paulista LGMAI_St_08, o qual obteve um número muito maior de grupos observados. Muito provavelmente isso é decorrente do maior tamanho genômico deste último (3,5 Mbp) em comparação aos outros dois, permitindo a identificação de novos grupos ortólogos que não foram encontrados nas estirpes mineiras.

A análise de genômica comparativa realizada com o servidor OrthoVenn2 obteve resultados que corroboram com os padrões observados nas análises de anotação de grupos

ortólogos. Essa análise mostrou um grande compartilhamento de subgrupos funcionais entre as estirpes LGMAI, com uma maior similaridade entre as estirpes mineiros quando em comparação a paulista. Além de mostrar maior conservação para proteínas de função molecular e de componentes celulares, com maior variação em proteínas associadas a processos metabólicos.

Ainda, é interessante ressaltar que a estirpe LGMAI_St_08 apresentou 69 subgrupos funcionais exclusivos que não eram compartilhados com as estirpes mineiras. Esses 69 subgrupos funcionais são, em sua maioria, pertencentes ao grupo de processos biológicos, que podem estar associados a adaptações diferenciais ao ambiente de origem dessa estirpe, assim como estar associados aos grupos KEGG funcionais extras de doenças humanas identificados exclusivamente na estirpe LGMAI_St_08, indicando um maior potencial de patogenicidade em humanos para essa estirpe. Estirpe, na qual, também foi exclusivamente identificado o gene de virulência *cfa/cfb*.

6 CONCLUSÃO

A presente dissertação de mestrado assumiu como objetivo dar continuidade ao trabalho com as estirpes de *Streptococcus agalactiae* isolados de mastite bovina, produtoras de substâncias antimicrobianas, buscando encontrar a origem de sua atividade antimicrobiana.

Para tal, foi reavaliado o potencial de antibiose relatado por Silva (2003) de 4 estirpes de *Streptococcus agalactiae* isolados de mastite bovina contra uma estirpe de *Streptococcus pyogenes* e realizado novos testes contra estirpes de *Staphylococcus aureus* e *Klebsiella pneumoniae* com e sem resistência antimicrobiana, além de *Micrococcus* sp. e *Cellulomonas fimi*. Com isso, observou-se que as quatro estirpes foram igualmente capazes de inibir *Streptococcus pyogenes*, *Micrococcus* sp. e *Cellulomonas fimi*, essas duas últimas espécies importantes na indústria alimentícia por deterioração de produtos, mas não foram capazes de inibir *Staphylococcus aureus* nem *Klebsiella pneumoniae*, com ou sem resistência antimicrobiana.

Com o sequenciamento desses genomas, pretendeu-se, primeiramente, realizar a identificação taxonômica das estirpes por meio de análise por sequências multilocus, segundo a metodologia da máxima verossimilhança. As estirpes já haviam sido avaliadas por testes bioquímicos e classificadas como *Streptococcus agalactiae* pela pesquisadora Maria Aparecida Vasconcelos Paiva Brito da Embrapa, responsável por seu isolamento. As análises taxonômicas aqui realizadas obtiveram resultados claros que permitiram a identificação das três estirpes como também pertencentes à espécie *Streptococcus agalactiae*, corroborando com dados prévios. Ademais, por meio de protocolo de sorotipagem de *Streptococcus* grupo B para dados de sequenciamento do Centro de Controle e Prevenção de Doenças, foi possível a designação das três estirpes como pertencentes ao sorotipo III, contribuindo ainda mais para uma melhor identificação das mesmas.

Em segundo lugar, considerando a importância clínica e o potencial patogênico dessas estirpes, assim como as preocupantes taxas de resistência antimicrobiana relatadas para a espécie, análises de busca por fatores de virulência e presença de genes de resistência foram realizadas. Com essas análises, foi encontrado o gene *mreA* nos três genomas, o qual codifica uma flavoquinase e está associado à resistência contra eritromicina, resultado esse importante, uma vez que resistência à eritromicina foi encontrada em mais da metade das infecções por GBS relatadas no último relatório do CDC e, portanto, precisa ser monitorada. Em relação à virulência, foram encontrados praticamente todos os principais fatores de virulência de *Streptococcus agalactiae*,

excetuando-se somente os fatores de aderência, como FbsA, FbsB, Lmb e as ilhas de pilus e o gene antifagocitose *cba*.

Em terceiro lugar, com o objetivo de avaliar o potencial biotecnológico das estirpes de *Streptococcus agalactiae* isolados de mastite bovina, assim como buscar identificar a fonte da capacidade antimicrobiana das mesmas, relatada por Silva (2003), foram realizadas análises de busca por genes que codificam antimicrobianos e metabólitos secundários. Essas análises revelaram a presença de sítios putativos para antibióticos e metabólitos secundários com diversas atividades biológicas de interesse, como atividade antimicrobiana, antioxidante e de captação de ferro. No entanto, o sítio mais completo e que mais corrobora com os resultados anteriores desses isolados, é o sítio de produção da bacteriocina zoocina A, uma bacteriocina originalmente identificada em *Streptococcus equi* subsp. *zooepidemicus* que apresenta foco de inibição direcionado para espécies de *Streptococcus*. O mesmo padrão de inibição já havia sido notado desde os testes iniciais de antibiose realizados com estirpes LGMAI. Além do gene estrutural dessa enzima, também foi identificado gene com similaridade ao fator de imunidade Zif, o qual permite a sobrevivência das bactérias produtoras à zoocina A.

Esse trabalho apresenta, não obstante, algumas limitações, principalmente no que tange à necessidade por análises de bancada que corroborem com os resultados identificados por métodos *in silico*. A atuação do gene que codifica a enzima zoocina A, por exemplo, precisa de ser avaliada por meio de um nocaute gênico. Dessa forma, poderá ser observado o fenótipo de bactérias deficientes na expressão do mesmo e, assim, corroborar o seu papel na atividade antimicrobiana identificada nas estirpes de *Streptococcus agalactiae* isolados de mastite bovina.

Ainda, o isolamento e purificação das substâncias antimicrobianas produzidas pelas estirpes *agalactiae* LGMAI faziam parte dos delineamentos originais do projeto desta dissertação de mestrado. Contudo, por decorrência do grande tempo demandado e da dificuldade dessa tarefa, além, obviamente, das adversidades impostas pela epidemia global da COVID-19, esse objetivo foi por fim abandonado. Trabalhos futuros com essas estirpes precisam de almejar tal objetivo, para que se possa ser quantificada a capacidade inibitória dessas substâncias, assim como calculado seu peso molecular e sequenciada sua cadeia de aminoácidos.

Apesar das limitações supracitadas, a presente dissertação de mestrado possibilitou melhor entendimento da origem da atividade antimicrobiana identificada nas estirpes LGMAI, definindo um ponto de partida para futuras investigações. Revelou, inclusive, que as estirpes não

são capazes de inibir estirpes representantes de *Klebsiella pneumoniae* e *Staphylococcus aureus*, com ou sem resistência antimicrobiana. Contribuiu, ademais, com informações de virulência e monitoramento de resistência antimicrobiana em estirpes de *Streptococcus agalactiae* isolados de mastite bovina nos estados de São Paulo e Minas Gerais no Brasil. Além do mais, o sequenciamento dessas estirpes forneceu dados genômicos de isolados bovinos clínicos de *Streptococcus agalactiae* brasileiros, os quais poderão ser acessados por cientistas de todo o mundo e usado em pesquisas científicas diversas.

REFERÊNCIAS

- Agência Nacional de Vigilância Sanitária – Anvisa. **Boletim informativo – Segurança do Paciente e Qualidade dos Serviços de Saúde**, Ano V, n. 9, Brasil, 2014. Disponível em: <<http://portal.anvisa.gov.br>>.
- AKESSON, M.; DUFOUR, M.; SLOAN, G. L.; SIMMONDS, R. S. Targeting os streptococci by zoocin A. **FEMS Microbiology Letters**, v. 270, p. 155-161, 2007.
- ALIKHAN, N. F.; PETTY, N. K.; ZAKOUR, N. L. B.; BEATSON, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. **BMC Genomics**, v. 12, n. 402, 2011.
- AMINOV, R. History of antimicrobial drug discovery – Major classes and health impact. **Biochemical Pharmacology**, v. 133, p. 4-19, 2017.
- ANDREWS, S. 2010. **FastQC: A Quality Control Tool for High Throughput Sequence Data**. Disponível em < <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> >.
- ANGELOPOULOU, A.; WARDA, A. K.; HILL, C.; ROSS, P. Non-antibiotic microbial solutions for bovine mastitis – live biotherapeutics, bacteriophage, and phage lysins. **Critical Reviews in Microbiology**, v. 45, n. 5-6, p. 564-580, 2019.
- ARAUZ, L. J.; JOZALA, A. F.; MAZZOLA, P. G.; e PENNA, T. C. V. Nisin biotechnological production and application: a review. **Trends in Food Science and Technology**, v. 20, n. 3-4, p. 146-154, 2009.
- ARMISTEAD, B.; OLER, E.; WALDORF, K. A.; RAJAGOPAL, L. The double life of Group B *Streptococcus*: Asymptomatic colonizer and potent pathogen. **Journal of Molecular Biology**, v. 431, n. 16, p. 2914-2931, 2019.
- ARNDT, D.; GRANT, J. R.; MARCU, A.; SAJED, T.; PON, A.; LIANG, Y.; WISHART, D. S. PHASTER: a better, faster version of the PHAST phage search tool. **Nucleic Acids Research**, v. 44, n. 1, p. 16-21, 2016.
- AUFFERMANN, K. M.; GRIJALVA, F.; JACOB, F.; HUTZLER, M. Nisin and its usage in breweries: a review and discussion. **Journal of the Institute of Brewing**, v. 121, p. 309-319, 2015.
- AZARIAN, T.; HUANG, I. T.; HANAGE, W. P. Structure and Dynamics of Bacterial Populations: Pangenome Ecology. In: TETTELIN, H.; MEDINI, D. **The Pangenome: Diversity, Dynamics, and Evolution of Genomes**, Springer, p. 115-128, 2020.
- BACHMANN, B. O.; LANEN, S. G. V.; BALTZ, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? **Journal of Industrial Microbiology and Biotechnology**, v. 41, n. 2, p. 175-184, 2014.
- BAHAR, A. A.; REN, D. Antimicrobial Peptides. **Pharmaceuticals**, v. 6, n. 12, p. 1543-1575, 2013.

BALCIUNAS, E. M.; MARTINEZ, F. A. C.; TODOROV, S. D.; FRANCO, B. D. G. M.; CONVERTI, A.; e OLIVEIRA, R. P. S. Novel biotechnological applications of bacteriocins: a review. **Food Control**, v. 32, n. 1, p. 134-142, 2013.

BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A. A.; DVORKIN, M.; KULIKOV, A. S.; LESIN, V. M.; NIKOLENKO, S. I.; PHAM, S.; PRJIBELSKI, A. D.; PYSHKIN, A. V.; SIROTKIN, A. V.; VYAHHI, N.; TESLER, G.; ALEKSEYEV, M. A.; e PEVZNER, P. A. 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.

BARCELÓ, C. T. The disparate effects of bacteriophages on antibiotic-resistant bacteria. **Emerging Microbes & Infections**, v. 7, n. 1, p. 1-12, 2018.

BARON, M. J.; BOLDUC, G. R.; GOLDBERG, M. B.; AUPÉRIN, T. C.; MADOFF, L. C. Alpha C Protein of Group B Streptococcus Binds Host Cell Surface Glycosaminoglycan and Enters Cells by an Actin-dependent Mechanism. **Journal of Biological Chemistry**, v. 279, n. 23, p. 24714-24723, 2004.

BEHJATI, S.; TARPEY, P. S. What is next generation sequencing? **Archives of Disease in Childhood: Education & Practice**, v. 98, n. 6, p. 236-238, 2013.

BERTELSEN, K.; DOROSZ, J.; HANSEN, S. K.; NIELSEN, N. C.; VOSEGAARD, T. Mechanisms of Peptide-Induced Pore Formation in Lipid Bilayers Investigated by Oriented ³¹P Solid State NMR Spectroscopy. **PLoS ONE**, v. 7, n. 10, 2012.

BIANCHI, R. M.; SCHWERTZ, C. I.; CECCO, B. S.; PANZIERA, W.; LORENZO, C.; HECK, L. C.; SNEL, G. G. M.; LOPES, B. C.; SILVA, F. S.; PAVARINI, S. P.; DRIEMEIER, D. Pathological and microbiological characterization of mastitis in dairy cows. **Tropical Animal Health and Production**, v. 51, p. 2057-2066, 2019.

BLUM, M.; CHANG, H. Y.; CHUGURANSKY, S.; GREGO, T.; KANDASAAMY, S.; MITCHELL, A.; NUKA, G.; LAFOSSE, T. P.; QURESHI, M.; RAJ, S.; RICHARDSON, L.; SALAZAR, G. A.; WILLIAMS, L.; BORK, P.; BRIDGE, A.; GOUGH, J.; HAFT, D. H.; LETUNIC, I.; BAUER, A. M.; MI, H.; NATALE, D. A.; NECCI, M.; ORENCO, C. A.; PANDURANGAN, A. P.; RIVOIRE, C.; SIGRIST, C. J. A.; SILLITOE, I.; THANKI, N.; THOMAS, P. D.; TOSATTO, S. C. E.; WU, C. H.; BATEMAN, A.; FINN, R. D. The InterPro protein families and domains database: 20 years on. **Nucleic Acids Research**, v. 49, n. d1, p. d344-d354, 2021.

BROWN, C. K.; GU, Z. Y.; MATSUKA, Y. V.; PURUSHOTHAMAN, S. S.; WINTER, L. A.; CLEARY, P. P.; OLMSTED, S. B.; OHLENDORF, D. H.; EARHART, C. A. Structure of the streptococcal cell wall C5a peptidase. **PNAS**, v. 102, n. 51, p. 18391-18396, 2005.

CASTRO, G. A. C.; SILVA, J. R.; PAIVA, L. V.; CUSTÓDIO, D. A. C.; MOREIRA, R. O.; MIAN, G. F.; PRADO, I. A.; JUNIOR, A. C.; COSTA, G. M. Molecular epidemiology of *Streptococcus agalactiae* isolated from mastitis in Brazilian dairy herds. **Brazilian Journal of Microbiology**, v. 48, n. 3, p. 551-559, 2017.

Centers for Disease Control and Prevention – CDC. **Antibiotic Resistance Threats in the United States**, 2013. Atlanta, Geórgia: Departamento de Saúde e Serviços Humanos dos Estados Unidos, CDC; 2013.

Centers for Disease Control and Prevention – CDC. **Antibiotic Resistance Threats in the United States**, 2019. Atlanta, Geórgia: Departamento de Saúde e Serviços Humanos dos Estados Unidos, CDC; 2019.

CEPAS, J. H.; FORSLUND, K.; COELHO, L. P.; SZKLARCZYK, D.; JENSEN, L. J.; MERING, C. V.; BORK, P. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. **Molecular Biology and Evolution**, v. 34, n. 8, p. 2115-2122, 2017.

CEPAS, J. H.; SZKLARCZYK, D.; HELLER, D.; PLAZA, A. H.; FORSLUND, S. K.; COOK, H.; MENDE, D. R.; LETUNIC, I.; RATTEI, T.; JENSEN, L. J.; MERING, C. V.; BORK, P. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on organisms and 2502 viruses. **Nucleic Acids Research**, v. 47, n. D1, p. D309-314, 2019.

CHACHATY, E.; SAULNIER, P. Isolating Chromosomal DNA from Bacteria. *In*: RAPPLEY, R. **The Nucleic Acid Protocols Handbook**. Humana Press, 2000, p. 29-32.

CHEN, L.; ZHENG, D.; LIU, B.; YANG, J.; JIN, Q. VFDB 2016: hierarchical and refined dataset for big data analysis – 10 years on. **Nucleic Acids Research**, v. 44, n. D1, p. D694-D697, 2016.

CHEN, Y.; SIMMONDS, R. S.; TIMKOVICH, R. Proposed docking interface between peptidoglycan and the target recognition domain of zoocin A. **Biochemical and Biophysical Research Communications**, v. 441, n. 2, p. 297-300, 2013.

CHERNOV, V. M.; CHERNOVA, O. A.; MOUZYKANTOV, A. A.; LOPUKHOV, L. L.; AMINOV, R. I. Omics of antimicrobials and antimicrobial resistance. *Expert Opinion on Drug Discovery*, v. 14, p. 455-468, 2019.

CLANCY, J.; HAJJ, F. D.; PETITPAS, J. W.; YUAN, W. Cloning and Characterization of a Novel Macrolide Efflux Gene, *mreA*, from *Streptococcus agalactiae*. **Antimicrobial Agents and Chemotherapy**, v. 41, n. 12, p. 2719-2723, 1997.

CLAREBOUT, G.; VILLERS, C.; LECLERCQ, R. Macrolide Resistance Gene *mreA* of *Streptococcus agalactiae* Encodes a Flavokinase. **Antimicrobial Agents and Chemotherapy**, v. 45, n. 8, p. 2280-2286, 2001.

COTTER, P. D.; ROSS, R. P.; HILL, C. Bacteriocins – a viable alternative to antibiotics? **Nature Reviews Microbiology**, v. 11, p. 95-105, 2013.

FELDGARDEN, M.; BROVER, V.; HAFT, D. H.; PRASAD, A. B.; SLOTTA, D. J.; TOLSTOY, I.; TYSON, G. H.; ZHAO, S.; HSU, C. H.; MCDERMOTT, P. F.; TADESSE, D. A.; MORALES, C.; SIMMONS, M.; TILLMAN, G.; WASILENKO, J.; FOLSTER, J. P.; KLIMKE, W. Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates. **Antimicrobial Agents and Chemotherapy**, v. 63, n. 11, 2019.

FU, L.; NIU, B.; ZHU, Z.; WU, S.; LI, W. CD-HIT: accelerated for clustering the next-generation sequencing data. **Bioinformatics**, v. 28, n. 23, p. 3150-3152, 2012.

- GAO, X. Y.; ZHI, X. Y.; LI, H. W.; KLENK, H. P.; LI, W. J. Comparative Genomics of the Bacterial Genus *Streptococcus* Illuminates Evolutionary Implications of Species Groups. **PLoS ONE**, v. 9, n. 6, 2014.
- GARGIS, S. R.; GARGIS, A. S.; HEATH, H. E.; HEATH, L. S.; LEBLANC, P. A.; SENN, M. M.; BÄCHI, B. B.; SIMMONDS, R. S.; SLOAN, G. L. Zif, the Zoocin A Immunity Factor, Is a FemABX-like Immunity Protein with a Novel Mode of Action. **Applied and Environmental Microbiology**, v. 75, n. 19, p. 6205-6210, 2009.
- GHOSH, C.; SARKAR, P.; ISSA, R.; HALDAR, J. Alternatives to Conventional Antibiotics in the Era of Antimicrobial Resistance. **Trends in Microbiology**, v. 27, n. 4, p. 323-338, 2019.
- GRANT, J. R.; ARANTES, A. S.; STOTHARD, P. Comparing thousands of circular genomes using the CGView Comparison Tool. **BMC Genomics**, v. 13, n. 202, 2012.
- GRÜNING, B. A.; ERXLEBEN, A.; HÄHNLEIN, A.; GÜNTHER, S. Draft Genome Sequence of *Streptomyces viridochromogens* Strain Tü57, Producer of Avilamycin. **Genome Announcements**, v. 1, n. 3, 2013.
- GUPTA, S. K.; PADMANABHAN, B. R.; DIENE, S. M.; ROJAS, R. L.; KEMPF, M.; LANDRAUD, L.; ROLAIN, J. R. ARG-ANNOT, a New Bioinformatic Tool to Discover Antibiotic Resistance Genes in Bacterial Genomes. **Antimicrobial Agents and Chemotherapy**, v. 58, n. 1, p. 212-220, 2014.
- GUTEKUNST, H.; EIKMANN, B. J.; REINSCHEID, D. J. The Novel Fibrinogen-Binding Protein FbsB Promotes *Streptococcus agalactiae* Invasion into Epithelial Cells. **Infection and Immunity**, v. 72, n. 6, p. 3495-3504, 2004.
- HAENNI, M.; LUPO, A.; MADEC, J. Y. Antimicrobial Resistance in *Streptococcus* spp. **Microbiology Spectrum**, v. 6, n. 2, 2018.
- HAMMAMI, R.; ZOUHIR, A.; HAMIDA, J. B.; e FLISS, I. BACTIBASE: a new web-accessible database for bacteriocin characterization. **BMC Microbiology**, v. 7, n. 89, out. 2007.
- HASLAM, D. B.; GEME, J. W. S. Classification of Streptococci. *In*: LONG, S. S.; PROBER, C. G.; FISCHER, M. **Principles and Practice of Pediatric Infectious Diseases**, 5^a ed, p. 712-714, 2018.
- HEATH, L. S.; HEATH, H. E.; LEBLANC, P. A.; SMITHBERG, S. R.; DUFOUR, M.; SIMMONDS, R. S.; SLOAN, G. L. The streptococcolytic enzyme zoocin A is a penicillin-binding protein. **FEMS Microbiology Letters**, v. 236, n. 2, p. 205-211, 2004.
- HEATHER, Z.; HOLDEN, M. T. G.; STEWARD, K. F.; PARKHILL, J.; SONG, L.; CHALLIS, G. L.; ROBINSON, C.; POYNTER, N. D.; WALLER, A. S. A novel streptococcal integrative conjugative element involved in iron acquisition. **Molecular Microbiology**, v. 70, n. 5, p. 1274-1292, 2008.
- HÉCHARD, Y.; SAHL, H. G. Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria. **Biochimie**, v. 84, n. 5-6, p. 545-557, 2002.

- HENG, N. C. K.; BURTENSHAW, G. A.; JACK, R. W.; TAGG, J. R. Ubericin A, a Class IIa Bacteriocin Produced by *Streptococcus uberis*. **Applied and Environmental Microbiology**, v. 73, n. 23, p. 7763-7766, 2007.
- HERNANDEZ, D.; FRANÇOIS, P.; FARINELLI, L.; ØSTERÅS, M.; SCHRENZEL, J. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. **Genome Research**, v. 18, p. 802-809, 2008.
- IJAZ, M.; AMEEN, F.; ALFOTEIH, Y. A.; SHAMIM, S.; ALSHEHRI, W. A.; MURTAZA, G. Dissecting *Streptococcus pyogenes* interaction with human. **Archives of Microbiology**, v. 202, p. 2023-2032, 2020.
- JAIN, C.; RODRIGUEZ-R, L. M.; PHILLIPPY, A. M.; KONSTANTINIDIS, K. T.; ALURU, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. **Narute Communications**, v. 9, n. 5114, 2018.
- JANS, C.; MEILE, L.; LACROIX, C.; STEVENS, M. J. A. Genomics, evolution, and molecular epidemiology of the *Streptococcus bovis*/*Streptococcus equinus* complex (SBSEC). **Infection, Genetics and Evolution**, v. 33, p. 419-436, 2015.
- JERLSTRÖM, P. G.; TALAY, S. R.; WEIGAND, P. V.; TIMMIS, K. N.; CHHATWAL, G. Identification of an Immunoglobulin A Binding Motif Located in the β -Antigen of the c Protein Complex of Group B Streptococci. **Infection and Immunity**, v. 64, n. 7, p. 2787-2793, 1996.
- JIA, B.; RAPHENYA, A. R.; ALCOCK, B.; WAGLECHNER, N.; GUO, P.; TSANG, K. K.; LAGO, B. A.; DAVE, B. M.; PEREIRA, S.; SHARMA, A. N.; DOSHI, S.; COURTOT, M.; LO, R.; WILLIAMS, L. E.; FRYE, J. G.; ELSAYEGH, T.; SARDAR, D.; WESTMAN, E. L.; PAWLOWSKI, A. C.; JOHNSON, T. A.; BRINKMAN, F. S. L.; WRIGHT, G. D.; MCARTHUR, A. G. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. **Nucleic Acids Research**, v. 45, n. D1, p. D566-D573, 2017.
- JÚNIOR, J. C. R.; TAMANINI, R.; OLIVEIRA, A. L. M.; ALFIERI, A. A.; BELOTI, V. Genetic diversity of thermophilic spoilage microorganisms of milk from Brazilian dairy farms. **Journal of Dairy Science**, v. 101, n. 8, p. P6927-P6936, 2018.
- JÜRGENS, D.; STERZIK, B.; FEHRENBACH, F. J. Unspecific binding of group B streptococcal cocytolysin (CAMP Factor) to immunoglobulins and its possible role in pathogenicity. **Journal of Experimental Medicine**, v. 165, n. 3, p. 720-732, 1987.
- KANEHISA, M.; SATO, Y. KEGG Mapper for inferring cellular functions from protein sequences. **Protein Science**, v. 29, n. 1, p. 28-35, 2020.
- KAPATAI, G.; PATEL, D.; EFSTRATIOU, A.; CHALKER, V. J. Comparison of molecular serotyping approaches of *Streptococcus agalactiae* from genomic sequences. **BMC Genomics**, v. 18, n. 429, 2017.
- KOO, H. B.; SEO, J. Antimicrobial peptides under clinical investigation. **Peptide Science**, v. 111, n. 5, 2019.
- KRIVENTSEVA, E. V.; KUZNETSOV, D.; TEGENFELDT, F.; MANNI, M.; DIAS, R.; SIMÃO, F. A.; ZDOBNOV, E. M. OrthoDB v10: sampling the diversity of animal, plant, fungal,

protest, bacterial and viral genomes for evolutionary and functional annotations of orthologs. **Nucleic Acids Research**, v. 47, n. D1, p. D807-D811, 2019.

KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C.; TAMURA, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547-1549, 2018.

KUMARIYA, R.; GARSA, A. K.; RAJPUT, Y. S.; SOOD, S. K.; AKHTAR, N.; PATEL, S. Bacteriocins: classification, synthesis, mechanism of action and resistance development in food spoilage causing bacteria. **Microbial Pathogenesis**, v. 128, p. 171-177, 2019.

LANG, S.; PALMER, M. Characterization of *Streptococcus agalactiae* CAMP factor as a Pore-forming Toxin. **Journal of Biological Chemistry**, v. 278, n. 40, p. 38167-38173, 2003.

LARKIN, M. A.; BLACKSHIELDS, G.; BROWN, N. P.; CHENNA, R.; MCGETTIGAN, P. A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R.; THOMPSON, J. D.; GIBSON, T. J.; HIGGINS, D. G. Clustal W and Clustal X version 2.0. **Bioinformatics**, v. 23, n. 21, p. 1947-1948, 2007.

LAROCK, D. L.; RUSSELL, R.; JOHNSON, A. F.; WILDE, S.; LAROCK, C. N. Group A *Streptococcus* Infection of the Nasopharynx Requires Proinflammatory Signaling through the Interleukin-1 Receptor. **Infection and Immunity**, v. 88, n. 10, 2020.

LAXMINARAYAN, R.; BOECKEL, T. V.; FROST, I.; KARIUKI, S.; KHAN, E. A.; LIMMATHUROTSAKUL, D.; LARSSON, D. G. J.; HARA, G. L.; MENDELSON, M.; OUTTERSON, K.; PEACOCK, S. J.; ZHU, Y. G. *The Lancet Infectious Diseases* Commission on antimicrobial resistance: 6 years later. **The Lancet Infectious Diseases**, v. 20, n. 4, p. 51-60, 2020.

LAZZARO, B. P.; ZASLOFF, M.; ROLFF, J. Antimicrobial peptides: Application informed by evolution. **Science**, v. 368, n. 6490, 2020.

LEELAHAPONGSATHON, K.; SCHUKKEN, Y. H.; SRITHANASUWAN, A.; SURIYASATHAPORN, W. Molecular epidemiology of *Streptococcus uberis* intramammary infections: Persistent and transient patterns of infection in a dairy herd. **Journal of Dairy Science**, v. 103, n. 4, p. 3565-3576, 2020.

LICHVARIKOVÁ, A.; SOLTYS, K.; SZEMES, T.; SLOBODNIKOVA, L.; BUKOVSKA, G.; TURNA, J.; DRAHOVSKA, H. Characterization of Clinical and Carrier *Streptococcus agalactiae* and Prophage Contribution to the Strain Variability. **Viruses**, v. 12, n. 11, 2020.

LIU, L.; LI, Y.; LI, S.; HU, N.; HE, Y.; PONG, R.; LIN, D.; LU, L.; LAW, M. Comparison of Next-Generation Sequencing Systems. **Journal of Biomedicine and Biotechnology**, v. 2012, n. 251364, 2012.

MÄKINEN, V.; SALMELA, L.; YLINEN, J. Normalized N50 assembly metric using gap-restricted co-linear chaining. **BMC Bioinformatics**, v. 13, n. 255, 2012.

MCEWEN, S. A.; COLLIGNON, P. J. Antimicrobial Resistance: a One Health Perspective. **Microbiology Spectrum**, v. 6, n. 2, 2017.

- MEDEMA, M. H.; BLIN, K.; CIMERMANCIC, P.; JAGER, V.; ZAKRZEWSKI, P.; FISCHBACH, M. A.; WEBER, T.; TAKANO, E.; BREITLING, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, v. 39, p. W339-W346, 2011.
- MEIRA, M. M.; LACROIX, C.; LAPOINTE, G.; LAVOIE, M. C. Purification and structure of mutacin B-Ny266: a new lantibiotic produced by *Streptococcus mutans*. **FEBS Letters**, v. 410, p. 275-279, 1997.
- MELSTED, P.; HALLDÓRSSON, B. KmerStream: Streaming algorithms for *k*-mer abundance estimation. **Bioinformatics**, v. 30, n. 24, p. 3541-3547, 2014.
- METCALF, B. J.; CHOCHUA, S.; GERTZ JR., R. E.; HAWKINS, P. A.; RICALDI, J.; LI, Z.; WALKER, W.; TRAN, T.; RIVERS, J.; MATHIS, S.; JACKSON, D.; GLENNEN, A.; LYNFIELD, R.; MCGEE, L.; BEALL, B.; Active Bacterial Core Surveillance team. Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of currant invasive *Streptococcus agalactiae* recovered in the USA. **Clinical Microbiology and Infection**, v. 23, n. 8, p. 574.e7-574.e14, 2017.
- MIETHKE, M.; MARAHIEL, M. A. Siderophore-Based Iron Acquisition and Pathogen Control. **Microbiology and Molecular Biology Reviews**, v. 71, n. 3, p. 413-451, 2007.
- MIKHEENKO, A.; PRJIBELSKI, A.; SAVELIEV, V.; ANTIPOV, D.; GUREVICH, A. Versatile genome assembly evaluation with QUAST-LG. **Bioinformatics**, v. 34, n. 13, p. i142-i150, 2018.
- MORIYA, Y.; ITOH, M.; OKUDA, S.; YOSHIZAWA, A. C.; KANEHISA, M. KAAS: an automatic genome annotation and pathway reconstruction server. **Nucleic Acids Research**, v. 35, p. W182-W185, 2007.
- MOUSA, W. K.; ATHAR, B.; MERWIN, N. J.; MAGARVEY, N. A. Antibiotics and specialized metabolites from the human microbiota. **Natural Products Reports**, v. 34, n. 11, p. 1302-1331, 2017.
- NASCIMENTO, C. S.; SANTOS, N. F. B.; FERREIRA, R. C. C.; TADDEI, C. R. *Streptococcus agalactiae* in pregnant women in Brazil: prevalence, serotypes, and antibiotic resistance. **Brazilian Journal of Microbiology**, v. 50, p. 943-952, 2019.
- NES, I. F.; DIEP, D. B.; HOLO, H. Bacteriocin Diversity in *Streptococcus* and *Enterococcus*. **Journal of Bacteriology**, v. 189, n. 4, p. 1189-1198, 2007.
- NOVÁK, J.; CAUFIELD, P. W.; MILLER, E. J. Isolation and Biochemical Characterization of a Novel Lantibiotic Mutacin from *Streptococcus Mutans*. **Journal of Bacteriology**, v. 176, n. 14, p. 4316-4320, 1994.
- O'CONNOR, P. M.; KUNIYOSHI, T. M.; OLIVEIRA, R. P. S.; HILL, C.; ROSS, R. P.; COTTER, P. D. Antimicrobials for food and feed; a bacteriocin perspective. **Current Opinion in Biotechnology**, v. 61, p. 160-167, 2020.
- OGAKI, M. B.; FURLANETO, M. C.; MAIA, L. F. Revisão: Aspectos Gerais das Bacteriocinas. **Brazilian Journal of Food Technology**, v. 18, n. 4, p. 267-276, 2015.

- OKU, Y.; KUROKAWA, K.; ICHIHASHI, N.; SEKIMIZU, K. Characterization of the *Staphylococcus aureus mprF* gene, involved in lyzinylation of phosphatidylglycerol. **Microbiology**, v. 150, n. 1, 2004.
- O'NEILL, J. The Welcome Trust and the UK Department of Health. **Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations**. Reino Unido, 2014.
- ORTEGA, M. G. V.; GALLEGOS, A. C. F.; HERNÁNDEZ, J. L. M.; AGUILAR, C. N.; MOORILLÓN, G. V. N. Production of Bioactive Peptides from Lactic Acid Bacteria: A Sustainable Approach for Healthier Foods. **Comprehensive Reviews in Food Science and Food Safety**, v. 18, n. 4, p. 1039-1051, 2019.
- PARADA, J. L.; CARON, C. R.; MEDEIROS, A. B. P.; SOCCOL, C. R. Bacteriocins from Lactic Acid Bacteria: Purification, Properties and use as Biopreservatives. **Brazilian Archives of Biology and Technology**, v. 50, n. 3, p. 521-542, 2007.
- PATEL, I. B.; VAUGHN, R. H. Cellulolytic Bacteria Associated with Sloughing Spoilage of California Ripe Olives. **Applied Microbiology**, v. 25, n. 1, p. 62-69, 1973.
- PETERS, J.; PRICE, J.; LLEWELYN, M. Staphylococcal and streptococcal infections. **Bacterial Infections**, v. 45, n. 12, p. 727-734, 2017.
- PINGITORE, E. V.; SALVUCCI, E.; SESMA, F.; MACÍAS, M. E. N. Different strategies for purification of antimicrobial peptides from Lactic Acid Bacteria (LAB). *In*: VILAS, A. M. **Communicating Current Research and Educational Topics and Trends in Applied Microbiology**, FORMATEX 2007, p. 557-568.
- POMPILIO, A.; BONAVENTURA, G. D.; GHERARDI, G. An Overview on *Streptococcus bovis/Streptococcus equinus* Complex Isolates: Identification to the Species/Subspecies Level and Antibiotic Resistance. **International Journal of Molecular Sciences**, v. 20, n. 3, 2019.
- QI, F.; CHEN, P.; CAUFIELD, P. W. Purification of Mutacin III from Group III *Streptococcus mutans* UA787 and Genetic Analyses of Mutacin III Biosynthesis Genes. **Applied and Environmental Microbiology**, v. 65, n. 9, p. 3880-3887, 1999.
- QI, F.; CHEN, P.; CAUFIELD, P. W. The Group I Strain of *Streptococcus mutans*, UA140, Produces Both the Lantibiotic Mutacin I and a Nonlantibiotic Bacteriocin, Mutacin IV. **Applied and Environmental Microbiology**, v. 67, n. 1, p. 15-21, 2001.
- RAABE, V. N.; SHANE, A. L. Group B *Streptococcus* (*Streptococcus agalactiae*). **Microbiology Spectrum**, v. 7, n. 2, 2019.
- RAZZAQ, L. A.; SABRI, M. Isolation and Characterization of *Streptococcus agalactiae* from Woman Patients with Vaginitis in Hilla Province. **Medical Journal of Babylon**, v. 2, n. 3, 2005.
- REPKA, L. M.; CHEKAN, J. R.; NAIR, S. K.; DONK, W. A. V. D. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. **Chemical Reviews**, v. 117, p. 5457-5520, 2017.
- RICHARDS, V. P.; PALMER, S. R.; BITAR, P. D. P.; QIN, X.; WEINSTOCK, G. M.; HIGHLANDER, S. K.; TOWN, C. D.; BURNE, R. A.; STANHOPE, M. J. Phylogenomics and the Dynamic Genome Evolution of the Genus *Streptococcus*. **Genome Biology and Evolution**, v. 6, n. 4, p. 741-753, 2014.

ROCA, I.; AKOVA, M.; BAQUERO, F.; CARLET, J.; CAVALERI, M.; COENEN, S.; COHEN, J.; FINDLAY, D.; GYSEENS, I.; HEURE, O. E.; KAHLMETER, G.; KRUSE, H.; LAXMINARAYAN, R.; LIÉBANA, E.; LÓPEZ-CERERO, L.; MACGOWAN, A.; MARTINS, M.; RODRÍGUEZ-BAÑO, J.; ROLAIN, J. M.; SEGOVIA, C.; SIGAUQUE, B.; TACCONELLI, E.; WELLINGTON, E.; VILA, J. The global threat of antimicrobial resistance: science for intervention. **New Microbes and New Infections**, v. 6, p. 22-29, 2015.

ROSS, K. F.; RONSON, C. W.; TAGG, J. R. Isolation and Characterization of the Lantibiotic Salivaricin A and Its Structural Gene *salA* from *Streptococcus salivarius* 20P3. **Applied and Environmental Microbiology**, v. 59, n. 7, p. 2014-2021, 1993.

SANDERSON, J. M. Peptide-lipid interactions: insights and perspectives. **Organic and Biomolecular Chemistry**, v. 3, n. 2, p. 201-212, 2005.

SCHRÖNER, T. A.; GASSEL, S.; OSAWA, A.; TOBIAS, N. J.; OKUNO, Y.; SAKAKIBARA, Y.; SHINDO, K.; SANDMANN, G.; BODE, H. B. Aryl Polyenes, a Highly Abundant Class of Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids. **ChemBioChem**, v. 17, n. 3, p. 247-253, 2016.

SCHUBERT, A.; ZAKIKHANY, K.; SCHREINER, M.; FRANK, R.; SPELLERBERG, B.; EIKMANN, B. J.; REINSCHIED, D. J. A fibrinogen receptor from group B *Streptococcus* interacts with fibrinogen by repetitive units with novel ligand binding sites. **Molecular Biology**, v. 46, n. 2, p. 557-569, 2002.

SCHUBERT, M.; LINDGREEN, S.; ORLANDO, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. **BMC Research Notes**, v. 9, n. 88, 2016.

SCHWALEN, C. J.; HUDSON, G. A.; KILLE, B.; MITCHELL, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. **Journal of American Chemical Society**, v. 140, n. 30, 2018.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, v. 15, n. 30, p. 2068-2069, 2014.

SEPPEY, M.; MANNI, M.; ZDOBNOV, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. *In*: KOLLMAR, M. **Gene Prediction: Methods and Protocols**, Springer, p. 227-245, 2019.

SIEIRO, P. A.; LÓPEZ, M. M.; MU, D.; KUIPERS, O. P. Bacteriocins of lactic acid bacteria: extending the family. **Applied Microbiology and Biotechnology**, v. 100, p. 2939-2951, 2016.

SILVA, F. M. **Caracterização de bacteriocinas produzidas por *Streptococcus agalactiae* isolados de mastite bovina**. 2003. 50 f. Trabalho de conclusão de curso (Bacharelado em Microbiologia e Imunologia) – Instituto de Microbiologia Paulo de Góes/UFRJ, Rio de Janeiro, 2003.

SIMMONDS, R. S.; SIMPSON, W. J.; TAGG, J. R. Cloning and sequence analysis of *zooA*, a *Streptococcus zooepidemicus* gene encoding a bacteriocin-like inhibitory substance having a domain structure similar to that of lysostaphin. **Gene**, v. 189, n. 2, p. 255-261, 1997.

SOARES, S. C.; GEYIK, H.; RAMOS, R. T. J.; SÁ, P. H. C. G.; BARBOSA, E. G. V.; BAUMBACH, J.; FIGUEIREDO, H. C. P.; MIYOSHI, A.; TAUCH, A.; SILVA, A.;

- AZEVEDO, V. GIPSY: Genomic Island Prediction Software. **Journal of Biotechnology**, v. 232, p. 2-11, 2016.
- SPELLERBERG, B.; ROZDZINSKI, E.; MARTIN, S.; HEYNEMANN, J. W.; SCHNITZLER, N.; LÜTTICKEN, R.; PODBIELSKI, A. Lmb, a Protein with Similarities to the LraI Adhesin Family, Mediates Attachment of *Streptococcus agalactiae* to Human Laminin. **Infection and Immunity**, v. 67, n. 2, p. 871-878, 1999.
- STAUBITZ, P.; NEUMANN, H.; SCHNEIDER, T.; WIEDEMANN, I.; PESCHEL, A. MprF-mediated biosynthesis of lysylphosphatidylglycerol, an important determinant in staphylococcal defensin resistance. **FEMS Microbiology Letters**, v. 231, n. 1, p. 67-71, 2004.
- STOYANCHEVA, G.; MARZOTTO, M.; DELLAGLIO, F.; TORRIANI, S. Bacteriocin production and gene sequencing analysis from vaginal *Lactobacillus* strains. **Archives of Microbiology**, v. 196, n. 9, p. 645-653, 2014.
- STREICHER, L. M. Exploring the future of infectious disease treatment in a post-antibiotic era: A comparative review on alternative therapeutics. **Journal of Global Antimicrobial Resistance**, v. 24, p. 285-295, 2021.
- The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480-D489, 2021.
- THITIANANPAKORN, K.; AIBA, Y.; TAN, X. E.; WATANABE, S.; KIGA, K.; SATO'O, Y.; BOONSIRI, T.; LI, F. Y.; SASAHARA, T.; TAKI, Y.; AZAM, A. A.; ZHANG, Y.; CUI, L. Association of *mprF* mutations with cross-resistance to daptomycin and vancomycin in methicillin-resistant *Staphylococcus aureus* (MRSA). **Scientific Reports**, v. 10, n. 16107, 2020.
- THOMPSON, C. C.; EMMEL, V. E.; FONSECA, E. L.; MARIN, M. A.; VINCENTE, A. C. P. Streptococcal taxonomy based on genome sequence analysis. **F1000 Research**, 2:67, 2013.
- VAN HEEL, A. J.; JONG, A.; SONG, C.; VIEL, J. H.; KOK, J.; KUIPERS, O. P. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. **Nucleic Acids Research**, v. 46, p. W278-W281, 2018.
- VAN HEEL, A. J.; KLOOSTERMAN, T. G.; LOPEZ, M. M.; DENG, J. J.; PLAT, A.; BAUDU, B.; HENDRIKS, D.; MOLL, G. N.; KUIPERS, O. P. Discovery, production and modification of 5 novel lantibiotics using the promiscuous nisin modification machinery. **ACS Synthetic Biology**, v. 5, n. 10, p. 1146-1154, 2016.
- VÉLASQUEZ, J. E.; DONK, W. A. V. D. Genome mining for ribosomally synthesized natural products. **Current Opinion in Chemical Biology**, v. 15, p. 11-21, 2011.
- VILLANI, F.; PEPE, O.; MAURIELLO, G.; SALZANO, G.; MOSCHETTI, G.; COPPOLA, S. Antilisterial activity of thermophilin 347, a bacteriocin produced by *Streptococcus thermophilus*. **International Journal of Food Microbiology**, v. 25, p. 179-190, 1995.
- VOGEL, V.; SPELLERBERG, B. Bacteriocin Production by Beta-Hemolytic Streptococci. **Pathogens**, v. 10, n. 7, 867, 2021.

- VORNHAGEN, J.; WALDORF, K. M. A.; RAJAGOPAL, L. Perinatal Group B Streptococcal Infections: Virulence Factors, Immunity and Prevention Strategies. **Trends in Microbiology**, v. 25, n. 11, p. 919-931, 2017.
- VUYST, L. D.; LEROY, F. Bacteriocins from Lactic Acid Bacteria: Production, Purification, and Food Applications. **Journal of Molecular Microbiology and Biotechnology**, v. 13, p. 194-199, 2007.
- WESCOMBE, P. A.; UPTON, M.; DIERKSEN, K. P.; RAGLAND, N. L.; SIVABALAN, S.; WIRAWAN, R. E.; INGLIS, M. A.; MOORE, C. J.; WALKER, G. V.; CHILCOTT, C. N.; JENKINSON, H. F.; TAGG, J. R. Production of the Lantibiotic Salivaricin A and Its Variants by Oral Streptococci and Use of a Specific Induction Assay to Detect Their Presence in Human Saliva. **Applied and Environmental Microbiology**, v. 72, n. 2, p. 1459-1466, 2006.
- WESCOMBE, P. A.; UPTON, M.; RENAULT, P.; WIRAWAN, R. E.; POWER, D.; BURTON, J. P.; CHILCOTT, C. N.; TAGG, J. R. Salivaricin 9, a new lantibiotic produced by *Streptococcus salivarius*. **Microbiology**, v. 157, n. 5, p. 1290-1299, 2011.
- WIRAWAN, R. E.; SWANSON, K. M.; KLEFFMANN, T.; JACK, R. W.; TAGG, J. R. Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*. **Microbiology**, v. 153, n. 5, p. 1619-1630, 2007.
- World Health Organization – WHO. **Antimicrobial resistance and primary health care**. Technical Series on Primary Health Care, 2018.
- XIAO, H.; CHEN, X.; CHEN, M.; TANG, S.; ZHAO, X.; HUAN, L. Bovicin HJ50, a novel lantibiotic produced by *Streptococcus bovis* HJ50. **Microbiology**, v. 150, n. 1, p. 103-108, 2004.
- XU, L.; DONG, Z.; FANG, L.; LUO, Y.; WEI, Z.; GUO, H.; ZHANG, G.; GU, Y. Q.; DERR, D. C.; XIA, Q.; WANG, Y. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. **Nucleic Acids Research**, v. 47, n. 1, p. 52-58, 2019.
- YANG, S. C.; LIN, C. H.; SUNG, C. T.; e FANG, J. Y. Antibacterial activities of bacteriocins: application in foods and pharmaceuticals. **Frontiers in Microbiology**, v. 5, n. 241, 2014.
- YU, D.; XU, F.; ZENG, J.; ZHAN, J. Type III Polyketide Synthases in Natural Product Biosynthesis. **IUBMB Life**, v. 64, n. 4, p. 285-295, 2012.
- YU, G. Using ggtree to Visualize Data on Tree-Like Structures. **Current Protocols in Bioinformatics**, 2020.
- ZANKARI, E.; HASMAN, H.; COSENTINO, S.; VESTERGAARD, M.; RASMUSSEN, S.; LUND, O.; AARESTRUP, F. M.; LARSEN, M. V. Identification of acquired antimicrobial resistance genes. **Journal of Antimicrobial Chemotherapy**, v. 67, n. 11, p. 2640-2644, 2012.
- ZHENG, J.; GÄNZLE, M. G.; LIN, X. B.; RUAN, L.; SUN, M. Diversity and dynamics of bacteriocins from human microbiome. **Environmental Microbiology**, v. 17, n. 6, p. 2133-2143, 2015.

ANEXOS

**ANEXO A – RESULTADOS DOS TESTES DE ATIVIDADE ANTIMICROBIANA DE
SILVA (2003)**

| Estirpes de <i>Streptococcus</i> <i>agalactiae</i> | Indicadoras | | | | | |
|--|----------------|-------------------|------------------|--------------------------|--------------------|------------------|
| | <i>C. fimi</i> | <i>L. innocua</i> | <i>S. aureus</i> | <i>S. agalactiae</i> II4 | <i>S. pyogenes</i> | <i>L. lactis</i> |
| 3770 | - | - | - | + | - | - |
| 3771 | - | - | - | + | + | - |
| 3772 | - | - | - | + | + | - |
| 3773 | - | - | - | + | + | - |
| 3774 | - | - | - | + | + | - |
| 3775 | - | - | - | + | + | - |
| 3776 | - | - | - | + | + | - |
| 3797 | - | - | - | + | + | - |
| 3798 | - | - | - | + | + | - |
| 3799 | - | - | - | + | + | - |
| 3800 | - | - | - | + | + | - |
| 3801 | - | - | - | + | + | - |
| 3802 | - | - | - | + | + | - |
| 3804 | - | - | - | + | + | - |

OBS: Somente foram incluídas na tabela as estirpes que inibiram pelo menos uma das indicadoras.

Fonte: Caracterização de bacteriocinas produzidas por *Streptococcus agalactiae* isolados de mastite bovina (SILVA, 2003).

**ANEXO B – RESULTADOS DOS TESTES DE ATIVIDADE ANTIMICROBIANA
ADICIONAIS COM AS ESTIRPES 3772, 3774, 3799 E 3800**

| Indicadoras | | Estirpes BLS ⁺ | | | |
|----------------------|----------|---------------------------|------|------|------|
| Espécies | Estirpes | 3772 | 3774 | 3799 | 3800 |
| <i>S. agalactiae</i> | 3725 | + | + | + | + |
| | 3726 | + | + | + | + |
| | 3727 | + | + | + | + |
| | 3728 | + | + | + | + |
| | 3729 | + | + | + | + |
| | 3730 | + | + | + | + |
| | 3731 | + | + | + | + |
| | 3739 | + | + | + | + |
| | 3740 | + | + | + | + |
| | 3741 | + | + | + | + |
| | 3742 | + | + | + | + |
| | 3743 | + | + | + | + |
| | 3744 | - | - | - | - |
| | 3745 | + | + | + | + |
| | 3751 | + | + | + | + |
| 3752 | + | + | + | + | |
| <i>S. porcinus</i> | 628 | - | - | - | - |
| | 662 | + | + | + | + |
| | 790 | - | - | - | - |
| | 857 | + | + | + | + |
| | 1058 | + | + | + | + |
| | 1124 | + | + | - | - |
| | 1317 | + | + | + | + |
| | 1451 | - | + | - | - |
| | 2378 | - | - | - | - |
| | 3123 | + | + | + | + |
| | 3176 | - | - | - | - |
| 3658 | - | - | - | - | |
| <i>S. pyogenes</i> | 1465 | - | - | - | - |
| | 1471 | - | - | - | - |
| | 1474 | - | - | - | - |
| | 1996 | - | - | - | - |
| | 2009 | - | - | - | - |
| | 2586 | + | + | + | + |
| | 2587 | + | + | + | + |
| | 2588 | + | + | + | + |
| | 2590 | - | - | + | + |
| | 2591 | + | + | + | - |
| | 2593 | - | - | - | - |
| 2606 | - | - | - | - | |
| 2608 | - | - | - | - | |

Continua

| Indicadoras | | Conclusão | | | |
|--------------------|----------|-----------|------|------|------|
| Espécies | Estirpes | 3772 | 3774 | 3799 | 3800 |
| <i>S. pyogenes</i> | 2612 | - | - | - | - |
| | 2615 | - | - | - | - |
| | 2617 | - | - | - | - |
| | 2618 | + | + | + | + |
| | 2620 | - | - | - | - |
| | 2625 | + | - | - | - |
| <i>S. uberis</i> | 602 | + | + | - | - |
| | 752 | - | - | - | - |
| | 959 | - | - | - | - |
| | 2825 | + | - | - | - |
| | 3355 | + | - | - | + |
| | 3351 | + | + | + | + |
| | 3354 | + | + | + | + |
| | 3431 | - | - | - | - |
| | 3485 | + | - | + | - |
| | 3670 | + | - | + | + |
| | 3671 | + | + | + | + |
| 3724 | - | - | - | - | |

Fonte: Caracterização de bacteriocinas produzidas por *Streptococcus agalactiae* isolados de mastite bovina (SILVA, 2003).

ANEXO C – PROTOCOLO DE EXTRAÇÃO DE DNA POR FENOL-CLOROFÓRMIO-ÁLCOOL ISOAMÍLICO (CHACHATY; SAULNIER, 2000) MODIFICADO

1. Crescer 25 mL de cultura bacteriana *overnight*;
2. Centrifugar o crescimento em centrífuga refrigerada CT-5000R para tubos *falcon* a 4°C e 5.000 RPM por 15 minutos, descartar o sobrenadante em seguida;
3. Lavar o *pellet* com 1 mL de água destilada e centrifugá-lo em microcentrífuga refrigerada NT805 a 4°C e 6.000 RPM por 15 minutos, descartar o sobrenadante em seguida;
4. Ressuspender o *pellet* em 450 µL de tampão TE (Tris 50 mM, EDTA 50 mM) e adicionar 25 µL de SDS 20%, lisozima na concentração final de 0,5 mg/mL, 20 µL de Rnase (200µg/mL) e 2,5 µL de protease (20 mg/mL);
5. Incubar o material a 37°C por 2 horas;
6. Adicionar 220 µL de NaCl 5M, homogeneizar e adicionar 700 µL de solução clorofórmio:álcool isoamílico na proporção 24:1;
7. Agitar o material por inversão por 15 minutos e centrifugá-lo a 4°C e 10.000 xg por 30 minutos;
8. Remover a fase aquosa superior, evitando contato com a fase proteica intermediária;
9. Adicionar 700 µL de solução clorofórmio:álcool isoamílico na proporção 24:1 e repetir a etapa 7 e 8, visando maior grau de purificação do material extraído;
10. Adicionar 70 µL de acetato de sódio 3M pH 7,0 e 700 µL de isopropanol gelado, e incubar o material a -20°C por 1 hora;
11. Centrifugar a 4°C e 10.000 xg por 10 minutos para precipitar o DNA;
12. Lavar o *pellet* com 800 µL de etanol 70% gelado e centrifugar a 4°C e 10.000 xg por 5 minutos, descartar o sobrenadante em seguida;
13. Deixar o etanol evaporar completamente;
14. Ressuspender o *pellet* em 100 µL de água nuclease-*free* e armazenar a -20°C.

ANEXO D – RESULTADOS COMPLETOS DA ANÁLISE DE ANI.

| Espécies de referência | Estirpe | | |
|--|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| <i>S. agalactiae</i> NCTC 8181 | 98,87% | 98,81% | 98,77% |
| <i>S. thermophilus</i> NCTC 12958 | 82,90% | 80,61% | 80,69% |
| <i>S. pyogenes</i> NCTC 8198 | 82,69% | 82,95% | 82,78% |
| <i>S. equi</i> subsp. <i>ruminatorum</i> CECT 5772 | 82,03% | 81,98% | 81,87% |
| <i>S. dysgalactiae</i> subsp. <i>dysgalactiae</i> NCTC 13731 | 81,99% | 82,10% | 82,51% |
| <i>S. constellatus</i> subsp. <i>pharyngis</i> SK1060 | 81,39% | 81,80% | 81,49% |
| <i>S. dysgalactiae</i> subsp. <i>equisimilis</i> NCTC 13762 | 80,79% | 80,59% | 80,65% |
| <i>S. equi</i> subsp. <i>zoepidemicus</i> NCTC 4676 | 80,27% | 80,39% | 80,23% |
| <i>S. pseudoporcinus</i> NCTC 13786 | 79,82% | 80,08% | 79,68% |
| <i>S. anginosus</i> subsp. <i>anginosus</i> NCTC 10713 | 79,81% | 79,57% | 79,76% |
| <i>S. oralis</i> subsp. <i>tigurinus</i> AZ 3a | 79,70% | 80,23% | 79,70% |
| <i>S. equi</i> subsp. <i>equi</i> ATCC 33398 | 79,68% | 79,19% | 79,24% |
| <i>S. uberis</i> NCTC 3858 | 79,59% | 79,68% | 79,72% |
| <i>S. salivarius</i> subsp. <i>salivarius</i> NCTC 8618 | 79,56% | 79,72% | 79,89% |
| <i>S. constellatus</i> subsp. <i>constellatus</i> NCTC 11325 | 79,47% | 79,72% | 79,59% |
| <i>S. equinus</i> NCTC 12969 | 79,36% | 79,68% | 79,39% |
| <i>S. pseudopneumoniae</i> CCUG 49455 | 79,32% | 79,39% | 79,70% |
| <i>S. gordonii</i> NCTC 7865 | 79,32% | 79,51% | 79,43% |
| <i>S. anginosus</i> subsp. <i>whileyi</i> CCUG 39159 | 79,30% | 79,02% | 79,05% |
| <i>S. mitis</i> NCTC 12261 | 79,27% | 79,51% | 79,65% |
| <i>S. parasanguinis</i> ATCC 15912 | 79,23% | 79,55% | 79,43% |
| <i>S. sanguinis</i> NCTC 7863 | 79,16% | 79,77% | 79,49% |
| <i>S. vestibularis</i> ATCC 49124 | 79,13% | 79,11% | 79,16% |
| <i>S. gallolyticus</i> subsp. <i>macedonicus</i> NCTC 13767 | 79,12% | 79,52% | 79,15% |
| <i>S. pneumoniae</i> NCTC 7645 | 79,06% | 79,46% | 79,34% |
| <i>S. oralis</i> subsp. <i>oralis</i> NCTC 11427 | 79,03% | 79,78% | 79,14% |
| <i>S. gallolyticus</i> subsp. <i>pasteurianus</i> NCTC 13784 | 78,95% | 78,86% | 78,79% |
| <i>S. downei</i> NCTC 11391 | 78,92% | 79,63% | 79,15% |
| <i>S. sobrinus</i> NCTC 12279 | 78,76% | 79,30% | 79,18% |
| <i>S. oralis</i> subsp. <i>dentisani</i> CECT 7747 | 78,74% | 79,01% | 78,78% |
| <i>S. gallolyticus</i> subsp. <i>gallolyticus</i> DSM 16831 | 78,72% | 78,89% | 78,78% |
| <i>S. intermedius</i> NCTC 11324 | 78,59% | 78,68% | 78,97% |

Continua

| Espécies de referências | Conclusão | | |
|-----------------------------------|-------------|-------------|-------------|
| | LGMAI_St_08 | LGMAI_St_11 | LGMAI_St_14 |
| <i>S. criceti</i> HS-6 | 78,20% | 78,88% | 78,72% |
| <i>S. mutans</i> NCTC 10449 | 78,19% | 78,18% | 78,15% |
| <i>S. acidominimus</i> NCTC 12957 | 79,27% | 79,75% | 79,49% |
| <i>S. suis</i> S 735 | 79,08% | 79,54% | 79,50% |

ANEXO E: RESULTADOS COMPLETOS DA ANOTAÇÃO PELA FERRAMENTA BAGEL4.

| Estirpe (Região) | Localização | Gene/ORF | Função |
|---------------------------------------|----------------|---------------|---|
| LGMAI_St_08 (NODE_1_length_142099) | 86166-89891 | Orf00001 | Fosforibosilformilglicinamidina sintase subunidade PurL OS= <i>Thermoplasma volcanium</i> (ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1) OX=273116 GN=purL PE=3 SV=1 |
| | 90125-91579 | Orf00002 | Amidofosforibosiltransferase OS= <i>Staphylococcus aureus</i> (MW2) OX=196620 GN=purF PE=3 SV=1 |
| | 91607-92632 | Orf00003 | Fosforibosilformilglicinamidina ciclo-ligase OS= <i>Streptococcus agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purM PE=3 SV=1 |
| | 92895-93466 | Orf00005 | Fosforibosilglicinamida formiltransferase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=purN PE=3 SV=2 |
| | 93469-94221 | Orf00006 | N-acetiltransferase não caracterizada YdfB OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=ydfB PE=3 SV=1 |
| | 94241-95788 | Orf00008 | Proteína de biossíntese de purina bifuncional PurH OS= <i>Streptococcus agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purH PE=3 SV=1 |
| | 95981-96880 | 93.3;Zoocin_A | Zoocina A |
| | 97027-98331 | Orf000010 | |
| | 98578-99276 | Orf000012 | N-acetilmanosamina-6-fosfato 2-epimerase putativa OS= <i>Streptococcus agalactiae</i> sorotipo Ia (ATCC 27591 / A909 / CDC SS700) OX=205921 GN=nanE PE=3 SV=1 |
| | 99323-100639 | Orf000013 | Proteína de ligação ao substrato transportadora ABC putativa YesO OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yesO PE=1 SV=2 |
| | 100760-101614 | Orf000014 | Provável permease de transportador ABC protein YurN OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yurN PE=3 SV=1 |
| | 101624-102454 | Orf000015 | Proteína permease do sistema de transporte de L-arabinose AraQ OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=araQ PE=2 SV=2 |
| | 102467-102910 | Orf000016 | Proteína de biofilme de toxina-antitoxina TabA OS= <i>Escherichia coli</i> (K12) OX=83333 GN=tabA PE=1 SV=1 |
| | 102930-103592 | Orf000017 | |
| | 103589-104506 | Orf000019 | N-acetilneuraminato liase OS= <i>Salmonella newport</i> (SL254) OX=423368 GN=nanA PE=3 SV=1 |
| | 104523-105404 | Orf000021 | Proteína não caracterizada CPE0188 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0188 PE=3 SV=2 |
| | 105412-1063899 | Orf000022 | Cefalosporina-C desacetilase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=cah PE=1 SV=1 |

Continua

| Estirpe (Região) | Localização | Continuação | |
|---------------------------------------|---------------|---------------|---|
| | | Gene/ORF | Função |
| LGMAI_St_08 (NODE_1_length_142099) | 106412-106864 | Orf000023 | Regulador transcricional não caracterizado do tipo HTH CPE0189 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0189 PE=4 SV=1 |
| | 92674-92796 | sORF_1 | |
| | 89995-90105 | sORF_2 | |
| | 89897-89998 | sORF_3 | |
| LGMAI_St_11 (L2__8) | 3347-7072 | Orf00001 | Subunidade da fosforibosilformilglicinamidina sintase PurL OS= <i>Thermoplasma volcanium</i> (ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1) OX=273116 GN=purL PE=3 SV=1 |
| | 7306-8760 | Orf00002 | Amidofosforibosiltransferase OS= <i>Staphylococcus aureus</i> (MW2) OX=196620 GN=purF PE=3 SV=1 |
| | 8788-9813 | Orf00003 | Fosforibosilformilglicinamidina ciclo-ligase OS= <i>Streptococcus agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purM PE=3 SV=1 |
| | 10076-10627 | Orf00005 | Fosforibosilglicinamida formiltransferase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=purN PE=3 SV=2 |
| | 10650-11402 | Orf00006 | N-acetiltransferase não caracterizada YdfB OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=ydfB PE=3 SV=1 |
| | 11422-12969 | Orf00008 | Proteína de biossíntese de purina bifuncional PurH OS= <i>Streptococcus</i> <i>agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purH PE=3 SV=1 |
| | 13162-14061 | 93.3;Zoocin_A | Zoocina A |
| | 14208-15512 | Orf000010 | |

Continua

| Continuação | | | |
|------------------------|-------------|-----------|--|
| Estirpe (Região) | Localização | Gene/ORF | Função |
| LGMAI_St_11 (L2__8) | 15759-16457 | Orf000012 | N-acetilmanosamina-6-fosfato 2-epimerase putativa OS= <i>Streptococcus agalactiae</i> sorotipo Ia (ATCC 27591 / A909 / CDC SS700) OX=205921 GN=nanE PE=3 SV=1 |
| | 16504-17820 | Orf000013 | Proteína de ligação ao substrato transportadora ABC putativa YesO OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yesO PE=1 SV=2 |
| | 17941-18795 | Orf000014 | Provável permease de transportador ABC protein YurN OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yurN PE=3 SV=1 |
| | 18805-19635 | Orf000015 | Proteína permease do sistema de transporte de L-arabinose AraQ OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=araQ PE=2 SV=2 |
| | 19648-20091 | Orf000016 | Proteína de biofilme de toxina-antitoxina TabA OS= <i>Escherichia coli</i> (K12) OX=83333 GN=tabA PE=1 SV=1 |
| | 20111-20773 | Orf000017 | |
| | 20770-21687 | Orf000019 | N-acetilneuraminato liase OS= <i>Salmonella newport</i> (SL254) OX=423368 GN=nanA PE=3 SV=1 |
| | 21704-22585 | Orf000021 | Proteína não caracterizada CPE0188 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0188 PE=3 SV=2 |
| | 22593-23570 | Orf000022 | Cefalosporina-C desacetilase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=cah PE=1 SV=1 |
| | 23593-24045 | Orf000023 | Regulador transcricional não caracterizado do tipo HTH CPE0189 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0189 PE=4 SV=1 |

Continua

| Continuação | | | |
|------------------------|-------------|-----------|--|
| Estirpe (Região) | Localização | Gene/ORF | Função |
| LGMAI_St_11 (L2__8) | 9855-9977 | sORF_1 | Regulador transcricional não caracterizado do tipo HTH CPE0189 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0189 PE=4 SV=1 |
| | 7176-7286 | sORF_2 | |
| | 7078-7179 | sORF_3 | |
| LGMAI_St_14 (L2__7) | 55629-56081 | Orf00001 | Regulador transcricional não caracterizado do tipo HTH CPE0189 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0189 PE=4 SV=1 |
| | 54629-55606 | Orf00002 | Cefalosporina-C desacetilase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=cah PE=1 SV=1 |
| | 53740-54621 | Orf00003 | Proteína não caracterizada CPE0188 OS= <i>Clostridium perfringens</i> (13 / Tipo A) OX=195102 GN=CPE0188 PE=3 SV=2 |
| | 52806-53723 | Orf00005 | N-acetilneuraminato liase OS= <i>Salmonella newport</i> (SL254) OX=423368 GN=nanA PE=3 SV=1 |
| | 52147-52809 | Orf00007 | |
| | 51684-52127 | Orf00008 | Proteína de biofilme de toxina-antitoxina TabA OS= <i>Escherichia coli</i> (K12) OX=83333 GN=tabA PE=1 SV=1 |
| | 50841-51671 | Orf00009 | Proteína permease do sistema de transporte de L-arabinose AraQ OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=araQ PE=2 SV=2 |
| | 49977-50831 | Orf000010 | Provável permease de transportador ABC protein YurN OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yurN PE=3 SV=1 |
| | 48540-49856 | Orf000011 | Proteína de ligação ao substrato transportadora ABC putativa YesO OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=yesO PE=1 SV=2 |

Continua

| Estirpe (Região) | Localização | Conclusão | |
|------------------------|-------------|---------------|---|
| | | Gene/ORF | Função |
| LGMAI_St_14 (L2__7) | 47795-48493 | Orf000012 | N-acetilmanosamina-6-fosfato 2-epimerase putativa OS= <i>Streptococcus agalactiae</i> sorotipo Ia (ATCC 27591 / A909 / CDC SS700) OX=205921 GN=nanE PE=3 SV=1 |
| | 46244-47548 | Orf000014 | |
| | 45198-46097 | 93.3;Zoocin_A | Zoocina A |
| | 43458-45005 | Orf000016 | Proteína de biossíntese de purina bifuncional PurH OS= <i>Streptococcus agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purH PE=3 SV=1 |
| | 42686-43438 | Orf000018 | N-acetiltransferase não caracterizada YdfB OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=ydfB PE=3 SV=1 |
| | 42112-42663 | Orf000019 | Fosforibosilglicinamida formiltransferase OS= <i>Bacillus subtilis</i> (168) OX=224308 GN=purN PE=3 SV=2 |
| | 40824-41849 | Orf000021 | Fosforibosilformilglicinamida ciclo-ligase OS= <i>Streptococcus agalactiae</i> sorotipo III (NEM316) OX=211110 GN=purM PE=3 SV=1 |
| | 39342-40796 | Orf000022 | Amidofosforibosiltransferase OS= <i>Staphylococcus aureus</i> (MW2) OX=196620 GN=purF PE=3 SV=1 |
| | 35383-39108 | Orf000023 | Subunidade da fosforibosilformilglicinamida sintase PurL OS= <i>Thermoplasma volcanium</i> (ATCC 51530 / DSM 4299 / JCM 9571 / NBRC 15438 / GSS1) OX=273116 GN=purL PE=3 SV=1 |
| | 39212-39322 | sORF_1 | |
| | 39114-39215 | sORF_2 | |
| 41891-42013 | sORF_3 | | |

ANEXO F: RESULTADOS COMPLETOS DA ANÁLISE DE VIRULÊNCIA REALIZADA COM A FERRAMENTA ABRICATE, USANDO O BANCO DE DADOS VFDB.

| Estirpe | Contig | Localização | Gene | Cobertura | Identidade | Produto |
|----------------|---------------|--------------------|-------------|------------------|-------------------|--|
| LGMAI_St_08 | L2__30 | 8026-9267 | <i>neuA</i> | 100% | 99,60% | CMP-N-acetilneuramínico sintetase de ácido NeuA [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 9278-9907 | <i>neuD</i> | 100% | 99,68% | Proteína neuD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 9904-11058 | <i>neuC</i> | 100% | 98,96% | UDP-N-acetilglucosamina-2-epimerase NeuC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 11135-12160 | <i>neuB</i> | 100% | 99,71% | Sintetase de ácido N-acetilneuramínico NeuB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 12160-13560 | <i>cpsL</i> | 100% | 97,50% | Proteína de biossíntese de polissacarídeo CpsL [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 13557-14513 | <i>cpsK</i> | 100% | 98,33% | Proteína de biossíntese de polissacarídeo CpsK (V) [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18157-18606 | <i>cpsF</i> | 100% | 99,56% | Proteína de biossíntese de polissacarídeo CpsF [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18630-20009 | <i>cpsE</i> | 99,35% | 99,35% | Glicosiltransferase CpsE [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|----------------------|-------------|----------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_08 | L2__30 | 20025-20723 | <i>cpsD</i> | 100% | 98,43% | Autoquinase CpsD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 20734-21426 | <i>cpsC</i> | 100% | 99,71% | Proteína CpsC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 21435-22166 | <i>cpsB</i> | 100% | 99,86% | Fosfatase CpsB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 22172-23629 | <i>cpsA</i> | 100% | 99,45% | CpsA responsável pela transferência de CPS do lipídeo da membrana para o peptidoglicano da parede celular [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | L2__5 | 34345-37563 | <i>hylB</i> | 100% | 98,94% | Hialuronidase [Hialuronato-liase (VF0281)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | L2__61 | 16506-17273 | <i>cfa/cfb</i> | 100% | 98,96% | Fator CAMP [Fator CAMP (VF0280)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | NODE_12_length_51642 | 7665-8240 | <i>cylK</i> | 100% | 99,13% | Fosfopantetoinil transferase CylK putativa [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 8209-9420 | <i>cylJ</i> | 100% | 98,84% | Glicosiltransferase CylJ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus</i> <i>agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|----------------------|-------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_08 | NODE_12_length_51642 | 9425-11620 | <i>cylI</i> | 100% | 99,18% | 3-cetoacil-ACP sintase putativa CylI [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 11617-12570 | <i>cylF</i> | 100% | 99,06% | Aminometiltransferase CylF putativa [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 14567-15445 | <i>cylB</i> | 100% | 99,43% | Transportador CylB ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 15438-16367 | <i>cylA</i> | 100% | 100% | Transportador CylA ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 16367-16833 | <i>cylZ</i> | 100% | 100% | 3R-hidroximiristoil ACP desidratase CylZ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 16817-17122 | <i>acpC</i> | 100% | 100% | Proteína transportadora de acil AcpC [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 17115-17837 | <i>cylG</i> | 100% | 100% | 3-cetoacil-ACP-redutase CylG [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|----------------------|---------------------|-------------|-----------|------------|--|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_08 | NODE_12_length_51642 | 17834-18682 | <i>cylD</i> | 100% | 100% | Malonil-CoA-ACP transacilase CylD [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18682-18987 | <i>cylX</i> | 100% | 100% | Acetil coenzima A (CoA) carboxilase CylX [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| LGMAI_St_11 | L2__29 | NODE_9_length_52512 | <i>hylB</i> | 100% | 98,94% | Hialuronidase [Hialuronato-liase (VF0281)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 8026-9267 | <i>neuA</i> | 100% | 99,60% | CMP-N-acetilneuramínico sintetase de ácido NeuA [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 9278-9907 | <i>neuD</i> | 100% | 99,68% | Proteína neuD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 9904-11058 | <i>neuC</i> | 100% | 98,96% | UDP-N-acetilglucosamina-2-epimerase NeuC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 11135-12160 | <i>neuB</i> | 100% | 99,71% | Sintetase de ácido N-acetilneuramínico NeuB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 12160-13560 | <i>cpsL</i> | 100% | 97,50% | Proteína de biossíntese de polissacarídeo CpsL [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|--------|-------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_11 | L2__29 | 13557-14513 | <i>cpsK</i> | 100% | 98,33% | Proteína de biossíntese de polissacarídeo CpsK (V) [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18158-18607 | <i>cpsF</i> | 100% | 99,56% | Proteína de biossíntese de polissacarídeo CpsF [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18631-20010 | <i>cpsE</i> | 99,35% | 99,35% | Glicosiltransferase CpsE [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 20026-20724 | <i>cpsD</i> | 100% | 98,43% | Autoquinase CpsD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 20735-21427 | <i>cpsC</i> | 100% | 99,71% | Proteína CpsC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 21436-21167 | <i>cpsB</i> | 100% | 99,86% | Fosfatase CpsB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 22173-23630 | <i>cpsA</i> | 100% | 99,45% | CpsA responsável pela transferência de CPS do lipídeo da membrana para o peptidoglicano da parede celular [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | L2__9 | 33852-37070 | <i>hylB</i> | 100% | 98,94% | Hialuronidase [Hialuronato-liase (VF0281)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|---------------------|-----------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_11 | NODE_9_length_51476 | 32659-32964 | <i>cylX</i> | 100% | 100% | Acetil coenzima A (CoA) carboxilase CylX [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 32964-33812 | <i>cylD</i> | 100% | 100% | Malonil-CoA-ACP transacilase CylD [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 33809-34531 | <i>cylG</i> | 100% | 100% | 3-cetoacil-ACP-redutase CylG [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 34524-34829 | <i>acpC</i> | 100% | 100% | Proteína transportadora de acil AcpC [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 34813-35289 | <i>cylZ</i> | 100% | 100% | 3R-hidroximiristoil ACP desidratase CylZ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 35279-36208 | <i>cylA</i> | 100% | 100% | Transportador CylA ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 36201-37079 | <i>cylB</i> | 100% | 99,43% | Transportador CylB ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | Continua | | | | |

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|---------------------|-------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_11 | NODE_9_length_51476 | 39076-40029 | <i>cylF</i> | 100% | 99,06% | Aminometiltransferase CylF putativa [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 40026-42221 | <i>cylI</i> | 100% | 99,18% | 3-cetoacil-ACP sintase putativa CylI [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 42226-43437 | <i>cylJ</i> | 100% | 98,84% | Glicosiltransferase CylJ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 43406-43981 | <i>cylK</i> | 100% | 99,13% | Fosfopantetoinil transferase CylK putativa [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| LGMAI_St_14 | L2__29 | 5634-7091 | <i>cpsA</i> | 100% | 99,45% | CpsA responsável pela transferência de CPS do lipídeo da membrana para o peptidoglicano da parede celular [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 7097-7828 | <i>cpsB</i> | 100% | 99,86% | Fosfatase CpsB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 7837-8529 | <i>cpsC</i> | 100% | 99,71% | Proteína CpsC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 8540-9238 | <i>cpsD</i> | 100% | 98,43% | Autoquinase CpsD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|--------|-----------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_14 | L2__29 | 9254-10633 | <i>cpsE</i> | 99,35 | 99,35% | Glicosiltransferase CpsE [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 10657-11106 | <i>cpsF</i> | 100% | 99,56% | Proteína de biossíntese de polissacarídeo CpsF [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 14751-15707 | <i>cpsK</i> | 100% | 98,33% | Proteína de biossíntese de polissacarídeo CpsK (V) [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 15704-17104 | <i>cpsL</i> | 100% | 97,50% | Proteína de biossíntese de polissacarídeo CpsL [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 17104-18129 | <i>neuB</i> | 100% | 99,71% | Sintetase de ácido N-acetilneuramínico NeuB [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18206-19360 | <i>neuC</i> | 100% | 98,96% | UDP-N-acetilglucosamina-2-epimerase NeuC [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 19357-19986 | <i>neuD</i> | 100% | 99,68% | Proteína neuD [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 19997-21238 | <i>neuA</i> | 100% | 99,60% | CMP-N-acetilneuramínico sintetase de ácido NeuA [cápsula (VF0274)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | Continua | | | | |

| Estirpe | Contig | Localização | Continuação | | | Produto |
|-------------|--------|-------------|-------------|-----------|------------|--|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_14 | L2__8 | 16350-19568 | <i>hylB</i> | 100% | 98,94% | Hialuronidase [Hialuronato-liase (VF0281)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | L2__9 | 7863-8438 | <i>cylK</i> | 100% | 99,13% | Fosfopantetoinil transferase CylK putativa [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 8407-9618 | <i>cylJ</i> | 100% | 98,84% | Glicosiltransferase CylJ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus</i> <i>agalactiae</i> 2603V/R] |
| | | 9623-11818 | <i>cylI</i> | 100% | 99,18% | 3-cetoacil-ACP sintase putativa CylI [Beta- hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 11815-12768 | <i>cylF</i> | 100% | 99,06% | Aminometiltransferase CylF putativa [Beta- hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 14765-15643 | <i>cylB</i> | 100% | 99,43% | Transportador CylB ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 15636-16565 | <i>cylA</i> | 100% | 100% | Transportador CylA ABC (cassete de ligação de ATP) [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |

Continua

| Estirpe | Contig | Localização | Conclusão | | | Produto |
|-------------|--------|-------------|-------------|-----------|------------|---|
| | | | Gene | Cobertura | Identidade | |
| LGMAI_St_14 | L2_9 | 16555-17031 | <i>cylZ</i> | 100% | 100% | 3R-hidroximiristoil ACP desidratase CylZ [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 17015-17320 | <i>acpC</i> | 100% | 100% | Proteína transportadora de acil AcpC [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 17313-18035 | <i>cylG</i> | 100% | 100% | 3-cetoacil-ACP-redutase CylG [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18032-18880 | <i>cylD</i> | 100% | 100% | Malonil-CoA-ACP transacilase CylD [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |
| | | 18880-19185 | <i>cylX</i> | 100% | 100% | Acetil coenzima A (CoA) carboxilase CylX [Beta-hemolisina / citolisina (CVF171)] [<i>Streptococcus agalactiae</i> 2603V/R] |

ANEXO G: RESULTADOS COMPLETOS DA PREDIÇÃO DE ILHAS GENÔMICAS DE RESISTÊNCIA E PATOGENICIDADE, REALIZADA NO SOFTWARE GIPSY.

| Estirpe | Ilha genômica | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | Posição | Escore |
|----------------|----------------------|---------------------|--------------------------------|------------------------------|------------------------------|-----------------|---------------|
| LGMAI_St_08 | Resistência 1 | 21% | 73% | 15% | 36% | 174580-199440 | Normal |
| | Genômica 1 | 15% | 29% | 1% | 71% | 569353-667503 | NA |
| | Genômica 2 | 66% | 66% | 0% | 66% | 694948-735864 | NA |
| | Resistência 2 | 28% | 28% | 12% | 68% | 939710-987780 | Normal |
| | Genômica 3 | 30% | 0% | 7% | 38% | 1106326-1118347 | NA |
| | Resistência 3 | 27% | 36% | 18% | 45% | 1128892-1164378 | Normal |
| | Resistência 4 | 14% | 0% | 42% | 71% | 1259251-1268895 | Normal |
| | Genômica 4 | 3% | 69% | 3% | 81% | 1368910-1453486 | NA |
| | Resistência 5 | 13% | 71% | 8% | 86% | 1545086-1602299 | Forte |
| | Genômica 5 | 0% | 73% | 0% | 60% | 1627053-1638691 | NA |
| | Resistência 6 | 9% | 17% | 14% | 55% | 1652611-1843228 | Fraco |
| | Genômica 6 | 20% | 20% | 5% | 66% | 1848454-1901176 | NA |
| | Genômica 7 | 43% | 87% | 0% | 82% | 2110992-2183661 | NA |
| | Genômica 8 | 14% | 42% | 0% | 60% | 2198527-2220011 | NA |
| | Genômica 9 | 50% | 66% | 0% | 66% | 2319833-2325485 | NA |

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore |
|-------------|------------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_08 | Genômica 10 | 7% | 51% | 7% | 80% | 2363091-2468550 | NA |
| | Genômica 11 | 28% | 42% | 0% | 28% | 2482022-2493060 | NA |
| | Genômica 12 | 5% | 29% | 5% | 52% | 2532436-2555967 | NA |
| | Genômica 13 | 20% | 17% | 6% | 51% | 2726186-2758174 | NA |
| | Genômica 14 | 9% | 77% | 0% | 81% | 2847714-2876401 | NA |
| | Genômica 15 | 15% | 33% | 6% | 63% | 2913369-2964353 | NA |
| | Genômica 16 | 55% | 72% | 0% | 72% | 2971199-2991361 | NA |
| | Genômica 17 | 29% | 26% | 2% | 67% | 3023519-3066619 | NA |
| | Genômica 18 | 21% | 12% | 6% | 51% | 3094267-3134645 | NA |
| | Genômica 19 | 26% | 41% | 2% | 78% | 3237548-3371673 | NA |
| | Genômica 20 | 25% | 35% | 2% | 85% | 3403129-3461159 | NA |
| | Patogenicidade 1 | 21% | 73% | 57% | 36% | 174580-199440 | Normal |
| | Genômica 1 | 15% | 29% | 22% | 71% | 569353-667503 | NA |
| | Genômica 2 | 66% | 66% | 16% | 66% | 694948-735864 | NA |
| | Patogenicidade 2 | 28% | 28% | 64% | 68% | 939710-987780 | Normal |
| | Patogenicidade 3 | 30% | 0% | 76% | 38% | 1106326-1118347 | Normal |

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore |
|-------------|------------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_08 | Genômica 3 | 27% | 36% | 0% | 45% | 1128892-1164378 | NA |
| | Genômica 4 | 30% | 0% | 40% | 30% | 1282177-1292670 | NA |
| | Genômica 5 | 3% | 69% | 27% | 81% | 1368910-1453486 | NA |
| | Genômica 6 | 13% | 71% | 36% | 86% | 1545086-1602299 | NA |
| | Genômica 7 | 0% | 73% | 33% | 60% | 1627053-1638691 | NA |
| | Patogenicidade 4 | 9% | 17% | 57% | 55% | 1652611-1843228 | Normal |
| | Genômica 8 | 20% | 20% | 23% | 66% | 1848454-1901176 | NA |
| | Patogenicidade 5 | 8% | 11% | 61% | 58% | 1906774-1954836 | Normal |
| | Genômica 9 | 43% | 87% | 14% | 82% | 2110992-2183661 | NA |
| | Genômica 10 | 14% | 42% | 39% | 60% | 2198527-2220011 | NA |
| | Genômica 11 | 50% | 66% | 16% | 66% | 2319833-2325485 | NA |
| | Genômica 12 | 7% | 51% | 35% | 80% | 2363091-2468550 | NA |
| | Genômica 13 | 28% | 42% | 14% | 28% | 2482022-2493060 | NA |
| | Genômica 14 | 5% | 29% | 11% | 52% | 2532436-2555967 | NA |
| | Patogenicidade 6 | 11% | 7% | 76% | 80% | 2607886-2641064 | Normal |
| Genômica 15 | 20% | 17% | 37% | 51% | 2726186-2758174 | NA | |

Continua

| Estirpe | Ilha genômica | Desvio de GC | Continuação | | | Posição | Escore |
|---------------|------------------|---------------|-------------------------|-----------------------|-----------------------|-----------------|-------------|
| | | | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_08 | Genômica 16 | 9% | 77% | 22% | 81% | 2847714-2876401 | NA |
| | Genômica 17 | 15% | 33% | 18% | 63% | 2913369-2964353 | NA |
| | Genômica 18 | 55% | 72% | 38% | 72% | 2971199-2991361 | NA |
| | Genômica 19 | 29% | 26% | 32% | 67% | 3023519-3066619 | NA |
| | Genômica 20 | 21% | 12% | 30% | 51% | 3094267-3134645 | NA |
| | Patogenicidade 7 | 13% | 0% | 79% | 93% | 3196181-3234263 | Normal |
| | Patogenicidade 8 | 26% | 41% | 47% | 78% | 3237548-3371673 | Normal |
| | Patogenicidade 9 | 25% | 35% | 57% | 85% | 3403129-3461159 | Normal |
| | LGMAI_St_11 | Resistência 1 | 14% | 0% | 42% | 71% | 47614-57258 |
| Genômica 1 | | 15% | 28% | 7% | 69% | 116993-145639 | NA |
| Genômica 2 | | 69% | 61% | 7% | 69% | 232878-251528 | NA |
| Resistência 2 | | 15% | 0% | 30% | 46% | 476730-496378 | Fraco |
| Resistência 3 | | 4% | 23% | 19% | 52% | 602358-622312 | Fraco |
| Genômica 3 | | 30% | 0% | 7% | 38% | 773514-785536 | NA |
| Genômica 4 | | 37% | 71% | 5% | 79% | 825557-879835 | NA |
| Genômica 5 | | 0% | 26% | 3% | 80% | 885139-917748 | NA |

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore |
|-------------|----------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_11 | Resistência 4 | 31% | 68% | 21% | 47% | 1135404-1153950 | Normal |
| | Genômica 6 | 0% | 90% | 0% | 72% | 1167870-1179508 | NA |
| | Resistência 5 | 21% | 27% | 14% | 63% | 1204262-1284068 | Normal |
| | Genômica 7 | 34% | 63% | 5% | 79% | 1420953-1470142 | NA |
| | Genômica 8 | 32% | 35% | 8% | 58% | 1627670-1658358 | NA |
| | Resistência 6 | 33% | 100% | 20% | 53% | 1760686-1785548 | Forte |
| | Resistência 7 | 17% | 29% | 11% | 41% | 2109442-2135980 | Fraco |
| | Genômica 9 | 25% | 12% | 0% | 37% | 2141086-2149066 | NA |
| | Genômica 10 | 12% | 68% | 4% | 84% | 2153807-2180303 | NA |
| | Genômica 11 | 83% | 66% | 0% | 66% | 2235148-2240800 | NA |
| | Genômica 12 | 7% | 35% | 0% | 57% | 2266163-2286944 | NA |
| | Genômica 13 | 61% | 23% | 0% | 76% | 2295771-2305125 | NA |
| | Genômica 14 | 34% | 46% | 2% | 85% | 2414896-2447492 | NA |
| | Genômica 15 | 20% | 37% | 0% | 85% | 2510608-2534984 | NA |
| | Genômica 16 | 21% | 34% | 9% | 62% | 2555887-2715279 | NA |
| | Patogenicidade | 15% | 28% | 58% | 69% | 116993-145639 | Normal |

1

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore |
|-------------------|------------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_11 | Genômica 1 | 69% | 61% | 38% | 69% | 232878-251528 | NA |
| | Patogenicidade 2 | 15% | 0% | 53% | 46% | 476730-496378 | Fraco |
| | Patogenicidade 3 | 4% | 23% | 42% | 52% | 602358-622312 | Fraco |
| | Patogenicidade 4 | 30% | 0% | 76% | 38% | 773514-785536 | Normal |
| | Genômica 2 | 37% | 71% | 30% | 79% | 825557-879835 | NA |
| | Genômica 3 | 0% | 26% | 30% | 80% | 885139-917748 | NA |
| | Patogenicidade 5 | 31% | 68% | 63% | 47% | 1135404-1153950 | Normal |
| | Patogenicidade 6 | 0% | 90% | 45% | 72% | 1167870-1179508 | Normal |
| | Genômica 4 | 21% | 27% | 30% | 63% | 1204262-1284068 | NA |
| | Patogenicidade 7 | 3% | 0% | 90% | 39% | 1312847-1410032 | Fraco |
| | Genômica 5 | 34% | 63% | 36% | 79% | 1420953-1470142 | NA |
| Patogenicidade 8 | 6% | 6% | 43% | 68% | 1517897-1535327 | Fraco | |
| Genômica 6 | 32% | 35% | 35% | 58% | 1627670-1658358 | NA | |
| Patogenicidade 9 | 33% | 100% | 80% | 53% | 1760686-1785548 | Forte | |
| Patogenicidade 10 | 3% | 0% | 98% | 21% | 1876789-1945272 | Fraco | |
| Patogenicidade 11 | 4% | 0% | 100% | 21% | 2023447-2045850 | Fraco | |

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore | |
|---------------|-------------------|---------------|-------------------------|-----------------------|-----------------------|-----------------|---------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | | |
| LGMAI_St_11 | Patogenicidade 12 | 17% | 29% | 52% | 41% | 2109442-2135980 | Normal | |
| | Genômica 7 | 25% | 12% | 25% | 37% | 2141086-2149066 | NA | |
| | Genômica 8 | 12% | 68% | 24% | 84% | 2153807-2180303 | NA | |
| | Genômica 9 | 83% | 66% | 16% | 66% | 2235148-2240800 | NA | |
| | Genômica 10 | 7% | 35% | 21% | 57% | 2266163-2286944 | NA | |
| | Patogenicidade 13 | 61% | 23% | 46% | 76% | 2295771-2305125 | Normal | |
| | Patogenicidade 14 | 17% | 5% | 43% | 76% | 2319728-2367935 | Normal | |
| | Patogenicidade 15 | 34% | 46% | 74% | 85% | 2414896-2447492 | Forte | |
| | Genômica 11 | 20% | 37% | 31% | 85% | 2510608-2534984 | NA | |
| | Genômica 12 | 21% | 34% | 31% | 62% | 2555887-2715279 | NA | |
| | LGMAI_St_14 | Genômica 1 | 34% | 31% | 8% | 57% | 58676-89364 | NA |
| | | Resistência 1 | 31% | 53% | 11% | 69% | 124909-188502 | Normal |
| Genômica 2 | | 38% | 21% | 8% | 65% | 690135-739229 | NA | |
| Resistência 2 | | 14% | 0% | 42% | 71% | 825250-834894 | Normal | |
| Genômica 3 | | 25% | 25% | 2% | 70% | 1005051-1055954 | NA | |
| Genômica 4 | | 30% | 0% | 7% | 46% | 1130321-1142343 | NA | |

Continua

| Estirpe | Ilha genômica | Continuação | | | | Posição | Escore |
|------------------|---------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | Desvio de GC | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_14 | Genômica 5 | 26% | 6% | 0% | 53% | 1346098-1363866 | NA |
| | Resistência 3 | 33% | 100% | 20% | 53% | 1405406-1430268 | Forte |
| | Genômica 6 | 22% | 44% | 0% | 44% | 1595401-1604678 | NA |
| | Genômica 7 | 0% | 100% | 0% | 81% | 1616117-1627755 | NA |
| | Genômica 8 | 35% | 57% | 4% | 84% | 1785304-1859423 | NA |
| | Resistência 4 | 56% | 62% | 12% | 68% | 1878275-1895276 | Forte |
| | Genômica 9 | 13% | 73% | 0% | 91% | 2227758-2251948 | NA |
| | Genômica 10 | 61% | 0% | 0% | 76% | 2367253-2376665 | NA |
| | Genômica 11 | 100% | 50% | 0% | 66% | 2437760-2443412 | NA |
| | Resistência 5 | 77% | 33% | 11% | 55% | 2470033-2477686 | Normal |
| | Genômica 12 | 16% | 33% | 0% | 80% | 2478984-2508032 | NA |
| | Genômica 13 | 23% | 26% | 0% | 83% | 2511034-2531428 | NA |
| | Genômica 14 | 20% | 28% | 8% | 72% | 2569921-2693965 | NA |
| | Genômica 1 | 34% | 31% | 34% | 57% | 58676-89364 | NA |
| | Genômica 2 | 31% | 53% | 36% | 69% | 124909-188502 | NA |
| Patogenicidade 1 | 21% | 0% | 78% | 47% | 458303-476783 | Normal | |

Continua

| Estirpe | Ilha genômica | Desvio de GC | Continuação | | | Posição | Escore |
|-------------|-------------------|--------------|-------------------------|-----------------------|-----------------------|-----------------|--------|
| | | | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_14 | Patogenicidade 2 | 38% | 21% | 51% | 65% | 690135-739229 | Normal |
| | Genômica 3 | 25% | 25% | 35% | 70% | 1005051-1055954 | NA |
| | Patogenicidade 3 | 30% | 0% | 76% | 46% | 1130321-1142343 | Normal |
| | Genômica 4 | 26% | 6% | 26% | 53% | 1346098-1363866 | NA |
| | Patogenicidade 4 | 33% | 100% | 80% | 53% | 1405406-1430268 | Forte |
| | Patogenicidade 5 | 3% | 0% | 100% | 19% | 1493232-1559323 | Fraco |
| | Genômica 5 | 22% | 44% | 22% | 44% | 1595401-1604678 | NA |
| | Patogenicidade 6 | 0% | 100% | 45% | 81% | 1616117-1627755 | Normal |
| | Patogenicidade 7 | 8% | 14% | 79% | 35% | 1641675-1737549 | Normal |
| | Genômica 6 | 35% | 57% | 32% | 84% | 1785304-1859423 | NA |
| | Genômica 7 | 56% | 62% | 31% | 68% | 1878275-1895276 | NA |
| | Patogenicidade 8 | 3% | 0% | 73% | 70% | 1960979-1988147 | Normal |
| | Genômica 8 | 13% | 73% | 21% | 91% | 2227758-2251948 | NA |
| | Patogenicidade 9 | 4% | 0% | 100% | 21% | 2272488-2294891 | Fraco |
| | Patogenicidade 10 | 61% | 0% | 46% | 76% | 2367253-2376665 | Normal |
| | Genômica 9 | 100% | 50% | 16% | 66% | 2437760-2443412 | NA |

Continua

| Estirpe | Ilha genômica | Desvio de GC | Conclusão | | | Posição | Score |
|----------------|----------------------|---------------------|--------------------------------|------------------------------|------------------------------|-----------------|--------------|
| | | | Desvio de uso de códons | Proteínas específicas | Proteínas hipotéticas | | |
| LGMAI_St_14 | Genômica 10 | 77% | 33% | 33% | 55% | 2470033-2477686 | NA |
| | Genômica 11 | 16% | 33% | 23% | 80% | 2478984-2508032 | NA |
| | Patogenicidade 11 | 23% | 26% | 80% | 83% | 2511034-2531428 | Normal |
| | Genômica 12 | 20% | 28% | 31% | 72% | 2569921-2693965 | NA |