



Universidade do Estado do Rio de Janeiro

Centro Biomédico

Faculdade de Ciências Médicas

Stephanie da Silva Rodrigues de Souza

**Estrutura populacional e cenário evolutivo do patógeno multirresistente
Enterococcus faecium em instituições hospitalares do Rio de Janeiro**

Rio de Janeiro

2020

Stephanie da Silva Rodrigues de Souza

**Estrutura populacional e cenário evolutivo do patógeno multirresistente
Enterococcus faecium em instituições hospitalares do Rio de Janeiro**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-graduação em Microbiologia, da Universidade do Estado do Rio de Janeiro. Área de concentração: Microbiologia Médica Humana.

Orientadoras: Prof.^a Dra. Vânia Lúcia Carreira Merquior
Prof.^a Dra. Lúcia Martins Teixeira

Rio de Janeiro

2020

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CB-A

S729 Souza, Stephanie da Silva Rodrigues de.
Estrutura populacional e cenário evolutivo do patógeno multirresistente
Enterococcus faecium em instituições hospitalares do Rio de Janeiro /
Stephanie da Silva Rodrigues de Souza. – 2020.
214 f.

Orientadoras: Vânia Lúcia Carreira Merquior; Lúcia Martins Teixeira.
Tese (Doutorado) - Universidade do Estado do Rio de Janeiro, Faculdade
de Ciências Médicas. Pós-graduação em Microbiologia.

1. *Enterococcus faecium* – Genética - Teses. 2. Farmacorresistência
bacteriana – Teses. 3. Epidemiologia molecular – Teses. 4. Evolução
molecular – Teses. I. Merquior, Vânia Lúcia Carreira. II. Teixeira, Lúcia
Martins . III. Universidade do Estado do Rio de Janeiro. Instituto de Biologia
Roberto Alcântara Gomes. IV. Título.

CDU 575.111:579.86

Bibliotecária: Kalina Silva CRB7/4377

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

Assinatura

Data

Stephanie da Silva Rodrigues de Souza

Estrutura populacional e cenário evolutivo do patógeno multirresistente *Enterococcus faecium* em instituições hospitalares do Rio de Janeiro

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-graduação em Microbiologia, da Universidade do Estado do Rio de Janeiro. Área de concentração: Microbiologia Médica Humana.

Aprovada em 14 de setembro de 2020.

Orientadora: Prof.^a Dra. Lúcia Martins Teixeira
Universidade Federal do Rio de Janeiro

Banca Examinadora:

Prof.^a Dra. Vânia Lúcia Carreira Merquior (Orientadora)
Faculdade de Ciências Médicas- UERJ

Prof. Dr. Robson de Souza Leão
Faculdade de Ciências Médicas - UERJ

Prof. Dr. Felipe Piedade Gonçalves Neves
Universidade Federal Fluminense

Dra. Ana Paula D'Alincourt Carvalho Assef
Fundação Oswaldo Cruz

Prof.^a Dra. Carla Alexandra Novais de Oliveira e Silva
Universidade do Porto

Rio de Janeiro

2020

DEDICATÓRIA

Ao meu querido e amado pai, Orlando Rodrigues, onde ele estiver.

AGRADECIMENTOS

Primeiramente a Universidade do Estado do Rio de Janeiro por RESISTIR!

Aos meus familiares que sempre me apoiaram em todas as minhas decisões, mesmo aquelas que me levaram para caminhos distantes. E, principalmente, meu eterno obrigado ao meu querido e amado pai, Orlando Rodrigues, que em vida foi um dos maiores incentivadores da minha educação. Este é o resultado final das tantas horas de paciência e atenção que começaram desde a alfabetização, lendo a, ainda viva na minha memória, cartilha “O sonho de Talita”.

Ao meu companheiro de vida, Márcio Taddei, que apareceu no meio dessa jornada, mas que hoje é a peça fundamental de tudo. Obrigada por todo suporte durante esse tempo, por aguentar a distância e por se manter presente e paciente mesmo quando nada era calmaria. Muito desta tese se deve a você, até porque eu não saberia como fazer as análises matemáticas sozinha. Minha eterna gratidão.

As minhas queridas orientadoras, Prof.^a Vânia Merquior e Prof.^a Lúcia Teixeira, primeiramente, por acolherem uma jovem biomédica, ainda na flor dos seus 20 anos de idade, e permitirem que hoje, um pouco mais cheia de rugas e com menos cabelo, saia uma doutora e cientista. Durante esses quase nove anos dentro do Laboratório de Apoio Biotecnológico tive a oportunidade não apenas de conviver com duas excelentes microbiologistas, mas com dois seres humanos extraordinários. Obrigada por toda confiança depositada, os puxões de orelha, os conselhos, as conversas e ensinamentos, e os exemplos diários que me permitiram crescer como profissional e como pessoa. Hoje, muito do que eu, Stephanie, sou, é fruto desses nove anos de convívio e aprendizado diário. Minha eterna admiração pelo esforço e dedicação de vocês em continuar fazendo ciência de qualidade e acreditando na educação. O laboratório 027 será sempre minha “segunda” casa.

Ao Dr. Paul Joseph Planet por me acolher em seu laboratório no *Childrens Hospital of Philadelphia* durante o período de doutorado sanduíche e por me proporcionar uma das experiências mais transformadoras. Meu muitíssimo obrigado ao Dr. Moustafa Ahmed, pela paciência e carinho durante os seis meses de Filadélfia, mas principalmente, por não me matar depois do décimo erro na linha de comando. Sou sua fã! Obrigada também, as queridas companheiras do PlanetLab: Kelsey, Prioty and Chanelle, por tornaram os seis meses mais amáveis e com sabor de nachos.

Aos professores do Departamento de Microbiologia, Imunologia e Parasitologia da

UERJ e a todos os outros professores que fizeram parte do meu crescimento e trajetória acadêmica. Minha imensa gratidão e admiração por escolherem um ofício tão nobre.

Aos membros da banca, a Prof.^a Dra. Ana Paula D'Alincourt, Prof.^a Dra. Carla Alexandra Novais de Oliveira e Silva e o Prof. Dr. Robson de Souza Leão que se dispuseram em participar da finalização desta etapa. Meu muito obrigada.

Ao Prof. Felipe Piedade Gonçalves Neves meu obrigada mais do que especial pela revisão desse trabalho e pela atenção, solicitude e ensinamento ao longo dos anos. Você é uma grande inspiração para mim.

À minha querida companheira científica Adriana, carinhosamente apelidada de "Babu". Meu eterno obrigada por esses anos de convívio harmonioso ao seu lado, dividindo uma mesa de 2x2 cm. Obrigada pelos conselhos científicos, amorosos e de vida, por todo conhecimento transmitido, pelas discussões sobre evolução e bactérias, por aturar minhas doideiras e piadas e por sempre me acompanhar no pastel do fim de tarde. Sua amizade foi um dos melhores presentes que esta jornada me trouxe.

À querida companheira de jornada Andréa, minha eterna gratidão por sua disposição em se fazer presente, ensinar, apoiar, ouvir e chorar as mágoas juntas. Sentirei saudade dos almoços que eram quase uma sessão terapêutica. Você é uma pessoa maravilhosa.

As outras amigas que o laboratório 027 me trouxe, Carolina Gomes, Priscilla Ribau e Lucina Milanez. Obrigada por todo o carinho, as risadas, o apoio e os quilos a mais que os nossos encontros sempre me trazem. Em especial à Beatriz Monteiro, que eu sei que ficaria chateada em ser meramente citada no meio de outros nomes, meu eterno agradecimento por se dispor a ensinar uma jovem IC nove anos atrás. Obrigada pelas conversas, conselhos e ensinamentos.

Aos companheiros do Laboratório 027, tanto os que hoje integram o grupo, quanto os muitos que passaram por lá ao longo dos últimos nove anos. Meu muito obrigada pelo convívio diário, pelos momentos de descontração e ajuda. A presença de todos, das mais variadas formas, em diferentes intensidades, com certeza, foi fundamental para essa jornada.

Aos amigos da vida, que vieram da escola ou da faculdade, amigos de longa data ou os mais recentes, mas que sempre me apoiam e se fazem presente.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) pelo suporte financeiro para a execução do projeto. O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Lança o barco contra o mar
Venha o vento que houver
E se virar,
Nada!
Pega a mala que couber
Roda a estrada sem saber
E se perder,
Calma! ...
... E se puder,
Voa!

Rubel

RESUMO

SOUZA, Stephanie da Silva Rodrigues. **Estrutura populacional e cenário evolutivo do patógeno multirresistente *Enterococcus faecium* em instituições hospitalares do Rio de Janeiro.** 2020. 214f. Tese (Doutorado em Microbiologia) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Enterococcus faecium se destaca como um patógeno oportunista frequentemente associado a infecções hospitalares em todo o mundo. O objetivo deste estudo foi determinar a estrutura populacional de amostras de *E. faecium* isoladas em instituições hospitalares do estado do Rio de Janeiro, durante um período consecutivo de oito anos (2009 a 2016), e avaliar os eventos genéticos relacionados à evolução de linhagens adaptadas ao ambiente hospitalar. As amostras bacterianas foram inicialmente caracterizadas quanto à suscetibilidade aos antimicrobianos, avaliada pelo método de disco-difusão, e à presença de determinantes genéticos associados à resistência a antimicrobianos e a virulência, investigada por técnicas de PCR. Com base nos resultados obtidos, 52 amostras foram selecionadas para realização de sequenciamento do genoma completo (WGS) com o auxílio da plataforma Illumina HiSeq. Além dessas, outras 22 amostras pertencentes à linhagem hospitalar ST78, anteriormente circulante no estado, foram incluídas para fins comparativos. Um extenso *pipeline* de ferramentas de bioinformática foi empregado para reconstruir as relações filogenéticas e investigar diferentes aspectos genéticos do processo evolutivo das 74 amostras analisadas. As análises filogenéticas indicaram a ocorrência de múltiplas alterações populacionais ao longo do tempo. A linhagem ST78 foi predominante no período de 2002 a 2008, sendo posteriormente substituída pela ST412. Essa linhagem predominou até 2012, quando então, duas novas surgiram: ST963 em 2013 e ST896 em 2015. As análises do genoma revelaram diferentes mecanismos genéticos associados à evolução dessas linhagens. A linhagem ST78 teve seu surgimento e expansão associados ao aumento do genoma acessório, com um elevado número de sequências de inserção (IS) e profagos. Além disso, a presença exclusiva de um sistema toxina-antitoxina, envolvido na estabilização de elementos genéticos móveis (EGM), corroborou os resultados encontrados. Por outro lado, o surgimento da linhagem ST412, observada em duas ocasiões distintas, foi associado a eventos de recombinação nos nodos ancestrais. Ademais, um número elevado de SNPs, possivelmente relacionado à formação de hipermutantes devido ao ganho de um códon de parada em um sistema de reparo de DNA, e o ganho exclusivo de genes de biossíntese da parede celular, também foram associados à expansão desta linhagem. Eventos de recombinação foram detectados nos nodos ancestrais das linhagens ST963 e ST896. O ST896 teve o menor número de genes acessórios, entretanto esteve fortemente associado a uma maior presença de genes de resistência aos antimicrobianos. A expansão da ST963 foi acompanhada, particularmente, do ganho exclusivo de genes relacionados ao metabolismo de aminoácidos e de metabólitos secundários. Um número elevado de unidades de novidade genética (GNU) foi observado nesse grupo. Algumas características, tais como a presença do gene *dfrF* (resistência ao trimetoprim), do determinante de virulência *hyl*, da *ISLgar5* e o alelo 44 do gene *purK*, também foram associadas às populações que substituíram a ST78. No geral, nossos resultados indicam que a população de *E. faecium* investigada sofreu sucessivas substituições de STs, associadas a um processo evolutivo dinâmico. Os dados evidenciam os diferentes eventos genéticos que podem ter desempenhado papel importante na emergência de subpopulações especializadas, e reafirmam a necessidade de constante monitoramento das infecções causadas por esse microrganismo.

Keywords: *Enterococcus faecium*. Genômica comparativa. Resistência aos antimicrobianos. Estrutura populacional. Evolução bacteriana. Sequenciamento do Genoma Completo.

ABSTRACT

SOUZA, Stephanie da Silva Rodrigues. **Populational structure and evolutionary scenario of the multidrug-resistant pathogen *Enterococcus faecium* in hospital institutions in Rio de Janeiro.** 2020. 214f. Tese (Doutorado em Microbiologia) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2020.

Enterococcus faecium stands out as an opportunistic pathogen frequently associated with hospital-acquired infections around the world. The aim of this study was to determine the population structure among *E. faecium* isolates recovered in hospitals of Rio de Janeiro city, Brazil, over a period of eight consecutive years (2009 to 2016) and to evaluate the genetic events related to the evolution of hospital-adapted clonal lineages. The bacterial isolates were initially characterized according to their susceptibilities to antimicrobial agents, evaluated by the disk-diffusion method, and to the presence of genetic determinants associated with resistance to antimicrobial agents and with virulence, investigated by PCR assays. Based on the results obtained, 52 strains were subsequently selected for whole-genome sequencing (WGS) using the Illumina HiSeq platform. In addition, another 22 strains belonging to the hospital-adapted lineage ST78, previously predominant in the city hospital institutions, were included for comparative purposes. An extensive pipeline of bioinformatics tools was implemented to reconstruct phylogenetic relationships and analyze different genetic aspects of the evolution of the 74 strains studied. Phylogenetic analyses indicated the occurrence of multiple lineage changes over time. ST78 was predominant among isolates recovered from 2002 to 2008, being replaced by the emergence of ST412. This lineage was predominant until 2012, when two new lineages have emerged: ST963 in 2013, and ST896 in 2015. Genome analyses revealed different mechanisms associated with the evolution of these lineages. ST78 had its emergence and expansion associated with the enlargement of the accessory genome, with a high number of insertion sequences (ISs) and prophages. Moreover, the exclusive presence of a toxin-antitoxin system, involved in the stabilization of mobile genetic elements (MGE), corroborated these findings. On the other side, the emergence of ST412, observed in two distinct waves, was associated with several recombination events in the ancestor nodes. Additionally, a high number of SNPs, possibly related to the formation of hypermutantes due to a stop codon acquisition in a DNA repair system, and the exclusive acquisition of cell wall biogenesis genes, were associated with the expansion of this lineage. Recombination events in the ancestor nodes were detected in the ST963 and ST896 lineages. ST896 had the lowest number of accessory genes, but a frequent presence of antibiotic resistance was strongly associated with this group. The expansion of ST963 was accompanied by the exclusive acquisition of genes related to amino acid metabolism and biosynthesis of secondary metabolites. A large number of genetic novelty units (GNU) was observed in this group. Some features such as the presence of *drfF* gene (trimethoprim resistance), virulence determinant *hyl*, the *ISLgar5* and the allele 44 to *purK* gene were also related to the lineages that replaced ST78. Overall, our results indicate that the *E. faecium* population investigated has undergone successive lineage replacement in a dynamic evolutionary process. The data highlighted the different genetic events that may play a role in the emergency of specialized subpopulations adapted to the hospital environment and endorse the need of continuous monitoring the infections caused by this microorganism.

Keywords: *Enterococcus faecium*. Comparative genomic. Antimicrobial Resistance. Population structure. Bacterial evolution. Whole-genome sequencing.

LISTA DE FIGURAS

Figura 1 –	Fluxo de trabalho realizado para a análise do genoma de 74 amostras de <i>Enterococcus faecium</i>	59
Figura 2 –	Árvore filogenética baseada em SNP construída através de Máxima Verossimilhança e enraizamento pelo ponto médio (<i>midpoint rooting</i>), dos 74 genomas de <i>Enterococcus faecium</i> estudados.....	76
Figura 3 –	Mapa de calor com arranjo hierárquico em cluster através do algoritmo UPGMA indicando os valores de ANI entre a amostra CL7611 e outras nove amostras pertencentes ao ST412.....	77
Figura 4 –	Mapa de calor com arranjo hierárquico em cluster através do algoritmo UPGMA indicando os valores de ANI entre a amostra CL7611 e outras nove amostras pertencentes ao ST78.....	78
Figura 5 –	Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs e enraizamento pelo ponto médio (<i>midpoint rooting</i>), demonstrando diferenças no conteúdo genético de 74 genomas de <i>Enterococcus faecium</i>	82
Figura 6 –	Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs e enraizamento pelo ponto médio (<i>midpoint rooting</i>) demonstrando o mobiloma associado ao genoma de <i>Enterococcus faecium</i>	84
Figura 7 –	Alinhamento para comparação das sequências nucleotídicas do gene <i>ant(6')-Ia</i> , e respectivas sequências de aminoácidos, entre a referência (AF330699.1) e uma das amostras de <i>Enterococcus faecium</i> apresentando resultados incongruentes para a susceptibilidade a estreptomicina, identificada no presente estudo.....	88
Figura 8 –	Modelo estrutural 3D da proteína ANT(6')-Ia nativa e mutada e respectivos sítios de ligação previstos.....	90
Figura 9 –	Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs com enraizamento pelo ponto médio (<i>midpoint rooting</i>) demonstrando a presença/ausência das três subunidades que compõem o sistema R-M.....	92

Figura 10 –	Distribuição dos genes preditos como exclusivos para cada um dos grupos de STs de <i>Enterococcus faecium</i> estudados de acordo com a sua categoria de função COG (<i>Cluster of Orthologous Group</i>).....	95
Figura 11 –	Distribuição dos genes preditos como exclusivos para cada um dos grupos pertencentes às duas emergências do ST412 de acordo com a sua categoria de função COG (<i>Cluster of Orthologous Group</i>).....	96
Figura 12 –	Avaliação das proteínas anotadas a partir da amostra de referência Aus0004 que tiveram o ganho prematuro de códons de parada nos 74 genomas de <i>Enterococcus faecium</i>	100
Figura 13 –	Avaliação das proteínas anotadas a partir da amostra de referência Aus0004 que tiveram a perda prematura de códons de parada nos 74 genomas de <i>Enterococcus faecium</i>	101
Figura 14 –	Distribuição das alterações observadas no gene que codifica a proteína ligadora de penicilina 5 (<i>pbp5</i>) em onze perfis distintos de amostras de <i>Enterococcus faecium</i>	105
Figura 15 –	Número de Unidades de Novidade Genética (GNU) nos 74 genomas de <i>Enterococcus faecium</i> estudados.....	107
Figura 16 –	Diagrama de Venn representando o compartilhamento de Unidades de Novidade Genética (GNU) entre os cinco grupos de STs estudados.....	108
Figura 17 –	Avaliação dos impactos de eventos de recombinação homóloga na reconstrução da filogenia de <i>Enterococcus faecium</i>	113
Figura 18 –	Árvore de filogenia construída através de Máxima Verossimilhança baseada no alinhamento de SNPs e enraizamento pelo ponto médio (<i>midpoint rooting</i>) demonstrando os eventos de recombinação homóloga em amostras de <i>Enterococcus faecium</i>	114
Figura 19 –	Árvore de filogenia construída através de Máxima Verossimilhança pelo software ClonalFrameML e enraizamento pelo ponto médio (<i>midpoint rooting</i>) detalhando as regiões de recombinação inferidas para os ramos ancestrais e terminais dos 74 genomas de <i>Enterococcus faecium</i>	116
Figura 20 –	Árvore de filogenia reconstruída através de análise bayesiana estimando o tempo de divergência entre as diferentes linhagens de <i>Enterococcus faecium</i> isolados no período de 2004 a 2016.....	119

Figura 21 – Linha do tempo reconstruindo os eventos de emergência de linhagens de <i>Enterococcus faecium</i> em instituições hospitalares do estado do Rio de Janeiro.....	120
---	-----

LISTA DE GRÁFICOS

Gráfico 1 –	Distribuição de 205 amostras de <i>Enterococcus faecium</i> de acordo com a suscetibilidade à 18 antimicrobianos.....	63
Gráfico 2 –	Correlação entre as taxas de não suscetibilidade a 14 antimicrobianos diferentes entre amostras não suscetíveis (VRE) e suscetíveis (VSE) à vancomicina.....	66
Gráfico 3 –	Distribuição dos fenótipos de não suscetibilidade a níveis elevados de aminoglicosídeos em amostras de <i>Enterococcus faecium</i>	67
Gráfico 4 –	Frequência de determinantes genéticos de resistência a níveis elevados de gentamicina identificados por metodologia de reação em cadeia da polimerase em amostras de <i>Enterococcus faecium</i>	69
Gráfico 5 –	Distribuição dos determinantes genéticos de virulência identificados por reação em cadeia da polimerase em amostras de <i>Enterococcus faecium</i> isoladas no período de 2009 a 2016.....	70
Gráfico 6 –	Correlação entre a presença de determinantes genéticos de virulência e a resistência à vancomicina em amostras de <i>Enterococcus faecium</i>	71
Gráfico 7 –	Composição do genoma central e do genoma acessório dos cinco grupos de STs de <i>Enterococcus faecium</i> observados no estudo.....	93

LISTA DE QUADROS

Quadro 1 –	Oligonucleotídeos iniciadores utilizados nas reações de PCR <i>multiplex</i> para detecção de determinantes de virulência e resistência em <i>Enterococcus faecium</i>	48
Quadro 2 –	Principais características dos diferentes STs encontrados nos 74 genomas de <i>Enterococcus faecium</i> circulantes em instituições hospitalares do estado Rio de Janeiro.....	74
Quadro 3 -	Alterações observadas nas duas subunidades da DNA girase (<i>gyrA</i> e <i>gyrB</i>) e da topoisomerase tipo IV (<i>parC</i> e <i>parE</i>) e sua distribuição nos 74 genomas de <i>Enterococcus faecium</i> avaliados.....	104
Quadro 4 –	Lista com os dez genomas presentes no banco do programa WhatsGNU que obtiveram mais <i>hits</i> com amostras de <i>Enterococcus faecium</i> dos diferentes grupos de STs estudados	110

LISTA DE TABELAS

Tabela 1 –	Distribuição das 205 amostras de <i>Enterococcus faecium</i> incluídas no presente estudo, de acordo com o espécime clínico de origem e ano de isolamento.....	62
Tabela 2 –	Distribuição do percentual de não suscetibilidade aos antimicrobianos de amostras de <i>Enterococcus faecium</i> no período de 2009 a 2016.....	65
Tabela 3 –	Perfis de suscetibilidade às quinolonas de 205 amostras de <i>Enterococcus faecium</i>	68
Tabela 4 –	Listagem das 32 sequências de inserção detectadas no genoma das 74 amostras de <i>Enterococcus faecium</i> estudadas e sua respectiva distribuição entre cinco grupos de STs distintos.....	86
Tabela 5 -	Valores obtidos para os diferentes eventos de mutação e seus efeitos nos cinco grupos de STs estudados.....	98
Tabela 6 –	Relação do número de eventos de recombinação e sua extensão nos diferentes nodos ancestrais da árvore de filogenia de <i>Enterococcus faecium</i>	112
Tabela 7 –	Parâmetros estimados na determinação do valor de D de Tajima para cinco genes de <i>Enterococcus faecium</i> avaliados.....	117
Tabela 8 –	Métricas obtidas para diversos parâmetros estudados no genoma de <i>Enterococcus faecium</i> e suas correlações entre os diferentes grupos de STs predominantes.....	121

LISTA DE ABREVIATURAS E SIGLAS

A	Adenina
AA	Aminoácidos
AAC	Acetiltransferases
Acm	Adesina de colágeno
Ala	Alanina
ANI	<i>Average nucleotide identity</i>
ANT	Nucleotidiltransferases
APH	Fosfotransferases
AREfm	<i>Ampicillin-resistant Enterococcus faecium</i>
Arg	Arginina
AS	Agar sangue
Asn	Asparagina
Asp	Ácido aspártico
ATCC	<i>American Type Culture Collection</i>
ATP _{SEP}	Adenosina trifosfato
BAPS	<i>Bayesian analysis of the population structure</i>
BLAST	<i>Basic local alignment search tool</i>
BR	Brasil
BURST	<i>Based upon related sequence types</i>
BWA	<i>Burrows-wheeler aligner</i>
C	Citosina
CA	Califórnia
CARD	<i>Comprehensive antibiotic resistance database</i>
CC	Complexo clonal
CDC	<i>Centers for Disease Control and Prevention</i>
CDS	<i>Coding sequence</i>
CGE	<i>Center for Genomic Epidemiology</i>
CIM	Concentração inibitória mínima
CLSI	<i>Clinical and Laboratory Standards Institute</i>
COG	<i>Cluster of orthologous group</i>

CRISPR-Cas	<i>Clustered regularly interspaced short palindromic repeats and associated cas protein</i>
D-Ala	D-alanina
D-Ala-D-Ala	D-alanil-D-alanina
D-Lac	D-lactato
DLV	<i>Double locus variant</i>
dNTP	Desoxirribonucleotídeo fosfatado
DNA	Ácido desoxirribonucleico
D-Ser	D-serina
DV	Desvio padrão
EB	Enzima bifuncional
EcbA	MSCRAMM ligadora de colágeno
EDTA	Ácido etilenodiaminotetracético
EGMs	Elementos genéticos móveis
EMAs	Enzimas modificadoras de aminoglicosídeos
ESKAPE	<i>Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter</i>
Esp	Proteína de superfície de enterococos
ESS	<i>Effective sample size</i>
EUA	Estados Unidos da América
EUCAST	<i>European Committee on Antimicrobial Susceptibility Testing</i>
ffn	<i>Protein coding portions of the genome segments</i>
G	Guanina
gbk	<i>GenBank</i>
Glu	Glutamato
Gln	Glutamina
Gly	Glicina
GNU	<i>Gene novelty unity</i>
GTR+G	<i>Generalised time-reversible + Gamma</i>
HCCA	Ácido α -ciano-4-hidroxicinâmico
His	<i>Histidina</i>
HKY	<i>Hasegawa-Kishino-Yano</i>
HLAR	<i>High-level aminoglycoside resistance</i>

HLGR	<i>High-level gentamicin resistance</i>
HLGSR	<i>High-level gentamicin and streptomycin resistance</i>
HLSR	<i>High-level streptomycin resistance</i>
HPD	<i>Highest posterior density</i>
Ile	Isoleucina
IRAS	Infecções relacionadas à assistência à saúde
IS	<i>Insertion sequence</i>
LAP	L-leucina- β -naftilamida
Leu	Leucina
LPSN	<i>List of procaryotic names with standing in nomenclature</i>
LPXTG	Leu-Pro-X-Thr-Gly
LS _A P	<i>Lincosamides, streptogramins A and pleuromutilins</i>
LTA	<i>Lipoteichoic acids</i>
Lys	Lisina
MA	Massachusetts
MALDI-TOF MS	<i>Matrix associated laser desorption-ionization – Time of flight</i>
MCC	<i>Maximum clade credibility</i>
MCMC	<i>Markov chain Monte Carlo</i>
MDR	<i>Multidrug-resistant</i>
Met	Metionina
MLST	<i>Multilocus sequence typing</i>
MO	Missouri
MRCA	<i>Most recent common ancestor</i>
MSCRAMMs	<i>Microbial surface components recognizing adhesive matrix molecules</i>
NCBI	<i>National Center for Biotechnology Information</i>
NJ	Nova Jersey
NT	Nucleotídeo
ORF	<i>Open reading frame</i>
PA	Pensilvânia
pb	Pares de bases
PBP	<i>Penicillin-binding proteins</i>
PCR	<i>Polimerase chain reaction</i>
PDR	<i>Pandrug-resistant</i>

PFGE	<i>Pulsed-field gel electrophoresis</i>
PHASTER	<i>Phage search tool enhanced release</i>
Phe	Fenilalanina
Pro	Prolina
PTS	<i>Phosphotransferase system</i>
PYR	L-pirrolidonil- β -naftilamida
Q/D	Quinupristina/Dalfopristina
QRDR	<i>Quinolone resistance-determining regions</i>
RJ	Rio de Janeiro
R-M	Restrição-Modificação
RNA	Ácido ribonucleico
rRNA	Ácido ribonucleico ribossomal
rRNA 16S	Ácido ribonucleico ribossomal 16S
Rrna 23S	Ácido ribonucleico ribossomal 23S
Scm	Segunda adesina de colágeno
Seg	Segundos
Ser	Serina
SgrA	Adesina de superfície de ligação ao colágeno
SLV	<i>Single locus variant</i>
SNP	<i>Single nucleotide polymorphism</i>
SP	São Paulo
ST	<i>Sequence type</i>
T	Timina
TA	Toxina-Antitoxina
TBE	Tampão composto de Tris, ácido bórico e EDTA
THG	Transferência horizontal de genes
TGI	Trato gastrointestinal
Thr	Treonina
Tn	Transposon
Trp	Triptofano
tRNA	Ácido ribonucleico transportador
TSB	<i>Trypticase soy broth</i>
Tyr	Tirosina

U	Unidade
UPGMA	<i>Unweighted-pair group method of averages</i>
v	Versão
Val	Valina
VCF	<i>Variant calling format</i>
VFDB	<i>Virulence factor database</i>
VRE	<i>Vancomycin-resistant Enterococci</i>
VREfm	<i>Vancomycin-resistant Enterococcus faecium</i>
VSE	<i>Vancomycin-susceptible Enterococci</i>
WGS	<i>Whole genome sequencing</i>
WTA	<i>Wall teichoic acids</i>

LISTA DE SÍMBOLOS

°C	grau Celsius
NaCl	Cloreto de Sódio
%	Porcentagem
β	Beta
=	Igual a
-	Negativo
h	Hora
\pm	Mais ou menos
μ l	Microlitro
α	Alfa
m/z	massa/carga
μ g	Micrograma
mL	Mililitro
μ l	Microlitro
s	segundos
Min	Minutos
rpm	Rotações por Minuto
mM	Milimolar
MgCl ₂	Cloreto de Magnésio
pmol	picomolar
>	Maior que
<	Menor que
Mb	Megabase
+	Positivo
\bar{x}	Média
\geq	Maior ou igual
\leq	Menor ou igual
Θ	Teta
\$	Cifrão
kb	Kilobase

SUMÁRIO

	INTRODUÇÃO	23
1	OBJETIVOS	40
1.1	Geral	40
1.2	Específicos	40
2	MATERIAL E MÉTODOS	41
2.1	Amostras Bacterianas	41
2.2	Caracterização Molecular por Espectrometria de Massa – MALDI TOF MS	42
2.3	Determinação da Suscetibilidade aos Antimicrobianos	42
2.4	Determinação da Concentração Inibitória Mínima (CIM) de Estreptomicina	43
2.5	Detecção dos Genes Associados à Resistência a Antimicrobianos e à Virulência	44
2.5.1	<u>Obtenção do DNA Bacteriano</u>	44
2.5.2	<u>Detecção dos Genótipos Associados à Resistência a Níveis Elevados de Aminoglicosídeos</u>	45
2.5.3	<u>Detecção dos Genótipos Associados à Resistência aos Glicopeptídeos</u>	46
2.5.4	<u>Detecção dos Genótipos Associados à Expressão de Fatores de Virulência</u>	47
2.5.5	<u>Eletroforese e Análise dos Produtos Amplificados</u>	47
2.6	Sequenciamento do Genoma de <i>E. faecium</i>	49
2.7	Análises de Bioinformática Aplicadas ao Estudo do Genoma de <i>E. faecium</i>	50
2.7.1	<u>Avaliação de Qualidade e “Trimagem” das Reads</u>	50
2.7.2	<u>Estudo do Conteúdo Genético e Análise do Pangenoma</u>	51
2.7.3	<u>Estudo das Mutações</u>	53
2.7.4	<u>Estudo dos Eventos de Recombinação</u>	56
2.7.5	<u>Análises Filogenômicas</u>	57
2.7.6	<u>Evolução Molecular</u>	58
2.7.7	<u>Estudo das Incongruências entre Fenótipo e Genótipo</u>	60
2.8	Análises Estatísticas	60
3	RESULTADOS	62

3.1	Caracterização da Amostragem.....	62
3.2	Suscetibilidade aos Antimicrobianos.....	63
3.3	Caracterização dos Genótipos Associados à Resistência.....	68
3.4	Caracterização dos Genótipos Associados à Expressão de Virulência.....	70
3.5	Análise do Sequenciamento do Genoma Completo de <i>E. faecium</i>.....	72
3.6	Determinação da Estrutura Populacional e Diversidade Genética.....	72
3.7	Reconstrução Filogenética.....	75
3.8	Estudo do Conteúdo Genético.....	79
3.9	Incongruência Fenotípica e Genotípica.....	87
3.10	Análise dos Mecanismos de Defesa.....	91
3.11	Análise do Pangenoma.....	92
3.12	Estudo das Mutações.....	97
3.13	Análise das Mutações Associadas à Resistência às Quinolonas e β- lactâmicos.....	102
3.14	Avaliação das Unidades de Novidade Genética.....	106
3.15	Estudo dos Eventos de Recombinação Homóloga.....	111
3.16	Evolução Molecular.....	117
3.17	Estimativa do Tempo de Divergência Clonal.....	118
4	DISCUSSÃO	122
	CONCLUSÃO.....	168
	REFERÊNCIAS	170
	APÊNDICE – Dados e Métricas do Genoma de 74 Amostras de <i>Enterococcus faecium</i>.....	211

INTRODUÇÃO

Enterococcus – Considerações Gerais e Importância

Os integrantes do gênero *Enterococcus* são classificados de acordo com suas características morfológicas como cocos Gram-positivos, dispostos isolados, arranjados aos pares ou em cadeias curtas. São anaeróbicos facultativos, crescem na faixa de temperatura de 10°C a 45°C, sendo a temperatura ótima de crescimento na faixa de 35°C a 37°C; a maior parte das amostras apresenta o ácido teicóico associado à parede celular (TEIXEIRA et al., 2019). Suas características biológicas permitem seu crescimento em meio contendo cloreto de sódio (NaCl) a 6,5%, e capacidade de hidrolisar esculina na presença de sais biliares. A maioria das espécies produz as enzimas pirrolidonicarboxilamidase e leucina aminopeptidase que hidrolisam os substratos L-pirrolidonicarboxilamida (PYR) e L-leucina- β -naftilamida (LAP), respectivamente. São homofermentativos e o ácido láctico é o produto final do metabolismo fermentativo da glicose ou de outros carboidratos. Não possuem a enzima citocromo oxidase, sendo negativos no teste da catalase; porém, podem produzir uma pseudocatalase (FACKLAM; CARVALHO; TEIXEIRA, 2002).

Durante muito tempo esses microrganismos fizeram parte do gênero *Streptococcus*, sendo também classificados como pertencentes ao grupo D de Lancefield (LANCEFIELD, 1933). A partir da década de 1980, com o advento das técnicas moleculares, o esquema de classificação dos enterococos foi reavaliado. Schleifer e Kilpper-Bälz (1984), conjugando técnicas de hibridização de DNA-DNA, observaram que as espécies *Streptococcus faecalis* e *Streptococcus faecium*, principais representantes do grupo dos enterococos, apresentavam pouca correlação genética com os demais integrantes do gênero *Streptococcus*. Assim, foi criado o gênero *Enterococcus*, e essas espécies foram renomeadas como *Enterococcus faecalis* e *Enterococcus faecium*.

Desde então, as ferramentas moleculares têm contribuído para melhor delinear o arranjo filogenético do gênero que, até o momento, ainda não está completamente elucidado. Neste ano de 2020, o gênero *Enterococcus* encontra-se composto por 68 espécies e duas subespécies, segundo o LPSN - *List of Prokaryotic Names with Standing in Nomenclature* (PARTE, 2018).

A identificação das espécies que no gênero é tradicionalmente baseada no uso de um esquema composto por diversos testes fisiológicos, como por exemplo: a produção de ácidos a partir do metabolismo de carboidratos (arabinose, manitol, metil-β-D-glicopiranosídeo, rafinose, sacarose, sorbitol e sorbose), a descarboxilação do aminoácido arginina, a utilização do piruvato de sódio, o crescimento em meio contendo 0,04 % de telurito de potássio, a produção de pigmento e a motilidade (TEIXEIRA et al., 2019). Em geral, estes testes permitem apenas distinguir as espécies mais comuns. O emprego de técnicas moleculares, principalmente as metodologias baseadas em PCR *multiplex* e, mais recentemente, das técnicas de espectrometria de massa por ionização e dessorção a laser assistida por matriz: tempo de voo (MALDI-TOF MS - do inglês, *matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*), contribuiu significativamente para a identificação das diversas espécies de *Enterococcus* spp., especialmente para a caracterização correta de amostras com comportamento atípico nos testes fenotípicos convencionais (GRIFFIN et al., 2012; JACKSON; FEDORKA-CRAY; BARRETT, 2004; SCHULTHESS et al., 2013).

Na atualidade, a Era Genômica, que tem garantido uma precisão e rapidez para a identificação bacteriana, particularmente quando associada ao diagnóstico clínico, tem revolucionado a Microbiologia Clínica (KWONG et al., 2015). A utilização do sequenciamento do genoma completo (WGS – do inglês, *whole-genome sequencing*), que antes parecia estar numa realidade muito distante, agora torna-se cada vez mais comum. O surgimento de diferentes plataformas de sequenciamento, incluindo os dispositivos portáteis, permitiu que os custos das técnicas diminuíssem consideravelmente, tornando-as cada vez mais acessíveis.

A utilização do sequenciamento do genoma completo permite não apenas a identificação das espécies bacterianas, mas também, com bastante propriedade, detectar genes de resistência e atributos de virulência, que muitas vezes são indicadores importantes no diagnóstico clínico, principalmente para as infecções de origem hospitalar (KÖSER et al., 2012). Além disso, essas técnicas fornecem uma resposta rápida e precisa em situações de surtos bacterianos (BROWN et al., 2019; MCGANN et al., 2016). Os enterococos se configuram como um dos grupos de microrganismos com maior número de genomas depositados em bancos de dados públicos, sendo reflexo da ampla utilização dessa metodologia em diversas investigações associadas a esse gênero bacteriano (MOUSTAFA; LAL; PLANET, 2020).

Com uma capacidade enorme de geração de dados, as técnicas de sequenciamento admitem grandes estudos de vigilância, tornando possíveis a avaliação e a caracterização de

Enterococcus spp. de diferentes origens e reforçando a ubiquidade destes microrganismos (ZAHEER et al., 2020). Os enterococos são considerados comensais do trato intestinal de seres humanos e de outros animais, como, por exemplo, cachorros, porcos, galinhas, aves selvagens, insetos e animais marinhos (COX; GILMORE, 2007; FREITAS, A.A et al., 2018; JUNG et al., 2018; MEDEIROS et al., 2017; NASER et al., 2005; POETA et al., 2005; SCHLOISSNIG et al., 2013). Na natureza, foram reportados em ambientes aquáticos, como água e sedimentos marinhos, efluentes de esgoto tratado e em diversos alimentos (BYAPPANAHALLI et al., 2012; CAMARGO et al., 2014; GOULIOURIS et al., 2019).

Representando cerca de 1% da população microbiana do intestino humano, essa é uma das principais portas de entrada para infecções enterocócicas monomicrobianas e polimicrobianas (LEBRETON; WILLEMS; GILMORE, 2014). Em situações de desestabilização da microbiota intestinal, como por exemplo decorrente da utilização de antimicrobianos de amplo espectro ou na presença de doenças de base, o risco de infecção por espécies do gênero se torna maior (ALEVIZAKOS et al., 2017; DONSKEY et al., 2000; PAPADIMITRIOU-OLIVGERIS et al., 2014). Considerados patógenos oportunistas, esses microrganismos vêm causando uma variedade de infecções nos mais diversos sítios anatômicos, que têm sido continuamente reportadas (FERNÁNDEZ-HIDALGO; ESCOLÀ-VERGÉ; PERICÀS, 2020; MADRAZO et al., 2020; RAJKUMARI; MATHUR; MISRA, 2014).

Nas últimas décadas, os enterococos, destacando-se principalmente as espécies *E. faecalis* e *E. faecium*, têm sido apontados entre os principais agentes de infecções relacionadas à assistência à saúde (IRAS) (BABADY, 2016; HIDRON et al., 2008; SCHABERG; CULVER; GAYNES, 1991). Dados obtidos de instituições brasileiras indicam que os enterococos estão entre os principais agentes de infecções do trato urinário e de corrente sanguínea (ALMEIDA; BRENDA; SILVA, 2014; BARROS; MARTINELLI; ROCHA, 2009). Um levantamento feito pelo Centro de Controle e Prevenção de Doenças dos Estados Unidos da América (CDC – do inglês, *Centers for Disease Control and Prevention*, Atlanta, Georgia, EUA) estimou que, no ano de 2019, o número de casos de infecções por *Enterococcus* spp. tenha sido de 54.500, com um total de 5.400 óbitos (CDC, 2019). No Reino Unido, a taxa de bacteremia causada por esses microrganismos foi de 13,3 para cada 100 mil habitantes em 2018, o maior índice observado nos últimos dez anos (PUBLIC HEALTH ENGLAND, 2019). Na América Latina, um estudo realizado de 1997 até 2016, pelo programa SENTRY de Vigilância de Resistência a Antimicrobianos, identificou que 23,5% das amostras bacterianas isoladas de infecções intra-abdominais pertenceram ao gênero *Enterococcus* (PFALLER et al., 2019).

Até duas décadas atrás, a espécie *E. faecalis* representava cerca de 90% das enterococcias. Porém, ao longo dos anos, diversos estudos apontaram a substituição parcial dessa espécie em decorrência do aumento do número de casos de IRAS causados pela espécie *E. faecium* (HIDRON et al., 2008; WILLEMS et al., 2012). Um estudo de vigilância global entre 2007-2008 apontou uma taxa em torno de 35% de isolamento de *E. faecium* a partir de diversos sítios (PUTNAM et al., 2010). Dados de vigilância epidemiológica mais recentes demonstraram que no Reino Unido houve um aumento na frequência de amostras de *E. faecium* isoladas de casos de bacteremia de 34% para 38% durante os cinco anos do estudo, de 2012 a 2016 (PUBLIC HEALTH ENGLAND, 2016). Em alguns países asiáticos, como a China por exemplo, *E. faecium* também foi responsável por 23,7% dos casos de bacteremia (CHEN et al., 2017). Em pacientes oncológicos, uma revisão sistemática da literatura identificou ao menos 35 surtos de *E. faecium*, com um total de 757 pacientes afetados e 77 mortes (ULRICH; VONBERG; GASTMEIER, 2017). Na Austrália, os dados obtidos no ano de 2018 para sepse indicaram que 54,2% das amostras de *Enterococcus* spp. pertenceram à espécie *E. faecalis*, enquanto o percentual de *E. faecium* foi de 39,3%. Entretanto, a taxa de mortalidade foi significativamente maior para este último, totalizando 27,2% dos casos contra 14,7% por *E. faecalis* (COOMBS et al., 2020).

A ascensão de Enterococcus faecium – Um modelo de adaptabilidade.

A ascensão de *E. faecium* no ambiente hospitalar pode ser explicada por alguns fatores, tais como: a resistência intrínseca a diversos antimicrobianos, a plasticidade genômica, permitindo a aquisição de elementos móveis (incluindo determinantes de resistência e de virulência), e a persistência e a expansão de linhagens altamente adaptadas.

Os membros da espécie *E. faecium* conseguem permanecer viáveis por longos períodos, em diferentes superfícies. Um dos primeiros estudos sobre o tema conseguiu recuperar amostras desta espécie da superfície de uma bancada após 58 dias da contaminação inicial (BONILLA; ZERVOS; KAUFFMAN, 1996). Outros estudos demonstraram a persistência de representantes dessa espécie no estofamento de cadeiras hospitalares após uma semana (NOSKIN et al., 2000). Além disso, sabe-se que *E. faecium* é capaz de sobreviver a processos de dessecação e estarvação, além de ser usualmente resistente a esterilização com radiação UV (LEBRETON et al., 2017; MARACCINI; FERGUSON; BOEHM, 2012).

Recentemente, estudos demonstraram a resistência de *E. faecium* aos processos de descontaminação com antissépticos como a clorexidina, que é bastante utilizada no ambiente hospitalar (GÚZMAN PRIETO et al., 2017). Posteriormente, outro estudo identificou o aumento, em até dez vezes, da tolerância de cepas de *E. faecium* ao álcool 70%, que é utilizado rotineiramente para a higienização das mãos (PIDOT et al., 2018).

Dentre os vários perfis intrínsecos de resistência de *E. faecium*, determinados por genes localizados no cromossomo, inclui-se a resistência aos β -lactâmicos, ao trimetoprim-sulfametoxazol e a baixas concentrações de aminoglicosídeos. Além disso, sua grande plasticidade genética permite uma alta frequência de aquisição de elementos móveis, através da transferência horizontal de genes (THG) que, quando expressos, induzem resistência a níveis elevados de β -lactâmicos, aminoglicosídeos (como gentamicina e estreptomicina), glicopeptídeos (vancomicina e teicoplanina), cloranfenicol, eritromicina, tetraciclina, clindamicina e fluoroquinolonas (por exemplo, ciprofloxacina) (LEE et al., 2019a). Todo esse processo leva ao surgimento de cepas multirresistentes.

Uma revisão sistemática publicada em 2019, com amostras isoladas de corrente sanguínea, demonstrou que 46% daquelas pertencentes a espécie *E. faecium* apresentaram resistência a níveis elevados de gentamicina; enquanto para penicilina, os percentuais obtidos foram superiores a 80% (JABBARI SHIADEH et al., 2019). No mesmo estudo, quando foram comparadas as diferentes regiões no mundo, foi observado que em países da África, a prevalência de amostras da espécie resistentes à ampicilina foi de 95%. Em outro estudo, que avaliou a frequência dessa característica em amostras isoladas em países da América Latina, como Peru e Equador, também foram encontrados percentuais elevados, sendo maiores que 85% (PANESSO et al., 2010).

Foi estimado que a caminhada para *E. faecium* se tornar um microrganismo multirresistente começou com a aquisição da resistência à tetraciclina e ao cloranfenicol durante os anos de 1950-60, seguida da resistência à gentamicina e à eritromicina na década de 1970 e à ampicilina e à vancomicina na década de 1980 (VAN TYNE; GILMORE, 2014).

Dentre os β -lactâmicos, as penicilinas possuem maior importância no tratamento das enterococcias, pois são utilizadas em associação a outros antimicrobianos, como a gentamicina, possuindo efeito bactericida (ARIAS; MURRAY, 2012). A resistência intrínseca a níveis baixos de penicilina nos enterococos está associada à produção de PBPs apresentando baixa afinidade para esses antimicrobianos, especialmente para a PBP5. A resistência a níveis elevados encontra-se mais associada às mutações pontuais e cumulativas nos sítios ativos das PBPs ou na hiperprodução de tais proteínas (RICE et al., 2004).

Outra classe de antimicrobianos com significativa importância no tratamento das enterococcias são os aminoglicosídeos que, em níveis elevados podem ser utilizados concomitantemente com os β -lactâmicos ou glicopeptídeos. Atuam fundamentalmente inibindo a síntese de proteínas ao se ligarem ao ribossomo. A resistência a níveis baixos ocorre pela redução do transporte ativo da droga para o interior da célula bacteriana, enquanto que a resistência em níveis elevados ocorre devido a existência de enzimas modificadoras de aminoglicosídeos (EMAs), além de mutações que levam a alterações na subunidade 30S do ribossomo (HERZSTEIN et al., 1984; MILLER et al., 2016; VAN TYNE; GILMORE, 2014).

As EMAs podem ser classificadas em três diferentes classes, dependendo de como atuam nas moléculas dos aminoglicosídeos: acetilação da molécula alvo a partir do acetilcoenzima A, sendo denominadas acetiltransferases (AAC); fosforilação da molécula de adenosina trifosfato (ATP), denominadas fosfotransferases (APH); ou adição de moléculas de adenina à estrutura do antimicrobiano (adenilação), a partir da molécula de ATP, denominadas nucleotidiltransferases (ANT) (CHOW, 2000; RAMIREZ; TOLMASKY, 2010).

Dentre os genótipos associados à resistência a níveis elevados de aminoglicosídeos (HLAR - do inglês, *high-level aminoglycoside resistance*) em *E. faecium*, o gene mais frequentemente identificado é o *aac(6')-Ie-aph(2')-Ia*, que codifica uma enzima bifuncional (EB), com propriedades de acetilação e fosforilação simultâneas. A presença da EB é de grande importância clínica, pois determina resistência em níveis elevados para todos os aminoglicosídeos disponíveis comercialmente, exceto a estreptomicina. Outros genes, já descritos nesses microrganismos são: *aph(2'')-Ib*, *aph(2'')-Ic*, *aph(2')-Id*, *aph(3')-IIIa*, *ant(4')-Ia*, *ant(6)-Ia*, *ant(3'')-Ia*, *ant(9')-Ia* e *ant(9')-Ib*. Tais genes são capazes de conferir resistência a diferentes antimicrobianos pertencentes à classe dos aminoglicosídeos. Entretanto, a resistência a níveis elevados de estreptomicina é codificada apenas pelos genes *ant(6)-Ia*, *ant(3'')-Ia*, *ant(9')-Ia* e *ant(9')-Ib* (CHOW, 2000; MAHBUB ALAM et al., 2005; PADMASINI; PADMARAJ; SRIVANI-RAMESH, 2014).

O surgimento de amostras apresentando resistência em níveis elevados tanto aos β -lactâmicos quanto aos aminoglicosídeos se tornou um desafio ao tratamento das infecções enterocócicas relacionadas ao ambiente hospitalar. Diante disso, antimicrobianos da classe dos glicopeptídeos começaram a ser amplamente utilizados para o tratamento das infecções graves. Porém, em sequência, foi observada a emergência de amostras de enterococos resistentes à vancomicina, que foi relatada em diferentes partes do mundo, sendo

acompanhada de significativo impacto no tratamento das enterococcias, diante da considerável redução de opções terapêuticas.

Na Europa, amostras de enterococos resistentes à vancomicina (VRE - do inglês, *vancomycin-resistant enterococci*) foram detectadas inicialmente em 1986. Acredita-se que os antimicrobianos utilizados como melhoradores de desempenho nas criações animais (como o glicopeptídeo avoparcina) tenham permitido a seleção de amostras VRE na comunidade. Nos Estados Unidos, as amostras VRE tiveram sua origem no ambiente hospitalar, estando, portanto, associadas ao uso de vancomicina nesse ambiente (ARIAS; MURRAY, 2012; JOHNSEN et al. 2005; MANCINI; BOTEN, 2005; TENOVER; MCDONALD, 2005).

A ocorrência de amostras de *E. faecium* resistentes à vancomicina no Brasil teve seu primeiro relato em 1997, em São Paulo (ZANELLA et al., 1999). Desde então, sua ocorrência em instituições hospitalares brasileiras ganhou destaque a partir dos vários relatos de surtos de VRE (CAMPOS et al., 2014; PALAZZO et al., 2011; MERQUIOR et al., 2012; ZANELLA et al., 2003). Em um estudo de Sader et al. (2016), a observação do fenótipo de resistência à vancomicina em amostras de *E. faecium* isoladas em instituições hospitalares brasileiras, no período de 2011 a 2014, foi de 71,7%.

Dados de vigilância epidemiológica demonstraram que, na América do Norte, o percentual de amostras de *E. faecium* resistentes à vancomicina, no período de 1997 até 2016, foi de 64,7% (PFALLER et al., 2019). Entretanto, em um outro estudo, foi observado que somente no período de 2011 a 2014, o percentual de amostras isoladas em hospitais norte-americanos apresentando essa característica de resistência foi em torno de 80% (WEINER et al., 2016). Na Alemanha, no período de 2012 até 2017, a proporção de amostras de VREfm (do inglês, *vancomycin-resistant Enterococcus faecium*) isoladas saltou de 11,2 % para 26,1% (MARKWART et al., 2019). Em países mais próximos do Brasil, como o Peru, já foi relatada uma taxa de 48% de amostras VRE (PANESSO et al., 2010).

A vancomicina, e a teicoplanina pertencem à classe dos glicopeptídeos e atuam ligando-se com alta afinidade ao pentapeptídeo precursor do peptideoglicano, componente básico da parede celular, especificamente na porção C-terminal do dipeptídeo D-alanil-D-alanina (D-Ala-D-Ala). Desta forma, impedem primariamente a transglicosilação dos demais precursores da cadeia de peptideoglicano e a subsequente ligação cruzada para a formação da parede celular, através de transpeptidação. Por consequência, esse componente da célula bacteriana não é formado. A resistência a esta classe ocorre devido à presença de genes que codificam enzimas para a síntese de precursores com baixa afinidade pelo antimicrobiano, nos

quais o resíduo C-terminal D-Ala é substituído por D-lactato (D-Lac) ou D-serina (D-Ser), modificando, portanto, o sítio de ligação da vancomicina (ZENG et al., 2016).

Vários genótipos de resistência aos glicopeptídeos já foram descritos em enterococos: *vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM* e *vanN*, sendo os genótipos *vanA* e *vanB* os mais frequentes (KLARE et al., 2003; LEBRETON et al., 2011; WERNER et al., 2008; XU et al., 2010). Os fenótipos VanA e VanB têm maior importância clínica, estando frequentemente associados a cepas de *E. faecium* (COURVALIN, 2006; KLARE et al., 2003). Isso porque amostras VREfm apresentam elevada capacidade de disseminação, sendo rapidamente transmitidas de paciente para paciente, ocasionando com frequência situações de surto e/ou endemidade. A identificação precoce desses microrganismos é fundamental para controlar a sua disseminação (MCGOWAN, 2004). Recomendações para a prevenção da disseminação da resistência adquirida à vancomicina incluem a vigilância rotineira das amostras das espécies *E. faecalis* e *E. faecium*, visando a detecção de amostras resistentes e o estabelecimento de um programa de triagem da colonização intestinal (TACCONELLI; CATALDO, 2008).

O gene *vanA* é parte integrante do Tn1546 ou transposons relacionados, que são geralmente localizados em plasmídeos transferíveis (ARTHUR et al., 1993; HOLLENBECK; RICE, 2012; PALEPOU et al., 1998). Os setes genes que participam da expressão do fenótipo VanA estão divididos em três grupos funcionais. Os genes *vanS* e *vanR* participam da regulação da resposta. Estes são os responsáveis pela expressão de proteínas relacionadas à detecção da presença do antimicrobiano no meio extracelular e na sinalização intracelular para ativação da transcrição dos genes *vanHAX*, que desempenham o papel efetor da resistência, e de *vanYZ* que são genes acessórios (ARTHUR; QUINTILIANI, 2001; KOTEVA et al., 2010).

Recomendada para o tratamento de infecções do trato urinário causadas por enterococos (CLINICAL LABORATORY STANDARDS INSTITUTE, 2019), a classe das quinolonas, na qual estão incluídas a ciprofloxacina, a norfloxacina e a levofloxacina, atua sobre duas enzimas essenciais no processo de replicação do DNA bacteriano: a DNA girase e a topoisomerase IV (CORREIA et al., 2017). No entanto, o surgimento de cepas resistentes já foi amplamente descrito (DESHPANDE et al., 2007; KIM et al., 2018; KIM; WOO, 2017; LÓPEZ et al., 2011; TORELL et al., 2003). O mecanismo de resistência a esses antimicrobianos é atribuído, principalmente, a mutações nas regiões chamadas de regiões determinantes de resistência às quinolonas (QRDR – do inglês, *quinolone resistance-determining regions*) nos genes *gyrA* e *gyrB* para DNA girase e *parC* e *parE* para

topoisomerase IV (AMIN; JALAL; WRETLIND, 1999; BRISSE et al., 1999). A resistência às quinolonas é uma das características mais associadas às cepas de *E. faecium* de origem hospitalar (WERNER et al., 2010).

Atualmente, a limitação de opções terapêuticas disponíveis para o tratamento de infecções graves por *E. faecium* resistentes à vancomicina é cada vez mais preocupante. Assim, vem sendo recomendado o uso de quinupristina/dalfopristina, uma combinação de estreptograminas, que atuam inibindo a síntese proteica através da interação com a subunidade 50S do ribossomo (NIGO, 2014). Registros da ocorrência de efeitos adversos e o fato de que muitas cepas dessa espécie carregam o gene *erm(B)*, que elimina o efeito bactericida desta combinação de antimicrobianos, levam a limitações importantes para sua utilização segura e mais ampla (LÓPEZ et al., 2010).

Dentre os antimicrobianos mais recentes aprovados para o tratamento de casos de infecção por *E. faecium* estão a linezolida e a daptomicina (BIRMINGHAM et al., 2003; CANTÓN et al., 2010). A primeira, pertencente à classe das oxazolidinonas, atua inibindo a síntese proteica (ARIAS; MURRAY, 2008). Entretanto, ainda que em percentuais não elevados, já foi relatado o surgimento de resistência, durante o tratamento com linezolida (SEEDAT et al., 2006). Os mecanismos de resistência envolvidos estão relacionados a mutações no rRNA 23S e à aquisição de uma metiltransferase codificada pelos genes *cfr* (ARIAS et al., 2008). A daptomicina é um lipopeptídeo cíclico com efeito bactericida, que atua nos canais de cálcio da membrana celular bacteriana, causando a sua consequente despolarização. A taxa de amostras não suscetíveis, em um estudo conduzido nos EUA, foi de 3,9% (EDELSBERG et al., 2014). Mutações nos genes *liaFSR*, responsáveis pela regulação da resposta ao estresse celular, já foram associadas a um aumento no valor da CIM (concentração inibitória mínima) para esse antimicrobiano (MUNITA et al., 2012).

Considerando a importância de *E. faecium* em IRAS, particularmente cepas multirresistentes (MDR – do inglês, *multi-drug resistant*), a espécie foi incluída no grupo dos microrganismos de risco à saúde humana, denominado “ESKAPE”, juntamente com *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* e *Enterobacter* spp. (SANTAJIT; INDRAWATTANA, 2016). Um documento preparado pela Organização Mundial da Saúde (WORLD HEALTH ORGANIZATION, 2017), definiu uma lista de patógenos bacterianos que inclui *E. faecium* resistentes à vancomicina como prioridade mundial, determinando a necessidade de esforços imediatos para pesquisa e desenvolvimento de novos antimicrobianos e seu efetivo controle.

Dentre os determinantes adquiridos por *E. faecium* através de processos de THG, estão também aqueles associados aos fatores de virulência. Embora alguns genes tenham sua importância já estabelecida em amostras relacionadas ao ambiente hospitalar, o conhecimento da virulência da espécie ainda é limitado. Entre os fatores secretados, o único considerado específico de *E. faecium* é a proteína SagA, que tem a função de se ligar amplamente às proteínas da matriz extracelular (GAO; HOWDEN; STINEAR, 2018; TENG et al., 2003). Regiões repetitivas dentro do gene *sag* foram observadas em amostras clínicas e contribuíram para a formação de biofilme (PAGANELLI et al., 2015). Em um estudo realizado no ano de 2011, os pesquisadores obtiveram resultados promissores para a utilização dessa proteína como eficaz candidata para compor uma futura vacina. Anticorpos criados especificamente contra a proteína de SagA se mostraram eficientes opsoninas, contribuindo assim na consequente destruição das células de *E. faecium* (KROPEC et al., 2011).

Dentre os genes associados a amostras de *E. faecium* de origem hospitalar, pode-se também destacar *hylefm*. Identificada erroneamente como uma hialuronidase, a proteína Hyl na verdade é uma glicosil hidrolase, que ainda não demonstrou evidências diretas quanto à sua contribuição na patogênese da espécie (LAVERDE GOMEZ et al., 2011; PANESSO et al., 2011). Outros fatores secretados, tais como uma citolisina, codificada pelo conjunto gênico *cyl* (sendo o gene *cylA* responsável pela expressão da atividade citolítica), gelatinase (codificada pelo gene *gelE*) e uma substância de agregação (codificada pelo gene *asaI*), também ainda não apresentaram um papel bem definido e raramente são identificadas em amostras dessa espécie. Para a maioria dos estudos disponíveis na literatura especializada, fatores de virulência em enterococos estão mais relacionados à espécie *E. faecalis* (CRETI et al., 2004; MADSEN et al., 2017; MUNDY; SAHM; GILMORE, 2000; VAN TYNE; GILMORE, 2014).

A espécie possui também um complexo arranjo de proteínas de superfície, sendo muitas delas envolvidas no processo de colonização e, portanto, requeridas para sua patogenicidade. Estruturas como os ácidos lipoteicóicos (LTA - do inglês, *lipoteichoic acids*), ácidos teicóicos de parede (WTA - do inglês, *wall teichoic acids*) e a cápsula polissacarídica (codificada no locus *cps*) são componentes associados à evasão da resposta imune inata mediada por complemento e morte por neutrófilos (GEISS-LIEBISCH et al., 2012; HANCOCK; GILMORE, 2002; WEIDENMAIER; PESCHEL, 2008). O LTA também já foi associado à formação de biofilmes (FABRETTI et al., 2006).

Uma extensa variedade de proteínas de ancoragem do tipo LPXTG (cerca de 24 distintas) foi descrita como estruturas de superfície na espécie *E. faecium*. Dessas, MEsp

(proteína de superfície de enterococos), Acm (adesina de colágeno), Scm (segunda adesina de colágeno), SgrA (adesina de superfície de ligação ao nidogênio) e EcbA [MSCRAMM (do inglês, *microbial surface components recognizing adhesive matrix molecules*) ligadora de colágeno] são descritas como importantes determinantes de virulência (FREITAS, A.R et al., 2018a; HENDRICKX et al., 2007).

As proteínas Acm e Esp contribuem na adesão inicial à célula hospedeira e sua participação na patogênese da endocardite causada pela espécie já foi demonstrada experimentalmente (HEIKENS et al., 2011; NALLAPAREDDY et al., 2008). A proteína Esp é o fator de virulência mais amplamente distribuído em amostras clínicas de *E. faecium*, sendo um marcador de linhagens associados ao ambiente hospitalar. Contribui para a formação de biofilmes e está associada à patogênese das infecções urinárias ascendentes (HEIKENS; BONTEN; WILLEMS, 2007; PANESSO et al., 2010; WILLEMS et al., 2001). Já as proteínas Acm e Scm se ligam ao colágeno, do tipo I e do tipo V, respectivamente, além de fibrinogênio (NALLAPAREDDY et al., 2008; SILLANPAA et al., 2008). A SgrA se liga ao nidogênio, componente da lâmina basal do tecido epitelial; enquanto EcbA se liga ao colágeno do tipo V. Além disso, ambas se ligam ao fibrinogênio. Mesmo não sendo ainda muito claro o papel dessas proteínas no processo infeccioso, estima-se que a sua expressão ocorra de forma organizada e dependente da fase de crescimento bacteriano (HENDRICKX et al., 2009).

Dentre os elementos que contribuem para a plasticidade genética da espécie estão elementos genéticos móveis (EGMs), como: plasmídeos, transposons, sequências de inserção (IS - do inglês, *insertion sequence*) e profagos (HEGSTAD et al., 2010; LEAVIS et al., 2007). Esses elementos contribuem para a alta diversidade genética de *E. faecium*, carreando diversos genes que conferem vantagens seletivas, tais como os determinantes de resistência e virulência já mencionados, além de sistemas toxina/antitoxina (MORITZ; HERGENROTHER, 2007), determinantes de resistência a metais pesados (LAVERDE GOMEZ et al., 2011) e genes envolvidos no metabolismo de carboidratos (ARREDONDO-ALONSO et al., 2020). A abundância de EGMs, principalmente em amostras clínicas humanas, revelou a importância do mobiloma como fator que impulsiona a emergência e evolução de linhagens associados ao ambiente hospitalar (LEAVIS et al., 2007; MIKALSEN et al., 2015; SANTAGATI; CAMPANILE; STEFANI, 2012).

Alguns sistemas, como o de repetições palindrômicas curtas agrupadas e regularmente inter espaçadas associadas à proteína cas (CRISPR-Cas - do inglês, *clustered regularly interspaced short palindromic repeats and associated cas protein*) e de restrição e modificação (R-M), foram muito bem descritos em espécies como *E. faecalis* por impedir a

aquisição de EGMs, principalmente determinantes associados à resistência aos antimicrobianos (PRICE et al., 2016). Entretanto, são raramente descritos em *E. faecium*, questionando, assim, sua função efetiva nesta espécie (PALMER; GILMORE, 2010). Por outro lado, alguns estudos têm demonstrado que esses sistemas, quando presentes, parecem contribuir para a estrutura e a divergência dentro da espécie (HUO et al., 2019; LEBRETON et al., 2013).

Histórico Evolutivo da Espécie *E. faecium*

Com o aumento da frequência de isolamento de *E. faecium* nos EUA, em um primeiro momento causado por amostras resistentes à ampicilina (AREfm - do inglês, *ampicillin-resistant Enterococcus faecium*), e posteriormente seguido pelo surgimento rápido de cepas VRE, um maior número de estudos de vigilância começou a ser realizado na tentativa de melhor compreender os aspectos envolvidos na persistência e disseminação dessas linhagens específicas, que naquele momento circularam no ambiente hospitalar (CHAVERS et al., 2003; FRIDKIN et al., 1998; MATO et al., 1996).

Nos estudos iniciais, a técnica de eletroforese em campo pulsado (PFGE - do inglês, *pulsed-field gel electrophoresis*) uma das precursoras, foi considerada como padrão-ouro para tipificação molecular desses microrganismos (DICUONZO et al., 2001; KUHN et al., 1995; MALATHUM et al., 1998; PEGUES et al., 1997; WERNER; KLARE; WITTE, 2007). Contudo, o uso de PFGE por laboratórios de rotina sempre foi considerado relativamente limitado, devido a necessidade de equipamentos específicos, que demandam elevados recursos financeiros. Adicionalmente, uma dificuldade inerente a essa metodologia é que os resultados obtidos por diferentes laboratórios não podem ser facilmente comparados, reduzindo a eficácia na reprodutibilidade dos ensaios. Além disso, os resultados expressos como padrões de bandas são apenas parcialmente informativos, no que diz respeito às relações filogenéticas (ABELE-HORN et al., 2006; JUREEN et al., 2004).

Assim, métodos baseados no sequenciamento de genes ou, ainda, do genoma completo, passaram a ser recomendados e muitas vezes considerados como padrão-ouro para caracterização do relacionamento genético de amostras bacterianas. A metodologia de tipificação de sequências em múltiplos loci (MLST - do inglês, *multilocus sequence typing*)

tem sido referida como a que agrega um maior número de vantagens para as análises de diversidade de microrganismo (JOLLEY; MAIDEN, 2010).

A técnica de MLST foi proposta por Maiden et al. (1998), como um método mais prático para identificar relações genéticas em bactérias. Essa metodologia é baseada na caracterização de sequências alélicas em um conjunto de genes de manutenção celular (no inglês, *housekeeping genes*), que caracterizam tipos (ou clones) e complexos clonais (CCs), de relevância clínica para diversas espécies bacterianas, que podem ser comparados em bancos de dados internacionais. Cada locus analisado apresenta uma determinada sequência que recebe um determinado número (número do alelo) com o propósito de identificação. Sequências que diferem em um ou mais nucleotídeo são consideradas como alelos diferentes. O conjunto de números de todos os alelos representa um perfil que define a sequência tipo (ST, do inglês, *sequence type*) de cada amostra bacteriana. Em técnicas baseadas na comparação de alelos, como no caso do MLST, cada mudança na sequência nucleotídica é contada como um evento genético único e independe do número de polimorfismos envolvidos, resultando em uma abordagem simplificada, porém informativa em nível de estrutura populacional (MAIDEN et al., 2013).

No esquema de MLST para a caracterização de amostras de *E. faecium* são analisados fragmentos internos dos seguintes sete genes de manutenção celular: *adK*, *atpa*, *ddl*, *gdh*, *gyd*, *pstS* e *purk*, que codificam para adenilato quinase, subunidade alfa da ATP sintetase, D-alanina-D-alanina ligase, glicose 6-fosfato desidrogenase, gliceraldeído 3-fosfato desidrogenase, proteína transportadora de cassete de ligação de ATP e fosforribosilaminoimidazol carboxilase, respectivamente (HOMAN et al., 2002). Um banco de dados internacional de linhagens de *E. faecium*, analisadas por MLST, está disponível através do PubMLST (JOLLEY; BRAY; MAIDEN, 2018).

Assim, os primeiros resultados utilizando a técnica de MLST demonstraram que a escalada desse microrganismo do anonimato comensal para o protagonismo no cenário das infecções hospitalares tem sido associada à presença de subpopulações especializadas em conquistar diferentes ambientes. As primeiras impressões sobre a estrutura populacional de *E. faecium* permitiram examinar os padrões de descendência evolutiva nesta espécie. Mesmo em estudos iniciais, os dados já apontavam a existência de uma subpopulação distinta relacionada às amostras hospitalares. Essa subpopulação foi posteriormente melhor caracterizada e descrita como um complexo clonal, denominado CC17 (WILLEMS et al., 2005).

Amostras vinculadas ao CC17 apareceram como responsáveis pela grande maioria das infecções nosocomiais e dos surtos causados por *E. faecium* nos cinco continentes. Seu sucesso no ambiente hospitalar foi, pelo menos em parte, relacionado com atributos de virulência e de resistência aos antimicrobianos. Entre as amostras pertencentes ao CC17, destacaram-se as resistentes à ampicilina e às quinolonas; que contam também, com a presença de uma putativa ilha de patogenicidade onde se inserem os genes *esp* e *hyl*. A presença destas características possivelmente torna as amostras capazes de sobreviver e de se disseminar, de forma mais eficiente, em ambientes que exigem adaptação contínua, como os de um hospital (CHANG et al., 2010; COQUE et al., 2002; LEAVIS; BONTEN; WILLEMS, 2006; WILLEMS et al., 2005).

A emergência da resistência à vancomicina na população de *E. faecium* foi detectada posteriormente ao estabelecimento do CC17 no âmbito hospitalar. Existem indícios de que no ganho do espaço hospitalar predominaram amostras resistentes à ampicilina e portadoras do gene *esp* (COQUE et al., 2002; LEAVIS; BONTEN; WILLEMS, 2006; WILLEMS et al., 2001; 2005), as quais subsequentemente adquiriram resistência à vancomicina. Além disso, a presença do alelo 1 da fosforribosilaminoimidazol carboxilase codificada pelo gene *purK*, que está incluído no esquema de tipificação por MLST, foi o inicialmente mais frequente encontrado nesta população e por isso considerado um importante marcador de amostras pertencentes ao CC17 (WILLEMS et al., 2001; WERNER et al., 2008; PANESSO et al., 2010).

Apesar da clara importância do CC17, como a principal subpopulação genética de *E. faecium* de origem hospitalar, a sua precisa estrutura filogenética permanecia incerta. Um dos primeiros algoritmos empregados nas análises dos dados obtidos por MLST foi o eBURST, que apontou que o CC17 apresentava uma única origem, sendo constituído pelo fundador primário ST17 e subfundadores, destacando-se as ramificações dos ST18 e ST78 que, juntamente com seus variantes de locus único (SLVs - do inglês, *single-locus variant*) e variantes de duplo locus (DLVs - do inglês, *double-locus variant*), foram responsabilizados pela ocorrência de surtos em diferentes partes do mundo (HSIEH et al., 2010, KHAN et al., 2010, RYAN et al., 2015; SILVA, 2012; STAMPONE et al., 2005; WERNER et al., 2008; ZHU et al., 2009).

Mais recentemente, diversos estudos têm apontado mudanças significativas na epidemiologia das infecções causadas por *E. faecium* em diferentes países. Tais mudanças relatam o surgimento e a predominância de amostras pertencentes ao ST412, um DLV do ST78 (DAMANI et al., 2010; PANESSO et al., 2010; RIOS et al., 2020).

No Brasil, os primeiros relatos do aparecimento de amostras do ST412 ocorreram na região sul do país (ALVES, 2017; PALAZZO et al., 2011; PRADO, 2016; SACRAMENTO, 2017; SILVA, 2012). Em um estudo anterior realizado pelo nosso grupo, foi observado que o ST78, apontado até então como um dos mais frequentes em infecções de origem hospitalar no estado do Rio de Janeiro, se manteve predominante entre amostras VREfm até o ano de 2007, quando foi constatado o surgimento e a disseminação do ST412, que passou a ser o predominante (FARIA, 2012). Cepas de *E. faecium* pertencentes ao ST412 têm sido reportadas na literatura albergando resistência à linezolida e daptomicina, que são os antimicrobianos mais recentes disponibilizados como alternativa para o tratamento das infecções enterocócicas causadas por amostras multirresistentes, incluindo a resistência à vancomicina (PRADO et al., 2016).

As constantes mudanças no curso epidemiológico de *E. faecium* que passaram a ser observadas indicaram que essa espécie está frequentemente submetida a uma elevada taxa de recombinação homóloga (TURNER et al., 2007; WILLEMS et al., 2010). Como altas taxas de recombinação podem levar a erros em análises filogenéticas, foi considerado que as abordagens empregando o algoritmo eBURST para definição da estrutura do CC17 eram sujeitas a consequências indesejadas. De fato, já foi reconhecido que a grande quantidade de recombinação a que esta espécie é submetida resultou em um agrupamento ilegítimo de diversas linhagens distintas em um único complexo clonal (TURNER et al., 2007, WILLEMS et al., 2010, 2011, 2012)

Diante disso, análises alternativas ao método de eBURST passaram a ser utilizadas e sugeriram que as linhagens componentes do CC17, como ST17, ST18 e ST78, são distintas. Dessa forma, essas linhagens de *E. faecium* de origem hospitalar poderiam não ter evoluído a partir de um único fundador (isto é, ST17) e, conseqüentemente, o CC17 como estaria até então sendo definido, não existiria (WILLEMS et al., 2012).

Uma das abordagens alternativas mais recentes utiliza a análise Bayesiana de estrutura populacional (BAPS - do inglês, *Bayesian Analysis of Population Structure*), que tem como princípio um modelo genético estatístico para particionar a variação molecular com base em padrões de ascendência e recombinações clonais identificadas, em vez de tentar recuperar informações filogenéticas ou implementar um modelo de agrupamento filogenético como os algoritmos anteriormente utilizados (WILLEMS et al., 2012). Esta abordagem também já foi realizada para examinar a estrutura populacional de alguns microrganismos, como *Streptococcus pneumoniae* (CHAGUZA et al., 2016), *Escherichia coli* (SUNDQVIST et al.,

2014), *Campylobacter jejuni* (SHEPPARD; MAIDEN, 2015) e *Neisseria* spp. (CORANDER et al., 2012).

O estudo realizado por Willems et al. (2012), utilizando BAPS para identificar grupos de amostras de *E. faecium*, sugeriu que linhagens associadas ao hospital estavam contidas em diferentes grupos que têm, no entanto, adquirido elementos adaptáveis semelhantes. Esse estudo demonstrou também que a grande maioria das amostras de *E. faecium* oriundas de infecções adquiridas no ambiente hospitalar foi reunida em dois grupos: BAPS 2-1, constituído pela linhagem 78, e BAPS 3-3 constituído pelas linhagens 17 e 18.

A análise BAPS indicou que amostras clínicas não estão relacionadas com estirpes de *E. faecium* que atuam como comensais nos seres humanos, mas, em vez disso, a evolução genética de linhagens hospitalares possivelmente envolve amostras animais. Assim, para a população de *E. faecium* hospitalar, foi proposto um modelo evolutivo em que as cepas com uma alta capacidade para colonizar e infectar pacientes hospitalizados surgiram através de transferência horizontal de genes. Uma vez adaptada para o nicho hospitalar, esta subpopulação se tornou isolada e a recombinação com outras populações declinou, resultando em um modelo de evolução convergente (WILLEMS et al., 2012).

Na atualidade, o emprego de WGS trouxe um grande avanço as mais diversas abordagens nas diferentes subáreas da Microbiologia. Análises por WGS para avaliação do pangenoma, que consiste da reunião do *core* genoma ou genoma central e genes acessórios (ou dispensáveis), têm contribuído para o estudo da dinâmica epidemiológica das infecções hospitalares. Por exemplo, estratégias empregando WGS já foram utilizadas para auxiliar na construção e implantação de esquemas de intervenção para conter o aumento da disseminação de amostras VRE em ambientes hospitalares (HOWDEN et al., 2013; JOHNSON et al., 2010). Raven et al. (2017) utilizaram a metodologia de WGS para obter uma resolução aprimorada da complexa rota de transmissão de *E. faecium* inter e intra-hospitalar, determinando a grandiosidade do problema, que não se mostrou restrito a uma escala local ou nacional, mas que se apresentou de importância global.

Outros achados, que apenas se tornaram possíveis devido à análise das sequências genômicas por WGS, apoiaram o conceito de um clado associado ao hospital que é geneticamente distinto dos demais, os quais incluem amostras comensais recuperadas a partir de animais e seres humanos. As primeiras reconstruções filogenéticas utilizando o genoma completo de *E. faecium* revelaram divisão da espécie em dois grandes grupos, designados como clado A e B, ou clado associado ao hospital e clado associado a comunidade (GALLOWAY-PEÑA et al., 2012; PALMER et al., 2012). No clado A estariam,

principalmente, as amostras clínicas humanas, incluindo aquelas que foram associadas ao CC17, enquanto no clado B estariam amostras isoladas de indivíduos saudáveis. Posteriormente, foi proposta uma subdivisão para o clado A, onde A1 agruparia as amostras clínicas humanas e A2 reuniria amostras oriundas de animais (LEBRETON et al., 2013). Entretanto, outros trabalhos refutaram a hipótese de subdivisão do clado A, demonstrando que a complexa reconstrução populacional de *E. faecium* ainda é objeto de investigação (RAVEN et al., 2016; RIOS et al., 2020).

Além disso, a utilização de WGS associado a ferramentas de bioinformática permite também a obtenção de informações sobre o caminho evolutivo da espécie, desde um microrganismo comensal a um reconhecido patógeno de infecções hospitalares. A construção de um modelo de relógio molecular permitiu, por exemplo, estimar que o surgimento dos enterococos foi a 425-500 milhões de anos, concomitante à passagem dos animais para a vida terrestre, através da seleção de características fisiológicas necessárias para sua chegada e sobrevivência no ambiente hospitalar (LEBRETON et al., 2017).

Diante do exposto, pode ser considerado que, dada a importância de *E. faecium* no cenário hospitalar em todo o mundo, o emprego de estratégias atuais, como as análises por sequenciamento do genoma completo, é de notável relevância para elucidar os processos evolutivos que resultam em uma maior adaptabilidade e, portanto, circulação mais ampla desses microrganismos. O entendimento da estrutura populacional e de eventos genéticos relacionados podem ajudar a entender os mecanismos de emergência de novas linhagens altamente adaptadas ao ambiente hospitalar, que apresentam um risco potencial à saúde humana.

1 OBJETIVOS

1.1 Objetivo Geral

Determinar a estrutura populacional de amostras de *E. faecium* resistentes a múltiplos antimicrobianos que circularam no período de 2009 a 2016 em instituições hospitalares do estado do Rio de Janeiro e avaliar diferentes aspectos genéticos e evolutivos associados à emergência e disseminação dessas populações no ambiente hospitalar.

1.2 Objetivos específicos:

- a) determinar a susceptibilidade a antimicrobianos em amostras de *E. faecium* isoladas em instituições durante um período de oito anos consecutivos;
- b) investigar a diversidade de mecanismos de resistência a antimicrobianos e de fatores de virulência através da detecção de elementos genéticos associados a essas características;
- c) realizar o sequenciamento do genoma completo de amostras de *E. faecium* e a análise por diferentes ferramentas de bioinformática;
- d) determinar a estrutura populacional através da análise do polimorfismo genético e caracterização os complexos clonais predominantes;
- e) determinar o conteúdo genético em relação à aquisição de elementos associadas à evolução de linhagens multirresistentes de *E. faecium* circulantes em nosso meio;
- f) determinar os efeitos de mutações e eventos de recombinação homóloga na emergência e expansão das diferentes populações no ambiente hospitalar;
- g) reconstruir a filogenia e determinar a estimativa do tempo de divergência clonal entre as diferentes populações de *E. faecium* circulantes no período.

2 MATERIAL E MÉTODOS

2.1 Amostras Bacterianas

Este trabalho incluiu a análise de 205 amostras bacterianas pertencentes à espécie *E. faecium*. Tais amostras foram obtidas ao longo de oito anos (2009 a 2016) de diferentes fontes clínicas, representativas de quadros de infecção e de colonização, provenientes de pacientes hospitalizados em instituições do Rio de Janeiro. A seleção amostral foi feita a partir de um universo de 547 amostras de *E. faecium* que foram recebidas em nosso laboratório no período estudado. O cálculo amostral significativo foi feito utilizando a fórmula abaixo:

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1 - p)}{Z^2 \cdot p \cdot (1 - p) + e^2 \cdot (N - 1)}$$

Onde **n** é amostra calculada; **N**, a população; **Z**, a variável padronizada do nível de confiança; **p**, a verdadeira probabilidade do evento e **e**, o erro amostral. Foi utilizado um valor de **p** de 50%, com erro amostral de 5% e nível de confiança de 95%.

Utilizando esta fórmula, o **n** amostral significativo foi estabelecido em 226 amostras, e dividindo-se esse valor pelos oito anos estudados, seria necessária a inclusão de 28,25 amostras por ano. Foi, então, estabelecido um critério de seleção e inclusão de 30 amostras por ano. Entretanto, considerando que em dois anos, compreendidos no período avaliado, não foi possível obter o número desejável de 30 amostras bacterianas (2011 = 11 amostras; 2012 = 14 amostras), este estudo passou a considerar um erro amostral de 5,5% com um nível de confiança de 95%, para um total de 205 amostras.

Os critérios para a seleção de amostras bacterianas neste estudo, visando contemplar todo o período, foram baseados na disponibilidade de informações quanto: ao sítio de isolamento - foram incluídas todas as amostras representativas tanto de infecção quanto de colonização; à instituição de origem; resultados preliminares de resistência à vancomicina, ampicilina, ciprofloxacina e aminoglicosídeos.

Todas as amostras estudadas pertencem à coleção de culturas do Laboratório de Apoio Biotecnológico, do Instituto de Microbiologia Paulo de Góes, da Universidade Federal do Rio

de Janeiro, sob a coordenação da Prof.^a Lúcia Martins Teixeira. As amostras bacterianas estavam mantidas congeladas a -20°C , em solução de leite desnatado a 10% (*Skim Milk*, BD Diagnostics, Franklin Lakes, NJ, EUA) acrescido de 10% de glicerol. As amostras já possuíam uma caracterização fenotípica prévia quanto a gênero e espécie, por um painel de testes fenotípicos convencionais (TEIXEIRA et al., 2019).

Para reativação, as amostras foram semeadas em meio ágar sangue de carneiro - AS (PlastLabor Ind e Com. Materiais e Equipamentos de Laboratório, Rio de Janeiro, RJ, BR), e incubadas por 18h-24h a $36\pm 1^{\circ}\text{C}$. Após esse período, as placas foram avaliadas para evidencição do crescimento característico de enterococos (morfologia colonial e atividade hemolítica) em culturas puras.

2.2 Caracterização Molecular por Espectrometria de Massa - MALDI TOF MS

Para confirmar a identificação das amostras bacterianas, foi utilizada a técnica de MALDI-TOF MS. Para tal, as amostras foram semeadas em placas de AS (PlastLabor) e em seguida passaram por incubação por 18h-24h a uma temperatura de $36 \pm 1^{\circ}\text{C}$. A partir do crescimento assim obtido, uma colônia foi depositada, com auxílio de uma haste de madeira, em uma das áreas delimitadas na placa de aço inoxidável polida (MSP 96 *target polished steel* BC, Bruker Daltonics®, Billerica, MA, EUA) e depois o material foi coberto com $1\mu\text{l}$ da matriz HCCA (ácido α -ciano-4-hidroxicinâmico, Bruker Daltonics®). As amostras foram avaliadas em um espectrômetro de massa Microflex LT (Bruker Daltonics®) em modo linear, utilizando-se o software MALDI Flex Control 3.3 no modo automático (MBT_FC.par). Os espectros de massas foram coletados na faixa de 2.000 a 20.000 m/z e analisados com auxílio do software MALDI Biotyper 2.0 (Bruker Daltonics®).

2.3 Determinação da Suscetibilidade aos Antimicrobianos

A suscetibilidade a 18 antimicrobianos foi avaliada pela metodologia de difusão em ágar (disco-difusão). Para tal, as amostras bacterianas foram previamente semeadas em placas de AS (PlastLabor) e incubadas por 18h-24h a uma temperatura de $36\pm 1^{\circ}\text{C}$. A preparação dos

inóculos foi realizada em solução salina estéril (NaCl 0,85%) com padrão de turbidez ajustado para a escala 0,5 de McFarland. Cada suspensão foi semeada em ágar Müeller-Hinton (Difco Labs., Franklin Lakes, NJ, USA) e, em seguida, os discos de antimicrobianos foram aplicados com o auxílio de uma pinça previamente esterilizada.

Os discos testados (Oxoid Brasil Ltda, São Paulo, SP, Brasil) abrangeram 13 classes distintas de antimicrobianos: β -lactâmicos (ampicilina 10 μ g e penicilina 30 μ g); ansamicinas (rifampicina 5 μ g); estreptograminas (quinupristina/dalfopristina 15 μ g); quinolonas (ciprofloxacina 5 μ g, norfloxacina 10 μ g e levofloxacina 5 μ g); fenicóis (cloranfenicol 30 μ g); macrolídeos (eritromicina 15 μ g); glicilciclinas (tigeciclina 15 μ g); fosfomicinas (fosfomicina 200 μ g); oxazolidinonas (linezolida 30 μ g); furantoínas (nitrofurantoína 300 μ g); tetraciclina (tetraciclina 30 μ g); glicopeptídeos (vancomicina 30 μ g e teicoplanina 30 μ g) e aminoglicosídeos em concentrações elevadas (gentamicina 120 μ g e estreptomina 300 μ g).

As placas foram incubadas por 18h, exceto para vancomicina e teicoplanina, onde o tempo de incubação padronizado foi de 24h, a uma temperatura de $36\pm 1^\circ\text{C}$ e, após este período, foram realizadas a leitura dos halos de inibição e a interpretação conforme determinação do *Clinical and Laboratory Standards Institute* (CLSI, 2019), exceto para a tigeciclina em que foram seguidas as recomendações de interpretação do *European Committee on Antimicrobial Susceptibility Testing* (EUCAST, 2019). Para fins de controle de qualidade foram empregadas as amostras de referência *Staphylococcus aureus* ATCC25923 e *Enterococcus faecalis* ATCC29212. As amostras foram classificadas como multirresistentes seguindo os critérios de Magiorakos et al. (2012).

2.4 Determinação da Concentração Inibitória Mínima (CIM) de Estreptomina

As amostras que apresentaram incongruências entre os resultados obtidos nos testes fenotípicos (determinação da susceptibilidade por disco-difusão) e genotípico (detecção de determinantes de resistência por PCR convencional e por sequenciamento do genoma completo) foram também avaliadas quanto à determinação da CIM, com o emprego do teste de diluição em ágar. A metodologia utilizada e os critérios de interpretação seguiram as recomendações de execução e interpretação estabelecidas pelos documentos M07-A10 e M100 do CLINICAL LABORATORY STANDARDS INSTITUTE (2015; 2019). Para a realização do teste, inicialmente foram preparadas placas de Müeller-Hinton Agar (Difco

Labs.) contendo diluições duplas seriadas de estreptomicina (Sigma-Aldrich Co., San Louis, MO, EUA), nas concentrações finais de 0,25 µg/mL a 4.096 µg/mL, além de placas sem a adição de antimicrobianos, que foram utilizadas como controle do crescimento bacteriano. Volumes de 1 µl de suspensões bacterianas confeccionadas em salina fisiológica (a partir de crescimento recente em AS), correspondendo à turvação equivalente a 0,5 McFarland, foram inoculados nas placas contendo as diferentes concentrações dos antimicrobianos, com auxílio de um multi-inoculador (replicador) do tipo Steers. A CIM foi considerada como a menor concentração onde não foi detectado visualmente crescimento bacteriano, após incubação a 36°C±1°C, por até 48h. As amostras de referência *E. faecalis* ATCC 29212 (sensível) e ATCC 51299 (que apresenta o fenótipo HLAR) foram utilizadas como controles dos testes.

2.5 Detecção dos Genes Associados à Resistência a Antimicrobianos e à Virulência

Para a detecção dos genes envolvidos na resistência aos aminoglicosídeos e aos glicopeptídeos e na virulência de *E. faecium*, foi utilizada a metodologia da reação em cadeia da polimerase otimizada em reações *multiplex* únicas ou em etapas sequenciais. Os protocolos utilizados para a detecção dos genes envolvidos na resistência a níveis elevados de aminoglicosídeos, glicopeptídeos e na virulência de *Enterococcus* spp. seguiram recomendações anteriores. Na seção 2.5.5 (Quadro 1) estão listados os genes pesquisados, bem como as sequências dos iniciadores utilizados, os tamanhos esperados dos fragmentos amplificados e as respectivas referências.

2.5.1 Obtenção do DNA Bacteriano

O DNA bacteriano foi obtido utilizando-se a resina Chelex 100 (Bio-Rad Laboratories, Hercules, CA, EUA) conforme protocolo proposto por Pinto et al. (2013). A partir de um crescimento bacteriano meio AS (PlastLabor) foram preparadas suspensões bacterianas em tubos tipo Eppendorf contendo 200 µl de uma solução aquosa da resina Chelex 100 a 5% e 0.2 mg/mL de proteinase K (Invitrogen, Carlsbad, CA, USA). As suspensões foram, primeiramente, incubadas a 50 °C por 1 h, em seguida homogeneizadas em vortex, por 10 s, e

submetidas a uma nova incubação a 95°C, durante 10 min. Após centrifugação a 12.000 rpm por 3 min, o sobrenadante contendo o DNA molde foi utilizado nos testes.

2.5.2 Detecção dos Genótipos Associados à Resistência a Níveis Elevados de Aminoglicosídeos

A presença de genes que determinam resistência a níveis elevados de estreptomicina foi pesquisada por PCR *multiplex*, de acordo com protocolo desenvolvido por nosso grupo (FARIA, 2017), utilizando iniciadores previamente descritos para os genes *ant(6')-Ia* (SWENSON et al., 1995), *ant(9')-Ia* e *ant(9')-Ib* (MAHBUB ALAM et al., 2005). A análise da presença dos genes que determinam a resistência a níveis elevados de gentamicina e outros aminoglicosídeos também foi pesquisada por PCR *multiplex*, seguindo protocolo descrito por Vakulenko et al. (2003), com algumas modificações. Foram pesquisados os seguintes determinantes genéticos: *aac(6')-Ie-aph(2'')-Ia*, *aph(2'')-Ib*, *aph(2'')-Ic*, *aph(2'')-Id*, *aph(3')-IIIa* e *ant(4')-Ia*.

Para a pesquisa dos genes associados à resistência à estreptomicina, foi feita uma reação de PCR *multiplex* com uma mistura no volume final de 25 µl. A reação foi constituída de 2,5 µl do extrato de DNA molde, 1x do tampão PCR (Ludwig Biotecnologia Ltda, Alvorada, RS, BRA), 3,0 mM de cloreto de magnésio – MgCl₂ (Ludwig Biotecnologia Ltda), 0,4 mM de cada dNTP (Invitrogen) e 1,25 U da enzima Taq DNA polimerase (Ludwig Biotecnologia Ltda), além dos oligonucleotídeos iniciadores conforme descritos na seção 2.5.5 (Quadro 1). Os ciclos de amplificação foram realizados em termociclador Veriti 96 Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA), sendo compostos de uma etapa inicial de desnaturação de 5 min a 95°C, seguida de 30 ciclos a 94°C por 1 min para desnaturação, uma fase de anelamento a 55°C por 1 min e uma fase de extensão a 72°C por 1 min. Em seguida foi realizada uma etapa de extensão final a 72°C por 10 min. As amostras utilizadas como controles positivos foram: *E. faecium* CL8287 para *ant(6')-Ia* e *ant(9')-Ia* e *E. faecalis* Fa-23 para *ant(9')-Ib*.

A pesquisa dos genes que codificam a resistência à gentamicina e outros aminoglicosídeos também foi realizada por PCR *multiplex*. A mistura da reação teve um volume final de 25 µl e foi constituída de 2,5 µl de DNA molde, 1x tampão PCR, 1,5 mM de MgCl₂, 0,1 mM de cada dNTP e 2 U da enzima Platinum Taq DNA polimerase, além dos

oligonucleotídeos como listados na seção 2.5.5 (Quadro 1). Os ciclos de amplificação foram realizados em termociclador (Applied Biosystems), partindo-se de uma etapa inicial de desnaturação a 94°C por 3 min, 35 ciclos compostos de desnaturação a 94°C por 40s, anelamento a 55°C por 40s e extensão a 72°C por 40s. Para a extensão final, foi realizado um ciclo de 72°C por 2 min. Todos os reagentes utilizados em ambos os protocolos foram obtidos da Invitrogen, Life Tech.

As seguintes amostras foram utilizadas como controles deste PCR *multiplex*: *E. faecium* SF11770 para os genes *aph(2'')*-Ib, *aph(3')*-IIIa e *ant(4')*-Ia; *E. gallinarum* SF9117 para o gene *aph(2'')*-Ic; *E. casseliflavus* SF11300 para o gene *aph(2'')*-Id e *E. casseliflavus* MBO4541E para o gene *aac(6')*-Ie-*aph(2'')*-Ia.

2.5.3 Detecção dos Genótipos Associados à Resistência aos Glicopeptídeos

A detecção da presença do gene *vanA*, que determina a resistência à vancomicina e teicoplanina, foi feita por PCR *uniplex*, utilizando-se oligonucleotídeos iniciadores descritos por Dutka-Malen et al. (1995). Cada mistura de reação tinha um volume final de 50 µl, sendo constituída de 4 µl do DNA molde, 1x tampão PCR (Ludwig Biotecnologia Ltda), 1,5 mM de MgCl₂ (Ludwig Biotecnologia Ltda), 0,1 mM de cada dNTP (Invitrogen) e 0,8 U de Taq DNA polimerase (Ludwig Biotecnologia), além das sequências iniciadoras listadas na seção 2.5.5 (Quadro 1).

Os ciclos de amplificação foram realizados em termociclador (Applied Biosystems), constituídos de uma etapa inicial de desnaturação a 94°C por 2 min, seguida de 35 ciclos a 94°C por 50s para desnaturação, 55°C por 30s para anelamento e 72°C por um minuto para extensão. Para a extensão final foi realizado um ciclo a 72°C por 10 min. Como controle positivo foi utilizada a amostra *E. faecalis* A256, portadora do gene *vanA*, e como controle negativo a amostra *E. faecalis* V583.

2.5.4 Detecção dos Genótipos Associados à Expressão de Fatores de Virulência

Foi avaliada a presença de cinco genes relacionados à expressão dos respectivos fatores envolvidos na virulência de *Enterococcus* spp.: *asaI*, substância de agregação; *cylA*, citolisina; *esp*, proteína de superfície de enterococos; *hyl*, glicosil hidrolase; *gelE*, gelatinase.

O ensaio de PCR *multiplex* seguiu as recomendações de Vankerckhoven et al. (2004). Para tal, foram feitas misturas de reação para um volume final de 25 µL, que constaram de 2,5 µL da suspensão do DNA molde, 1X tampão PCR, 3 mM de MgCl₂, 0,2 mM de cada dNTP, 1,25 U de Platinum Taq polimerase e os pares de iniciadores conforme descrito na seção 2.5.5 (Quadro 1). As condições de amplificação foram as seguintes: 1 ciclo de 94°C por 5 min, 30 ciclos de 94°C por 1 min, para desnaturação do DNA, 56°C por 1 min, para o anelamento dos iniciadores, e 72°C por 1 min, para extensão das novas fitas de DNA, e 1 ciclo de extensão final a 72°C por 10 min, em termociclador. Todos os reagentes utilizados em ambos os protocolos foram obtidos da Invitrogen, Life Tech.

As amostras de referência *E. faecalis* Ac-72 (*cylA*⁺, *gelE*⁺, *asaI*⁺), *E. gallinarum* Ac-19 (*hyl*⁺), *E. faecalis* Fa-52 (*gelE*⁺, *esp*⁺) foram utilizadas como controles positivos da reação.

2.5.5 Eletroforese e Análise dos Produtos Amplificados

Após as etapas em termociclador, os produtos de amplificação foram submetidos à eletroforese, por 1h a 100V, em gel de agarose (LE Agarose, Ludwig Biotecnologia Ltda) a 1,2% preparado em tampão TBE 0,5X (45mM Tris, 45mM ácido bórico, 1 mM EDTA, reagentes obtidos da Bio-Rad Labs.). Os géis foram corados com 0,5 µg/mL de brometo de etídio por 30 min, sendo em seguida visualizados e fotografados com auxílio do sistema MiniBis Pro (DNR Bio Imaging Systems Ltda, Jerusalém, ISR). A eletroforese foi acompanhada de um padrão de pares de base (100 pb Ladder – Ludwig Biotecnologia Ltda) utilizado para estimar o tamanho dos produtos amplificados.

Quadro 1 – Oligonucleotídeos iniciadores utilizados nas reações de PCR *multiplex* para detecção de determinantes de virulência e resistência em *Enterococcus faecium*

Característica	Gene	Sequências iniciadoras (Direção 5'-3')	Quantidade do iniciador (em pmol ^c)	Produto (em pb) ^d	Referências
HLRS ^a	<i>ant(6')-Ia</i>	F: ACTGGCTTAATCAATTTGGG R: GCCTTCCGCCACCTCACCG	1,25	577	Swenson et al., 1995
	<i>ant(9')-Ia</i>	F: GGTTTCAGCAGTAAATGGTGGT R: TGCCACATTCGAGCTAGGGTT	25	476	Mahbub Alam et al., 2005
	<i>ant(9')-Ib</i>	F: CGTCGTATCTGAACCATTGACA R: TCCATAATGGCTCTTCTCAC	25	351	
HLGR ^b	<i>aac(6')-Ie-aph(2'')-Ia</i>	F: CAGAGCCTTGGGAAGATGAAG R: CTCGTGTAATTCATGTCTGGC	25	348	Vakulenko et al., 2003
	<i>aph(2'')-Ib</i>	F: CTTGGACGCTGAGATATATGAGCAC R: TTTGTAGCAATTCAGAAACACCCTT	25	867	
	<i>aph(2'')-Ic</i>	F: CCACAATGATAATGACTCAGTTCCC R: CCACAGCTTCCGATAGCAAGAG	3,5	444	
	<i>aph(2'')-Id</i>	F: GTGGTTTTTACAGGAATGCCATC R: CCTCTTCATACCAATCCATATAACC	5,0	641	
	<i>aph(3')-IIIa</i>	F: GCTAAAATGAGAATATCACCGG R: CTTTAAAAAATCATACAGCTCGCG	3,0	523	
	<i>ant(4')-Ia</i>	F: AAACCTGCTAAATCGGTAGAAGCC R: GGAAAGTTGACCAGACATTACGAACT	2,0	294	
Vancomicina	<i>vanA</i>	F: GGGAAAACGACAATTGC R: GTACAATGCGGCCGTTA	25	732	Dutka-Malen et al., 1995
Virulência	<i>asaI</i>	F: GCACGCTATTACGAACATGA R: TAAGAAAGAACATCACACGA	40	375	Vankerckhoven et al., 2004
	<i>cylA</i>	F: ACTCGGGGATTGATAGGC R: GCTGCTAAAGCTGCGCTT	40	688	
	<i>esp</i>	F: AGATTTTCATCTTTGATTCTTGG R: AATTGATTCTTTAGCATCTGG	40	510	
	<i>gelE</i>	F: ACCCCGTATCATTGGTTT R: ACGCATTGCTTTTCCATC	40	405	
	<i>hyl</i>	F: ACAGAAGAGCTGCAGGAAATG R: GACTGACGTCCAAGTTCCAA	40	276	

Legenda: ^ahigh-level gentamicin resistance (HLGR); ^bhigh-level streptomycin resistance (HLRS); ^cpmol – quantidade de iniciadores em picomol para o volume final da respectiva reação; ^dpb - tamanho do produto em pares de base; F - ordem direta (do inglês, *forward*), R - ordem reversa (do inglês, *reverse*); Nucleotídeos: Adenina (A), Citosina (C), Guanina (G), Timina (T), K - apenas tautômeros de G ou de T.

Fonte: A autora, 2020.

2.6 Sequenciamento do Genoma de *E. faecium*

O sequenciamento do genoma completo foi realizado a partir de um projeto de colaboração com o Dr. Paul Joseph Planet – *University of Pennsylvania*, PA, EUA. No total, 52 amostras foram selecionadas para a realização do sequenciamento a partir de critérios baseados nos resultados obtidos preliminarmente, tais como: perfil fenotípico e genotípico de resistência, perfil genotípico de virulência e sítio de isolamento. Foram selecionadas amostras de todos os anos de estudo, bem como de todas as instituições hospitalares de origem.

Adicionalmente, como o objetivo principal deste trabalho foi elucidar a substituição populacional ocorrida pela troca do tipo MLST, de ST78 para ST412, foram incluídas no estudo 20 amostras bacterianas pertencentes ao ST78, obtidas no período de 2002 a 2008. Além disso, duas amostras pertencentes ao ST412 obtidas no ano de 2008, consideradas como os primeiros registros deste ST no banco de dados de nosso grupo, também foram incluídas nessas análises. Todas essas amostras foram oriundas do trabalho de Faria (2012) e também possuíam prévia caracterização fenotípica e genotípica. Sendo assim, no presente trabalho, 74 amostras de *E. faecium* tiveram seu genoma sequenciado e analisado.

O DNA das amostras foi obtido utilizando-se o kit *Wizard Genomic DNA Purification Kit* (Promega, Madison, WI, USA) de acordo com as instruções do fabricante. Para tal, as amostras foram cultivadas em meio tripton de soja (TSB - do inglês, *trypticase soy broth*) (BD Diagnostics). Após incubação a $36^{\circ}\text{C}\pm 1^{\circ}\text{C}$, por 18h-24h, 1 mL do crescimento obtido para cada amostra foi submetido a centrifugação (4.000 rpm / 2 min), o sobrenadante foi removido e as células ressuspensas em 480 μL de EDTA (50 mM) e 120 μL de uma solução aquosa de liozima (10 mg/mL). Após incubação ($36^{\circ}\text{C}\pm 1^{\circ}\text{C}$ / 1h), as suspensões foram centrifugadas (4.000 rpm / 10 min), os sobrenadantes removidos e as células ressuspensas em 600 μL de Solução de Lise (específica do kit, Promega) e incubadas a 80°C por 5 min. Posteriormente, foram adicionados 3 μL de solução de RNase e em sequência, após homogeneização, 200 μL da Solução para Precipitação de Proteínas (específica do kit, Promega). Após incubação por 5 min no gelo, os tubos foram centrifugados por duas vezes consecutivas: na primeira (4.000 rpm / 3 min) foram adicionados 600 μL de isopropanol ao sobrenadante; e na segunda (4.000 rpm / 2 min), 600 μL de etanol 70%. Após nova etapa de centrifugação, o sobrenadante foi desprezado e o sedimento contendo o DNA foi reidratado com 100 μL de Solução de Reidratação 9 (específica do kit, Promega). Após incubação (65°C / 1 h), o DNA foi estocado a 4°C , até o momento dos ensaios.

O DNA genômico extraído foi quantificado utilizando *Qubit 2.0 Fluorometer* (Invitrogen) e posteriormente 5 µg foram aplicadas em placas de 96 poços (MicroAmp Optical 96-Well Reaction plate; Applied Biosystems). As placas foram enviadas para o serviço de sequenciamento do PennCHOP – *Microbiome Program*, pertencente a *University of Pennsylvania* e o *Children's Hospital of Philadelphia*, com o apoio do Dr. Paul J. Planet. No local, as bibliotecas genômicas foram construídas utilizando-se Illumina Nextera XT DNA (150- bp *paired-end reads*) e então sequenciadas na plataforma de alto desempenho Illumina HiSeq 2500 (Illumina, San Diego, CA, EUA).

2.7 Análises de Bioinformática Aplicadas ao Estudo do Genoma de *E. faecium*

Para a análise do genoma completo das amostras de *E. faecium*, um *pipeline* com diversas ferramentas computacionais foi desenhado, conforme representado na seção 2.7.6 (Figura 1). O fluxo de trabalho foi dividido em diferentes etapas e consistiu primeiramente na avaliação da qualidade das sequências obtidas e processo de “trimagem”. Posteriormente, para a elucidação da estrutura populacional e dos processos evolutivos envolvendo a espécie, foram empregados diferentes métodos de análise com três enfoques distintos: o estudo do conteúdo genético e a análise do pangenoma; o estudo dos processos de mutação; e o estudo dos eventos de recombinação. Análises filogenômicas também foram empregadas para ajudar na elucidação das relações de ancestralidade entre as amostras. Por fim, por conta de alguns resultados obtidos, foi aplicado um viés proteômico, com a utilização de modelagem computacional para demonstração de estruturas proteicas.

2.7.1 Avaliação de Qualidade e “Trimagem” das Reads

Os arquivos brutos gerados do sequenciamento, chamados de *reads*, foram primeiramente analisados através do programa FastQC versão 0.11.8 (ANDREW, 2010), para avaliação do total de sequências geradas, tamanho e qualidade dos fragmentos gerados e conteúdo de adaptadores. A partir desses valores, foi possível então calcular a cobertura do sequenciamento, multiplicando-se o tamanho das *reads* (L) pelo número de *reads* (n) obtidas

e divididas pelo tamanho esperado para o genoma (G). Posteriormente, essas sequências brutas foram então “trimadas” com o software Trim Galore! versão 0.6.2 (KRUEGER, 2015), utilizando argumentos padrão, que removem das *reads* os adaptadores, assim como as extremidades de baixa qualidade.

2.7.2 Estudo do Conteúdo Genético e Análise do Pangenoma

A avaliação do conteúdo genético teve como objetivo identificar todos os genes que compõem o genoma das 74 amostras de *E. faecium*, com destaque, principalmente, para a presença de genes associados à resistência aos antimicrobianos (resistoma), a fatores de virulência (viruloma) e de estruturas associadas a elementos genéticos móveis (mobiloma), como, IS, plasmídeos e profagos. Todos os programas que serão descritos a seguir foram executados no modo padrão.

Para esta etapa, após o processo de “trimagem” das *reads*, as mesmas foram submetidas ao processo de montagem “*De novo*” através da plataforma online PATRIC 3.6.3 (WATTAM et al., 2017) utilizando-se a estratégia denominada “*auto*”, que consiste na montagem dos fragmentos através de três algoritmos diferentes: Velvet, Spades e IDBA, com posterior comparação e pontuação entre os resultados obtidos e indicação do melhor processo de montagem. O arquivo gerado, indicativo da melhor montagem, denominado *contigs.fasta*, foi utilizado em todas as análises posteriores. A avaliação da montagem do genoma foi feita através do software Quast v2 (GUREVICH et al., 2013), para a obtenção das métricas como o número de *contigs*, tamanho do genoma, N50 e conteúdo G+C.

Para a anotação do resistoma, foram utilizadas duas plataformas online diferentes, Resfinder versão 3.2 (ZANKARI et al., 2012) e *Comprehensive Antibiotic Resistance Database* (CARD) versão 3.0.8 (ALCOCK et al., 2020). A anotação dos determinantes de virulência foi realizada pelas plataformas VirulenceFinder versão 2.0 (JOENSEN et al., 2014) e *Virulence Factor Database* (VFDB) versão 2019 (LIU et al., 2019). Para anotação dos elementos que compõem o mobiloma, foram utilizadas as seguintes plataformas online: ISfinder (SIGUIER et al., 2006), para identificação de sequências de inserção, PlasmidFinder versão 2.1 (CARATTOLI et al., 2014), para detecção de plasmídeos, e *phage search tool enhanced release* (PHASTER) (ARNDT et al., 2016), para identificação de profagos. Para a validação dos resultados da detecção de IS, só foram anotadas aquelas sequências que

possuíam um *e-value* de 0.0, apresentaram similaridade igual ou superior a 95% e mais de 100 pb. Em relação à detecção da presença de profagos, somente aqueles que obtiveram uma pontuação superior a 90, sendo considerados pelo programa como profagos intactos, foram anotados.

As amostras sequenciadas tiveram a identificação dos respectivos STs através da tipificação de sequências em múltiplos loci, conforme painel proposto por Homan et al. (2002). As análises para os sete genes de manutenção *adk* (adenilato quinase), *atpA* (subunidade α da ATP sintetase), *ddl* (D-alanina-D-alanina ligase), *gdh* (Glicose 6-fosfato desidrogenase), *gyd* (gliceraldeído 3-fosfato desidrogenase), *pstS* (proteína transportadora de cassete de ligação de ATP fosfato) e *purK* (fosforribosilaminoimidazol carboxilase) foram realizadas por meio da plataforma online MLST versão 2.0 (LARSEN et al., 2012), pertencente ao servidor do CGE (*Center for Genomic Epidemiology*), com posterior confirmação utilizando-se o software mlst (SEEMANN, 2015).

Para a análise do pangenoma, as sequências foram primeiramente submetidas à anotação do genoma completo através do software Prokka versão 1.13.17 (SEEMANN, 2014). A partir dos arquivos *.gff gerados na anotação, o software Roary versão 3.11.2 (ANDREW et al., 2015) foi utilizado para estimar o pangenoma e o genoma central das 74 amostras (*core* genoma total). Além disso, as amostras foram divididas em cinco grupos distintos - ST78, ST412, ST896, ST963 e “Outros STs” (que incluíram os demais STs encontrados no estudo) - que, posteriormente, foram analisados separadamente pelo programa Roary. Foram denominados como genes centrais ou *core* genes de cada grupo ST, aqueles presentes em 95% ou mais amostras do respectivo grupo. Os genes presentes em 15% ou menos amostras de cada grupo foram considerados genes acessórios.

Com a finalidade de identificar genes exclusivos a cada um dos cinco grupos de amostras, foi utilizado o software Scoary versão 1.6.16 (BRYNILDSRUD et al., 2016) comparando-se o arquivo gerado por Roary, que continha uma lista dos genes identificados nos 74 genomas, com uma planilha binária que associou cada uma das amostras ao seu respectivo grupo. Foram considerados genes exclusivos de um grupo, aqueles que simultaneamente atenderam aos seguintes parâmetros: (i) valor de sensibilidade entre 80% e 100%; (ii) valor de especificidade de 90% a 100%; (iii) ser identificado exclusivamente em um único ST dentre os de interesse (ST78, ST412, ST896, ST963). Porém, não foi considerado como um fator de exclusão quando um determinado gene associado a um dos STs de interesse foi também identificado entre amostras pertencentes ao grupo nomeado como “Outros STs”.

Todos os genes preditos como exclusivos nas análises anteriores tiveram sua sequência extraída do arquivo de anotação para uma confirmação manual por meio da ferramenta de BLAST – *basic local alignment search tool* (ALTSCHUL et al., 1990), hospedado no NCBI (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, 2020) e o banco de dados do UniProt (THE UNIPROT CONSORTIUM, 2019). As categorias funcionais desses genes foram pesquisadas em bancos de dados como EggNOG versão 5.0.0 (HUERTA-CEPAS et al., 2019), e QuickGO versão 2020-04-04 (BINNS et al., 2009).

Foram, também, pesquisados os genes associados à resistência a metais pesados. Para tal, os genes *cadA* (resistência ao cádmio), *tcrB* e *cueO* (resistência ao cobre), e *merA* (resistência ao mercúrio) foram pesquisados de forma manual nos arquivos de anotação gerados pelo programa Prokka (arquivo.gbk), utilizando-se a ferramenta de busca simples para cada um dos 74 genomas. Os números de acesso para as sequências de nucleotídeos utilizadas como referência foram as seguintes: WP_010815218.1 (*cadA*), HM565232.1 (*cueO*), WP_049141803.1 (*merA*) e AY048044.2 (*tcrB*).

A detecção de sistema associado à imunidade bacteriana contra aquisição de material genético externo, tal como o sistema CRISPR-Cas, foi avaliada a partir da plataforma online CRISPRCasFinder versão 1.1.2 (COUVIN et al., 2018). Foram consideradas regiões associadas à presença de CRISPR apenas aquelas que tiveram valores entre 2-4 para o nível de evidência, conforme descrito no manual do programa. Esses valores atestam a confiabilidade dos resultados e se baseiam no número de espaçadores, na conservação das sequências repetitivas e em uma maior similaridade entre as matrizes CRISPR existentes. Uma análise do sistema R-M foi feita através da ferramenta de busca simples nos arquivos de anotação. Foram pesquisados manualmente em cada um dos genomas as três subunidades (R, M e S) que compõem o sistema. As sequências utilizadas como referência seguiram as recomendações do artigo de Huo et al. (2019).

2.7.3 Estudo das Mutações

Para o estudo das mutações, as 74 amostras foram alinhadas em relação ao genoma de referência *E. faecium* AUS0004 (CP003351). Esta amostra, utilizada como referência, pertence ao ST17, foi isolada no ano de 1998 de uma infecção de corrente sanguínea na Austrália e foi o primeiro genoma completo descrito para a espécie (LAM et al., 2012), sendo

até os dias atuais bastante utilizada como referência em análises de genômica comparativa dentro da espécie (BRODRICK et al., 2016; PINHOLT et al., 2015; VAN HAL et al., 2018)

A etapa de alinhamento foi realizada utilizando-se as *reads* previamente “trimadas” conforme descrito no ITEM 2.7.1 em relação à amostra de referência através do software *burrows-wheeler aligner* (BWA) versão 0.7.17 (LI; DURBIN, 2009). O software SAMtools versão 1.4 foi posteriormente utilizado para a chamada de variantes (LI et al., 2009). Foi empregado um filtro (BCFFILTER) para sítios que obtiveram uma baixa qualidade durante o alinhamento e a detecção de variantes utilizando o comando: -g 12.

Posteriormente, utilizando-se um script interno, gentilmente cedido pelo Dr. Paul Planet, denominado de *snp_matrixBuilder.pl*, os arquivos de chamada de variante (VCF - do inglês, *variant calling format*) gerados na análise anterior foram utilizados para criar uma matriz de polimorfismos de nucleotídeo único (SNPs - do inglês, *single nucleotide polymorphisms*) em relação a amostra de referência. Variantes que possuíam um valor de qualidade menor que 20 foram excluídas da matriz através da utilização do comando -q 20. Através dessa matriz, foi possível obter o número de SNPs de cada uma das 74 amostras em relação à referência e os resultados obtidos também foram posteriormente utilizados para a construção de árvores de filogenia.

Utilizando-se os mesmos arquivos VCF, as variantes encontradas foram anotadas em relação à amostra de referência Aus0004, para predição de seus efeitos nos genes através do software SnpEff versão 4.3T (CINGOLANI et al., 2012a). Para esse estudo, apenas efeitos de alto impacto nas proteínas, como o ganho ou perda de um códon de parada, foram considerados. Para tal, o software SnpSift versão 4.0 foi utilizado para filtrar esses dois tipos de efeito de cada um dos genomas (CINGOLANI et al., 2012b).

Com intuito de observar se os efeitos de alto impacto eram relacionados aos diferentes grupos de amostras bacterianas, uma planilha foi construída no Microsoft Excel®, confrontando todos os efeitos encontrados nos 74 genomas com os diferentes STs. Foram considerados efeitos de alto impacto associados a um determinado grupo de ST aqueles que foram identificados em mais de três amostras do grupo e que não estavam presentes em nenhum dos outros grupos de interesse, exceto em “Outros STs”.

Foram investigadas as mutações associadas a resistência aos antimicrobianos, tais como os β -lactâmicos e quinolonas. Para tal, foram avaliados os genes da PBP5 (*pbp5*), que codifica uma proteína ligadora de penicilina, e cujas mutações estão relacionadas à resistência à ampicilina e à penicilina nesses microrganismos, além dos genes que codificam a DNA girase (*gyrA* e *gyrB* das subunidades A e B da proteína, respectivamente) e a topoisomerase

tipo IV (*parC* e *parE* das subunidades A e B da proteína, respectivamente), cujas mutações são associadas à resistência às quinolonas e fluoroquinolonas

As sequências de nucleotídeos e aminoácidos dos respectivos genes foram extraídas a partir dos arquivos de anotação (*.ffn e *.gbk) gerados através da anotação feita pelo software Prokka. As sequências obtidas dos 74 genomas para cada um dos cinco genes foram então alinhadas junto à sua respectiva referência utilizando-se o aplicativo de domínio público BioEdit – *Biological Sequence Alignment v.7.04* (Ibis Biosciences, Carlsbad, CA, USA). As seguintes referências foram utilizadas: YP_006375753 (Sequência do gene completo) e AB017811.1 (Região QRDR), para o gene *parC*; YP_006374612.1 (Sequência do gene completo) e AF060881.1 (Região QRDR), para o gene *gyrA*; YP_006374611.1 e YP_006375752.1 para *gyrB* e *parE*, respectivamente, além da sequência X84860.1 para o gene da *pbp5*.

Para avaliar se as variações naturais dentro das sequências de proteínas (variações de alelos) poderiam ser consideradas como novidade genética (novos polimorfismos), foi utilizada a ferramenta WhatsGNU versão 1.0 (MOUSTAFA; PLANET, 2020). Para tal, é atribuída uma pontuação, ou score, GNU (do inglês, *gene novelty unity*) para cada uma das proteínas, a partir do número de combinações exatas encontradas dentro de um banco de dados contendo vários genomas. A pontuação GNU para cada proteína é inversamente proporcional à atribuição de novidade genética. Assim, uma dada sequência proteica com pontuação GNU elevada indica a alta frequência desse alelo no banco de dados sendo, provavelmente, uma proteína conservada. Por outro lado, sequências proteicas com pontuação GNU baixa são raras no banco de dados e aquelas com pontuação zero são as que não apresentam correspondência, sendo estas últimas consideradas novidades genéticas.

Para tal, foi obtida uma planilha, diretamente do NCBI, contendo o total de genomas da espécie (1.928 genomas) que estavam depositados até o momento da realização desta etapa do trabalho (março de 2020). Para as análises, foram excluídos desse número total aqueles genomas que estavam montados somente até o nível de *contigs*, que não possuíam anotação do genoma pelo GenBank ou que possuíam um tamanho total ou percentual de conteúdo G+C diferente do esperado para a espécie, sugerindo uma possível contaminação. Sendo assim, o banco de dados final, que foi construído e utilizado para essas análises, constou de 823 genomas diferentes de *E. faecium*.

Utilizando-se um script dentro da própria ferramenta WhatsGNU, foi realizado o download diretamente do GenBank da sequência de aminoácidos dos 823 genomas selecionados. Para uma abordagem mais completa dos dados, o ST dos 823 genomas foi

definido através da metodologia descrita no item 2.7.2. Os resultados então foram incorporados ao banco de dados final que foi posteriormente comprimido utilizando um script dentro da própria ferramenta.

Todos os 74 genomas presentes em nosso estudo tiveram suas proteínas avaliadas em relação ao banco de dados construído a partir do arquivo *.faa de anotação gerado pelo software Prokka. Uma planilha foi gerada com os resultados de cada uma das amostras. As análises foram realizadas filtrando apenas aquelas proteínas que obtiveram um GNU score igual a zero. Foram consideradas unidades de novidades genéticas exclusivas de um respectivo grupo, as proteínas com score zero presentes em mais de uma amostra do grupo.

Ainda utilizando a ferramenta WhatsGNU, foram também pesquisados os 10 genomas no bando de dados que apresentaram o maior número de *hits* genéticos com as proteínas presentes em cada uma das amostras de *E. faecium* incluídas neste estudo.

2.7.4 Estudo dos Eventos de Recombinação

A identificação de eventos de recombinação homóloga foi realizada através do software ClonalFrameML (DIDELOT; WILSON, 2015). Para tal, as *reads* “trimadas” dos 74 genomas incluídos neste estudo foram alinhadas e tiveram suas variações (SNPs) detectadas em relação à referência AUS0004 (CP003351) através da ferramenta Snippy (SEEMANN, 2015). Um arquivo de alinhamento denominado *core.snps.aln* foi gerado. Este arquivo incluiu os SNPs que estavam presentes em todas as amostras em relação à referência.

A partir do arquivo gerado, foi feita a reconstrução de filogenia utilizando-se o método de Máxima Verossimilhança no programa RAxML versão 8.2.4 (STAMATAKIS, 2014), com valor de *bootstrap* estipulado em 100 replicações e modelo evolutivo GTR+G. A árvore filogenética gerada pelo programa e o arquivo de alinhamento *core.snps.aln* serviram como entradas para a análise pelo ClonalFrameML. A análise foi replicada 100 vezes através do comando `emsim`. Os resultados gráficos foram produzidos por um script em linguagem R, presente no repositório do próprio programa. As árvores de filogenia, geradas antes e depois da demarcação dos eventos de recombinação, foram visualizadas pelo programa FigTree versão 1.4.4 (RAMBAUT, 2018).

2.7.5 Análises Filogenômicas

A reconstrução das relações filogenéticas foi realizada através de duas abordagens diferentes: Máxima verossimilhança e análise bayesiana (ver seção 2.7.6 – Figura 1). Utilizando a matriz de SNPs construída conforme descrição no ITEM 2.7.3, um script interno fornecido gentilmente pelo Dr. Paul Planet, denominado de *snp_invariant_screen.pl*, foi utilizado para corrigir os problemas com sítios invariantes, uma vez que o arquivo de matriz só possuía os sítios que variavam. As análises foram então realizadas com o programa RAxML versão 8.2.4 (STAMATAKIS, 2014), com valor de *bootstrap* estipulado em 100 replicações e modelo evolutivo GTR+G. As árvores foram construídas pelo método clássico de enraizamento pelo ponto médio (do inglês, *midpoint rooting*) e visualizadas na plataforma online iTOL versão 5 (LETUNIC; BORK, 2019), onde foi possível compilar os dados obtidos em diversas análises.

Para os casos pontuais nos quais as árvores produzidas apresentaram resultados discordantes do esperado (por exemplo, amostras pertencentes a STs diferentes compondo o mesmo ramo), foram utilizadas análises adicionais pelo cálculo de ANI (do inglês, *average nucleotide identity*). Para tal, foi utilizado o programa OrthoANI versão 0.90 (LEE et al., 2015), e, posteriormente, uma árvore foi construída pelo algoritmo *unweighted pair group method with arithmetic mean* (UPGMA), utilizando os valores percentuais obtidos de ANI.

Através da abordagem bayesiana para reconstrução da filogenia, foi possível estimar os tempos de divergência, pelo o emprego do software BEAST versão 2.5.2 (BOUCKAERT et al., 2019). Como a inferência bayesiana não leva em conta os eventos de recombinação homóloga, uma vez que, em geral, eles tendem a influenciar a reconstrução de relações de filogenia (DIDELOT; MAIDEN, 2010), um script em linguagem Python, denominado *cfml-maskrc.py* (KWONG, 2018), foi utilizado para mascarar as regiões de recombinação obtidas através da análise pela ferramenta ClonalFrameML (Item 2.8.4), no arquivo de alinhamento *core.snps.aln* gerado pelo programa Snippy (Item 2.8.4).

Este novo alinhamento, com as regiões de recombinação mascaradas, foi usado no programa BEAUti (distribuído junto com pacote BEAST) para o ajuste dos parâmetros e comandos necessários para à análise bayesiana. As datas de isolamento das amostras foram utilizadas como informação para calibrar as datas de divergência. O modelo de substituição escolhido foi HKY, com uma taxa de substituição estimada de 1.0, um modelo de relógio molecular estrito e com taxa estimada também de 1.0. O modelo *a priori* escolhido para

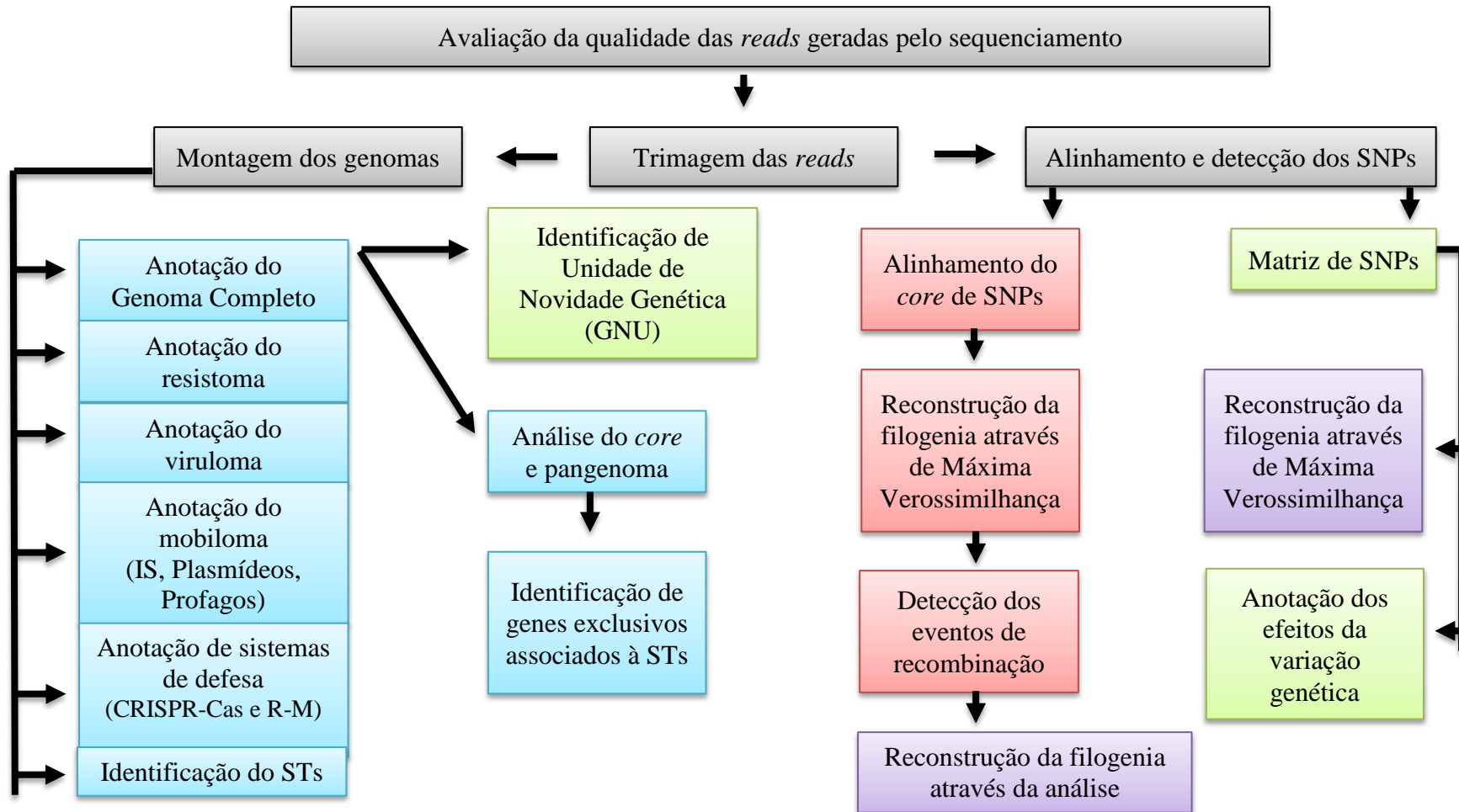
determinar a distribuição das árvores e dos comprimentos de ramos foi o coalescente em populações com tamanho constante (do inglês, *coalescent constant population*).

Análises matemáticas através do emprego do algoritmo de Monte Carlo via Cadeias de Markov (MCMC, do inglês *Markov chain Monte Carlo*) foram executadas por 100 milhões de gerações, com árvores e parâmetros amostrados a cada 1.000 gerações. Do total de árvores amostradas, 10.000 foram descartadas como *burn-in* (10%). Os arquivos de saída (*.log) foram então importados no software Tracer versão 1.7.1 (RAMBAUT et al., 2018), para avaliar se haviam atingido a convergência da distribuição posterior, através da análise dos valores de ESS (*effective sample size*) > 200. Posteriormente, uma árvore filogenética de consenso foi gerada pelo critério de máxima credibilidade de clados (MCC, do inglês *maximum clade credibility*) no software TreeAnnotator versão 2.5.2. As árvores foram visualizadas através do programa FigTree versão 1.4.4 (RAMBAUT, 2018)

2.7.6 Evolução Molecular

Para avaliar a evolução molecular, foi aplicado o teste D de Tajima (TAJIMA, 1989), para os genes *ddl*, *purK*, *pstS*, que compõem o esquema de MLST, e para os genes de resistência aos antimicrobianos *parC*, *gyrA* e *pbp5*, nas 61 amostras de *E. faecium* pertencentes aos ST78, ST412, ST963 e ST896. Este é um teste estatístico aplicado aos estudos de genética populacional, considerando-se uma população constante em evolução neutra (teste de neutralidade). As análises foram realizadas através do emprego do programa DnaSP v.6 (ROZAS et al., 2017). Valores de $p < 0,05$ foram considerados estatisticamente significativos.

Figura 1 – Fluxo de trabalho realizado para a análise do genoma de 74 amostras de *Enterococcus faecium*



Legenda: Azul - estudo do conteúdo genético; Verde - estudo das mutações; Lilás- análises filogenômicas; Vermelho - estudo dos eventos de recombinação; *Insertion Sequence* (IS); *Clustered regularly interspaced short palindromic repeats* (CRISPR-Cas); Restrição/Modificação (R-M); *Sequence types* (STs); *Gene novelty unity* (GNU); *Single nucleotide polymorphism* (SNPs).

Fonte: A autora, 2020.

2.7.7 Estudo das Incongruências entre Fenótipo e Genótipo

As amostras que apresentaram resultados incongruentes entre os testes fenotípicos (categoria sensível nos testes de susceptibilidade por disco-difusão e determinação da CIM) e genotípicos (presença do gene por PCR convencional e sequenciamento do genoma) para o antimicrobiano estreptomicina, após pelo menos duas repetições em dias distintos, foram selecionadas para uma avaliação mais detalhada. As sequências de nucleotídeos e aminoácidos relacionadas a presença do gene *ant(6')-Ia* foram obtidas e alinhadas com uma sequência de referência (AF330699.1) utilizando o aplicativo BioEdit.

Para a complementação dos dados obtidos com as amostras que foram submetidas ao sequenciamento do genoma (N=74), foi também investigada a presença do gene *ant(6')-Ia*, por PCR convencional (conforme item 2.5.2), em todas as demais amostras incluídas neste estudo que foram sensíveis à estreptomicina e não foram selecionadas para sequenciamento do genoma completo (N=124). Adicionalmente, foi feita uma avaliação da sequência de *primers* utilizada na técnica de PCR através da ferramenta Primer-BLAST, disponível online na plataforma do NCBI.

A predição da estrutura 3D da proteína codificada pelo gene *ant(6')-Ia* foi realizada através da sequência de com a utilização do servidor online PHYRE2 (KELLEY et al., 2015). Os sítios de ligação da proteína também foram preditos utilizando-se outra plataforma online chamada 3DLigandSite (WASS; KELLEY; STERNBERG, 2010). A exibição gráfica da estrutura foi obtida através do software EzMol 1.3 (REYNOLDS; ISLAM; STERNBERG, 2018).

2.8 Análises Estatísticas

A comparação entre as métricas obtidas para tamanho do genoma, percentual de conteúdo G+C, quantidade de SNPs, número de sequências codificantes (CDS - do inglês, *coding sequence*) anotados, quantidade de determinantes de virulência e resistência encontrados, número de ISSs, profagos e plasmídeos detectados nos diferentes grupos de STs foi realizada através da análise de variância monocaudal (*One-way ANOVA*) com posteriores comparações pelo teste de Welch, uma vez que foi assumido que os desvios-padrão para cada

grupo não eram semelhantes. Posteriormente, os grupos foram analisados pareadamente para cada uma das métricas através de teste t não pareado. A correlação entre a não suscetibilidade a 14 antimicrobianos distintos e a presença/ausência de resistência à vancomicina foi feita através da correlação de Pearson.

As análises foram realizadas com o aplicativo GraphPad Prism versão 8.0 (GraphPad Software Inc.). Foram considerados significativos valores de $p < 0,05$. Para as correlações de Pearson, valores entre 0,0 e 0,3 (0,0 e -0,3) indicavam correlação não significativa. Valores entre 0,3 e 0,5 (-0,3 e -0,5) indicavam baixa correlação positiva (negativa), entre 0,5 e 0,7 (-0,5 e -0,7), correlação positiva (negativa) moderada e, a partir de 0,7 (-0,7) alta correlação positiva (negativa) (MUKAKA, 2012).

3 RESULTADOS

3.1 Caracterização da Amostragem

Todas as 205 amostras incluídas no estudo, previamente caracterizadas por testes fisiológicos convencionais (TEIXEIRA et al. 2019) como pertencentes à espécie *E. faecium*, tiveram a identificação confirmada através da técnica de MALDI-TOF MS. Os escores de identificação foram >2.300, que garantem a classificação em gênero e espécie, de acordo com as recomendações do fabricante.

Em relação às fontes de isolamento, 42% (86/205) foram representativas de sítios de infecção, enquanto que 55,6% (114/205) foram provenientes de *swabs* e/ou fezes, representativos de culturas de vigilância. Apenas 2,4% (5/205) das amostras não tiveram sua origem de isolamento informada. Entre os espécimes clínicos representativos de infecção, os que mais se destacaram foram sangue (21,5%) e urina (11,7%). A Tabela 1 apresenta a distribuição das amostras nos oito anos de estudo de acordo com o espécime clínico de origem.

Tabela 1 – Distribuição das 205 amostras de *Enterococcus faecium* incluídas no presente estudo, de acordo com o espécime clínico de origem e ano de isolamento

Espécime Clínico	Ano/Nº de amostras								Total Nº (%)
	2009	2010	2011	2012	2013	2014	2015	2016	
Sangue	5	3	2	6	3	8	9	8	44 (21,5)
Urina	4	2	2	1	8	2	1	4	24 (11,7)
Secreções Abdominais	2	-	-	-	-	2	2	2	8 (3,9)
Fragmentos de Tecido	-	-	-	-	1	2	2	1	6 (2,9)
Cateteres e Drenos	-	1	-	-	-	1	-	-	2 (1)
Abscessos	-	-	1	-	-	-	-	-	1 (0,5)
Secreções Respiratórias	-	-	-	-	-	1	-	-	1 (0,5)
<i>Swabs</i> /Fezes	19	20	5	7	18	14	16	15	114 (55,6)
Não informado	-	4	1	-	-	-	-	-	5 (2,4)
Total/Nº (%)	30 (14,6)	30 (14,6)	11 (5,4)	14 (7)	30 (15)	30 (14,6)	30 (14,6)	30 (14,6)	205 (100)

Legenda: Não informado - não foi possível obter informação quanto ao espécime clínico de origem; - (nenhuma amostra).

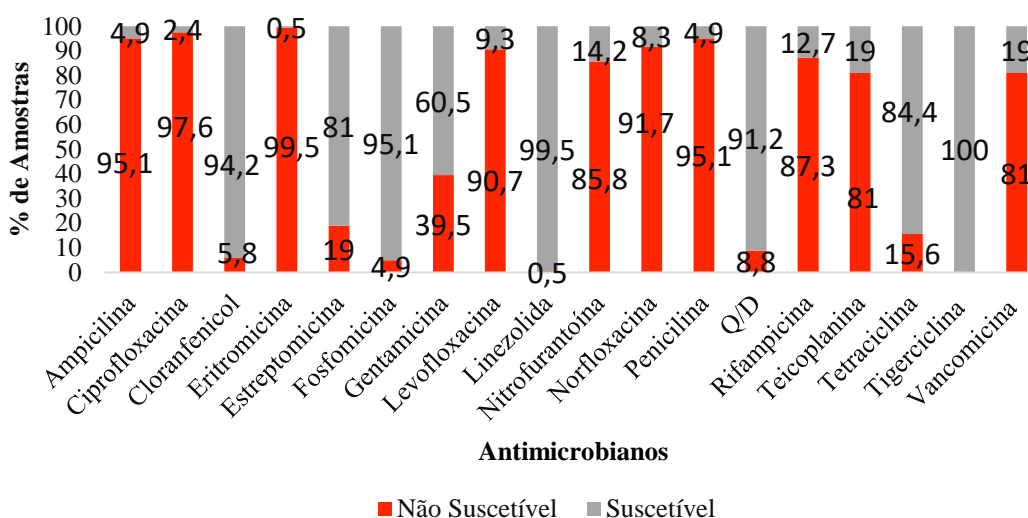
Fonte: A autora, 2020.

3.2 Suscetibilidade aos Antimicrobianos

Os resultados dos testes de suscetibilidade aos 18 antimicrobianos, considerando os critérios de interpretação do CLINICAL LABORATORY STANDARDS INSTITUTE (2019) e EUROPEAN COMMITTEE ON ANTIMICROBIAL SUSCEPTIBILITY TESTING (2019), estão demonstrados no Gráfico 1. Foram consideradas como amostras não suscetíveis aquelas classificadas como resistentes ou intermediárias, de acordo com os critérios de interpretação mencionados anteriormente.

Percentuais elevados de não suscetibilidade foram observados para ampicilina (95,1% - 195/205), ciprofloxacina (97,6% - 200/205), eritromicina (99,5% - 204/205), levofloxacina (90,7% - 186/205), nitrofurantoína (85,8% - 176/205), norfloxacina (91,7% - 188/205), penicilina (95,1% - 195/205), rifampicina (87,3% - 179/205), teicoplanina (81% - 166/205) e vancomicina (81% - 166/205). Por outro lado, taxas elevadas de sensibilidade foram encontradas para cloranfenicol (94,2% - 193/205), fosfomicina (95,1% - 195/205), linezolida (99,5% - 204/205), quinupristina/dalfopristina (91,2% - 187/205) e tetraciclina (84,4% - 173/205). Todas as amostras testadas foram suscetíveis à tigeciclina. A taxa de não suscetibilidade a níveis elevados de aminoglicosídeos foi de 19% (39/205) para estreptomicina e 39,5% (81/205) para gentamicina.

Gráfico 1 - Distribuição de 205 amostras de *Enterococcus faecium* de acordo com a suscetibilidade à 18 antimicrobianos



Legenda: Quinupristina/Dalfopristina (Q/D).

Fonte: A autora, 2020.

Foi observado que, durante quase que a totalidade do período de oito anos avaliado, houve a manutenção do percentual de amostras não suscetíveis à ampicilina, ciprofloxacina, eritromicina, levofloxacina, norfloxacina, penicilina, rifampicina, teicoplanina e vancomicina, com valores acima de 80%. Apenas nos anos de 2011 e 2012, reportados anteriormente por contemplarem um menor número de amostras, ocorreram variações nos perfis de não susceptibilidade (Tabela 2).

Em relação à vancomicina e a teicoplanina, a frequência de amostras não suscetíveis nesses dois anos foi de 36,4% (4/11) e 42,8% (6/14), respectivamente. Entre as quinolonas, o percentual de amostras não suscetíveis à norfloxacina foi de 45,4% (5/11) em 2011. Para a classe dos β -lactâmicos, 63,6% (7/11) das amostras foram não suscetíveis à penicilina em 2011. Neste mesmo ano, foi possível observar que 54,5% (6/11) das amostras foram não suscetíveis a níveis elevados de gentamicina.

No ano de 2015, o percentual de amostras não suscetíveis ao cloranfenicol e a quinupristina/dalfopristina foi de 20% (6/30) e 16,7% (5/30), respectivamente. O mesmo percentual de amostras não suscetíveis à quinupristina/dalfopristina foi observado em 2009. No ano de 2016, 33,3% (10/30) das amostras apresentaram não susceptibilidade a estreptomicina, enquanto 26,7% (8/30) foram não suscetíveis a tetraciclina. Embora não indicado na Tabela 2, a única amostra não suscetível à linezolida foi identificada em 2015.

A correlação entre amostras caracterizadas como resistentes à vancomicina (VRE) ou sensíveis (VSE) com a perda de susceptibilidade a outros antimicrobianos está demonstrada no Gráfico 2. Dos 14 antimicrobianos avaliados, as três quinolonas e os dois β -lactâmicos testados foram os únicos que obtiveram coeficiente de correlação superiores a 0,3, indicando uma fraca correlação positiva entre a presença de resistência à vancomicina com a não susceptibilidade a estas classes. Entre os β -lactâmicos, o coeficiente de correlação obtido para a ampicilina foi de 0,46, enquanto que para a penicilina foi de 0,40. As amostras resistentes à vancomicina apresentaram um percentual de não susceptibilidade à ampicilina e à penicilina de 100% (166/166) e 99% (165/166), respectivamente. Já entre as amostras sensíveis à vancomicina, o percentual de não susceptibilidade a estas drogas foi de 74% (29/39) e 77% (30/39).

Tabela 2 – Distribuição do percentual de não suscetibilidade aos antimicrobianos de amostras de *Enterococcus faecium* no período de 2009 a 2016

Antimicrobianos	Ano / N° de amostras (%)								Total
	2009	2010	2011	2012	2013	2014	2015	2016	
Ampicilina	29 (96,7)	28 (93,3)	9 (81,8)	13 (92,8)	29 (96,7)	28 (93,3)	30 (100)	29 (96,7)	195 (95,1)
Penicilina	29 (96,7)	28 (93,3)	7 (63,6)	14 (100)	29 (96,7)	28 (93,3)	30 (100)	30 (100)	195 (95,1)
Ciprofloxacina	30 (100)	30 (100)	10 (90,9)	12 (85,7)	29 (96,7)	29 (96,7)	30 (100)	30 (100)	200 (97,6)
Levofloxacina	26 (86,7)	26 (86,7)	10 (90,9)	12 (85,7)	27 (90)	27 (90)	29 (96,7)	29 (96,7)	186 (90,7)
Norfloxacina	27 (90)	29 (96,7)	5 (45,4)	12 (85,7)	28 (93,3)	28 (93,3)	30 (100)	29 (96,7)	188 (91,7)
Estreptomicina	7 (23,3)	6 (20)	2 (18,2)	3 (21,4)	5 (16,7)	0	6 (20)	10 (33,3)	39 (19)
Gentamicina	12 (40)	4 (13,3)	6 (54,5)	7 (50)	16 (53,3)	11 (36,7)	13 (43,3)	12 (40)	81 (39,5)
Teicoplanina	26 (86,7)	26 (86,7)	4 (36,4)	6 (42,8)	25 (83,3)	28 (93,3)	24 (80)	27 (90)	166 (81)
Vancomicina	26 (86,7)	26 (86,7)	4 (36,4)	6 (42,8)	25 (83,3)	28 (93,3)	24 (80)	27 (90)	166 (81)
Eritromicina	30 (100)	30 (100)	10 (90,9)	14 (100)	30 (100)	30 (100)	30 (100)	30 (100)	204 (99,5)
Q/D	5 (16,7)	2 (6,7)	0	1 (7,1)	1 (3,3)	0	5 (16,7)	4 (13,7)	18 (8,8)
Cloranfenicol	2 (6,7)	0	1 (9,1)	0	0	0	6 (20)	3 (10)	12 (5,8)
Fosfomicina	1 (3,3)	0	0	0	3 (10)	1 (3,3)	1 (3,3)	4 (13,3)	10 (4,9)
Nitrofurantóina	23 (76,7)	29 (96,7)	11 (100)	14 (100)	30 (100)	19 (63,3)	22 (73,3)	28 (93,3)	176 (85,8)
Rifampicina	27 (90)	30 (100)	10 (90,9)	13 (92,8)	22 (73,3)	25 (83,3)	26 (86,7)	26 (86,7)	179 (87,3)
Tetraciclina	7 (23,3)	2 (6,7)	2 (18,2)	1 (7,1)	2 (6,6)	4 (13,3)	6 (20)	8 (26,7)	32 (15,6)

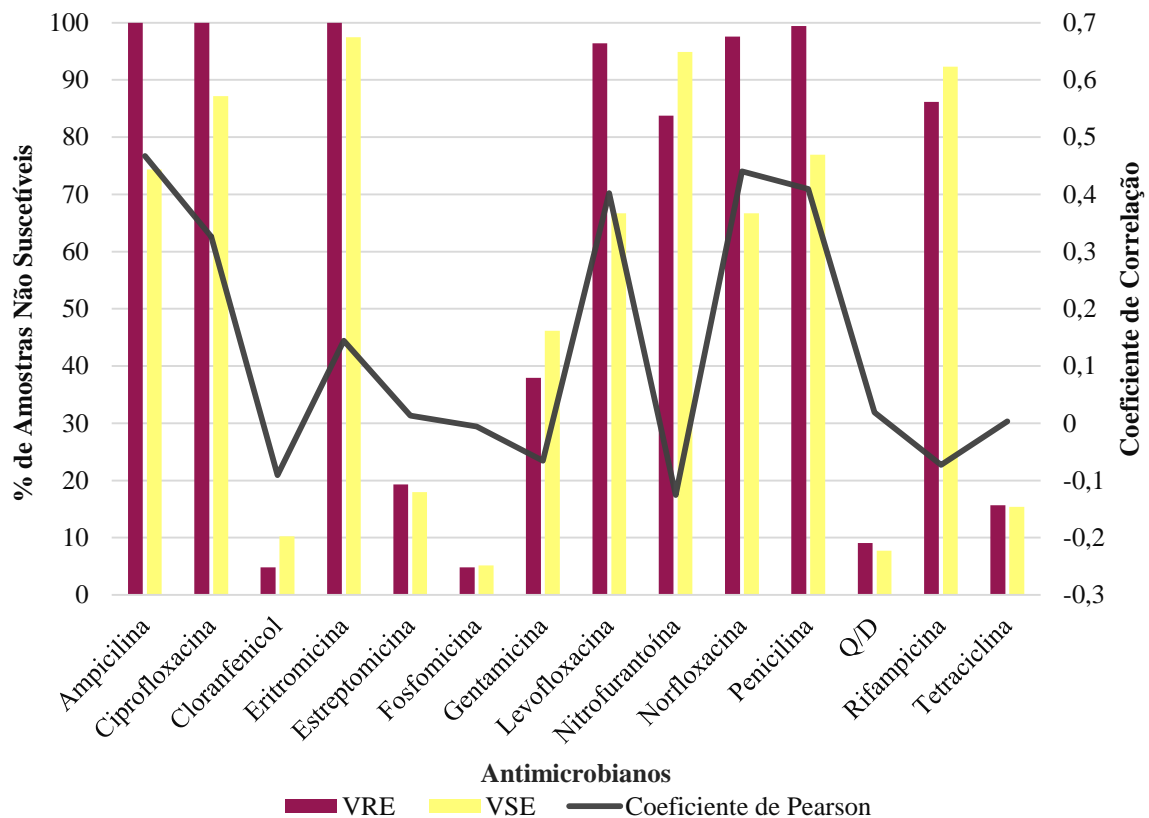
Legenda: Quinupristina/Dalfopristina (Q/D).

Fonte: A autora, 2020.

Entre as quinolonas, os coeficientes de correlação obtidos para ciprofloxacina, levofloxacina e norfloxacina foram 0,32, 0,40 e 0,44, respectivamente. Para ciprofloxacina, o percentual de amostras VRE não suscetíveis a esta droga foi de 100% (166/166), enquanto que para amostras VSE foi de 87% (34/39). Para essas amostras, o percentual de não suscetibilidade para levofloxacina e norfloxacina foi de 67% (26/39). Em contrapartida, as amostras VRE obtiveram percentuais de 96% (160/166) e 97,6% (162/166), respectivamente.

Para todos os outros antimicrobianos testados, os valores de correlação obtidos foram inferiores a 0,3, sendo considerados insignificantes e não permitindo associar a presença da resistência à vancomicina com a não suscetibilidade a essas drogas (Gráfico 2).

Gráfico 2 - Correlação entre as taxas de não suscetibilidade a 14 antimicrobianos diferentes entre amostras não suscetíveis (VRE) e suscetíveis (VSE) à vancomicina

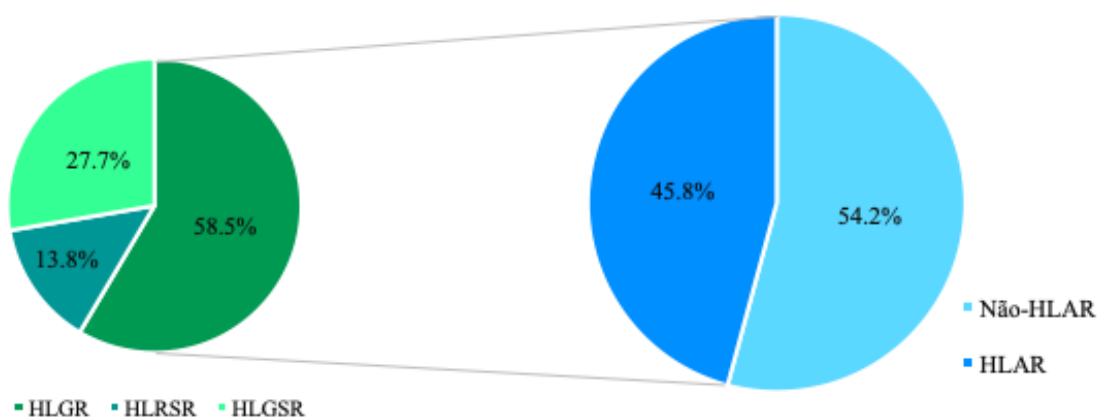


Legenda: Quinupristina/Dalfopristina (Q/D).

Fonte: A autora, 2020.

Em relação ao fenótipo HLAR, 45,8% (94/205) das amostras apresentaram este fenótipo. Dentre estas, o fenótipo HLGR (do inglês, *high-level gentamicin resistance*) foi predominante, representando 58,5% (55/94) das amostras (Gráfico 3). O fenótipo HLGSR (do inglês, *high-level gentamicin and streptomycin resistance*) e HLSR (do inglês, *high-level streptomycin resistance*) representaram 27,7% (26/94) e 13,8% (13/94) das amostras, respectivamente.

Gráfico 3 – Distribuição dos fenótipos de não suscetibilidade a níveis elevados de aminoglicosídeos em amostras de *Enterococcus faecium*



Legenda: *High-level gentamicin resistance* (HLGR); *High-level streptomycin resistance* (HLSR); *High-level gentamicin and streptomycin resistance* (HLGSR).

Fonte: A autora, 2020.

Em relação às quinolonas, um dos alvos de estudo do trabalho, 87,4% (180/205) das amostras foram não suscetíveis aos três antimicrobianos testados dentro da classe. Apenas 5,8% (12/205) foram não suscetíveis a dois antimicrobianos e 3,4% (7/205) somente a um antimicrobiano. Na Tabela 3, estão listados os fenótipos de suscetibilidade às quinolonas encontrados neste estudo.

Um total de 91,7% (188/205) das amostras foi considerada MDR, ou seja, não foram suscetíveis a pelo menos um agente dentro de três categorias de antimicrobianos diferentes. Nenhuma amostra foi caracterizada como PDR (do inglês, *pandrug-resistant*), isto é, os resultados deveriam apontar a não suscetibilidade a todos os antimicrobianos de todas as categorias testadas. Dezesete amostras não puderam ser classificadas seguindo os critérios mencionados, pois foram não suscetíveis a categorias de antimicrobianos desconsideradas pela classificação utilizada como referência, descrita por Magiorakos et al. (2012).

Tabela 3 - Perfis de suscetibilidade às quinolonas de 205 amostras de *Enterococcus faecium*

Antimicrobiano			Nº de amostras	% de amostras
Ciprofloxacina	Levofloxacina	Norfloxacina		
R	R	R	180	87,8
I	S	S	7	3,4
R	R	S	5	2,4
R	S	R	5	2,4
R	S	I	2	0,98
R	I	I	1	0,49
S	S	S	5	2,4

Legenda: Resistente (R); Intermediário (I); Sensível (S).

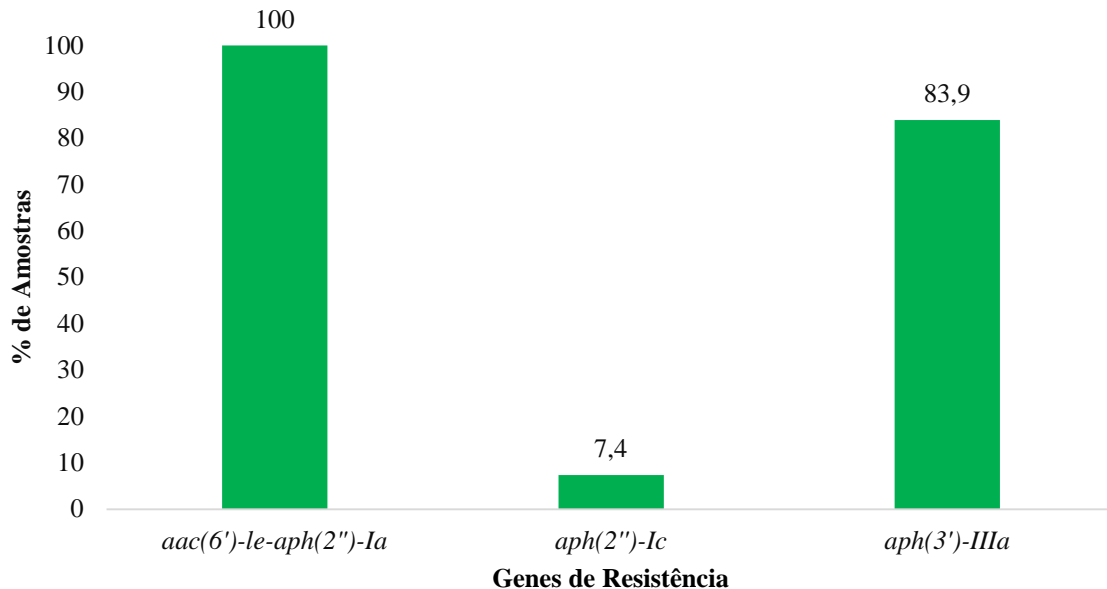
Fonte: A autora, 2020.

3.3 Caracterização dos Genótipos Associados à Resistência

Em relação à vancomicina, todas as amostras que apresentaram o fenótipo VRE (N=166) apresentaram produtos de amplificação para o gene *vanA*.

Todas as 94 amostras que apresentaram fenótipo HLAR foram avaliadas quanto à presença de genes relacionados a esta característica. O Gráfico 4 apresenta os resultados obtidos com as 81 amostras que apresentaram o fenótipo HLGR e HLGSR quanto à frequência dos genes que codificam resistência à gentamicina e a outros aminoglicosídeos na população estudada. O gene *aac(6')-Ie-aph(2'')-Ia*, que codifica uma enzima bifuncional, foi encontrado em todas as amostras, seguido pelo gene *aph(3')-IIIa*, que codifica resistência apenas à canamicina, e esteve presente em 83,9% (68/81) das amostras. Uma porcentagem pequena de amostras, 7,4% (6/81), apresentou produtos de amplificação relativos ao gene *aph(2'')-Ic*. Por outro lado, nenhuma amostra apresentou produtos de amplificação para os genes *aph(2'')-Ib*, *aph(2'')-Id* e *ant(4')-Ia*.

Gráfico 4 – Frequência de determinantes genéticos de resistência a níveis elevados de gentamicina identificados por metodologia de reação em cadeia da polimerase em amostras de *Enterococcus faecium*



Fonte: A autora, 2020.

Em relação a distribuição dos genes representativos do fenótipo HLAR, 62 (76,5%) amostras apresentaram amplificação dos genes *aac(6')-Ie-aph(2'')-Ia* e *aph(3')-IIIa* concomitantemente. A presença unicamente do gene *aac(6')-Ie-aph(2'')-Ia* foi detectada em 13 (16%) amostras. Somente seis amostras (7,4%) apresentaram produtos de amplificação relativos aos três genes avaliados.

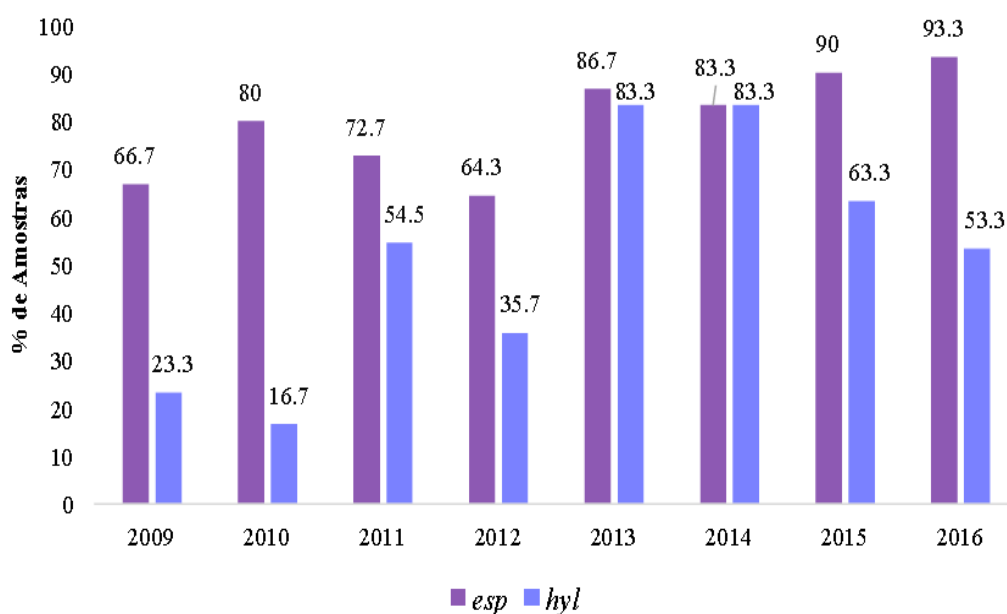
As 39 amostras que foram caracterizadas pelo teste de disco-difusão como não suscetíveis a níveis elevados de estreptomicina, também foram avaliadas quanto à presença de determinantes genéticos de resistência através de um protocolo de PCR *multiplex* para um painel específico de genes. Os resultados mostraram que, em quase sua totalidade (97,4% - 38/39), as amostras apresentaram produtos de amplificação correspondentes com o gene *ant(6')-Ia*. Apenas uma amostra, representando 2,6% do total, apresentou amplificação para o gene *ant(9')-Ia*. O gene *ant(9')-Ib* não foi identificado. Não houve concomitância na presença dos dois genes encontrados.

3.4 Caracterização dos Genótipos Associados à Expressão de Virulência

A análise, por PCR *multiplex*, da presença de cinco determinantes genéticos responsáveis pela expressão de fatores de virulência revelou que 82,4% (169/205) das amostras apresentaram algum dos fatores de virulência pesquisados. Dos cinco genes pesquisados, apenas dois foram observados. O gene *esp* foi o determinante de virulência que obteve a maior predominância, sendo encontrado em 81,4% (167/205) das amostras, enquanto que o gene *hyl* esteve presente em 52,7% (108/205) das amostras.

No Gráfico 5, é possível observar o percentual de amostras que possuíam os determinantes de virulência ao longo dos oito anos de estudo. Nos dois primeiros anos do período avaliado, 2009 e 2010, o percentual de amostras apresentando o gene *hyl* foi inferior a 25% [23,3% (7/30) e 16,7% (5/30), respectivamente]. Entretanto, nos anos seguintes, esse percentual foi superior a 50%. Nos anos de 2013 e 2014, 83,3% (25/30) das amostras carregavam o gene *hyl*. Em relação ao gene *esp*, durante todos os anos avaliados o percentual de amostras que apresentavam o gene foi superior a 60%.

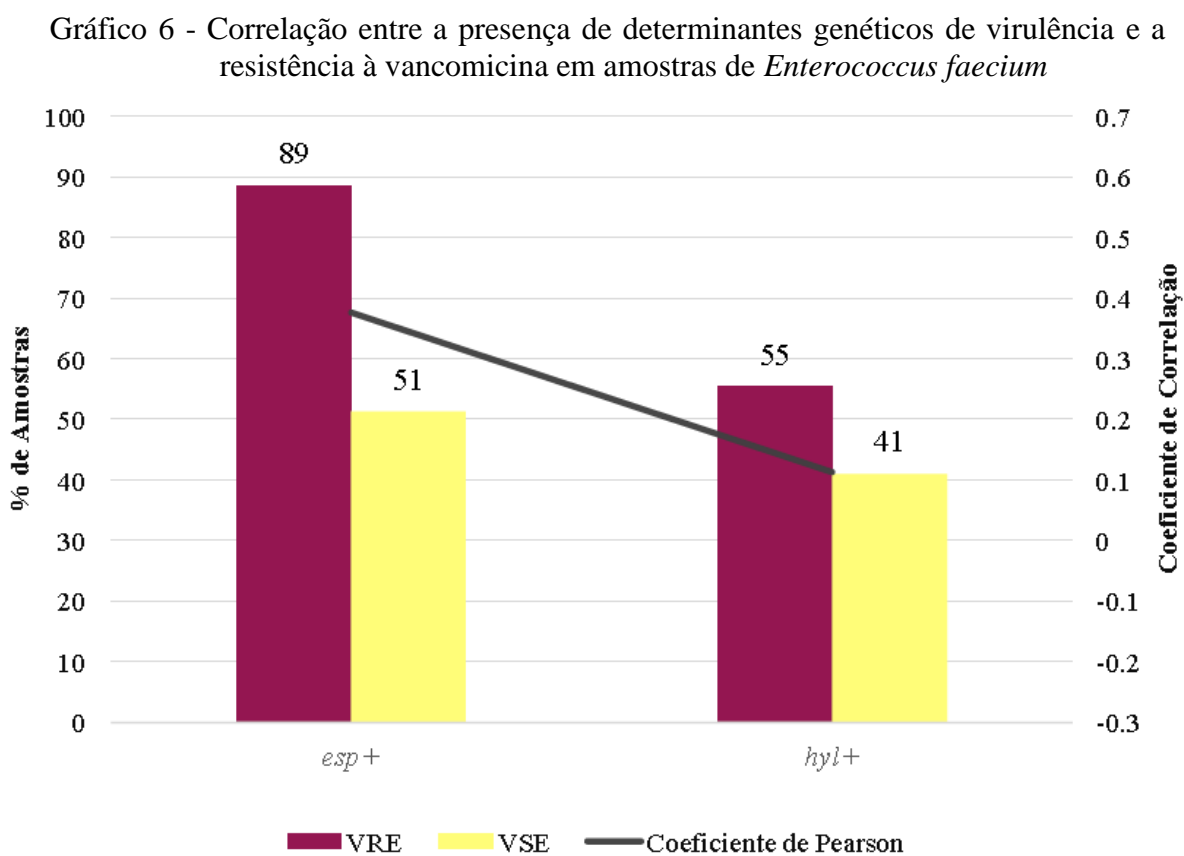
Gráfico 5 - Distribuição dos determinantes genéticos de virulência identificados por reação em cadeia da polimerase em amostras de *Enterococcus faecium* isoladas no período de 2009 a 2016



Fonte: A autora, 2020.

Em relação aos perfis genotípicos de virulência, 51,7% (106/205) das amostras incluídas no estudo apresentaram a presença concomitante dos genes *esp* e *hyl*. O segundo perfil genético mais encontrado, em 61 (29,7%) amostras, foi a presença unicamente do gene *esp*. Apenas duas amostras (0,97%) apresentaram o gene *hyl* somente. Trinta e seis amostras (17,6%) não apresentaram produtos de amplificação para nenhum dos determinantes avaliados.

A correlação entre a presença de determinantes de virulência com a resistência à vancomicina pode ser observada no Gráfico 6. Dentre as amostras VRE, 89% (147/166) apresentaram o gene *esp* e 55% (92/166), o gene *hyl*. As amostras VSE também apresentaram percentuais elevados para a presença do gene *esp* (51% - 20/39) e *hyl* (41% - 16/39). Sendo assim, não foi possível observar correlações significativas entre essas duas características. Para o gene *esp*, o coeficiente de Pearson foi igual a 0,37, indicando uma baixa correlação entre a presença do gene com a resistência ao antimicrobiano. Em relação a *hyl*, o coeficiente de correlação obtido foi ainda menor (0,11), considerado, por isso, não significativo.



3.5 Análise do Sequenciamento do Genoma Completo de *E. faecium*

Todas as 74 amostras, sendo 52 isoladas no período de 2009 a 2016 e 22 pertencentes a um estudo anterior do grupo, como mencionado na seção 2.6, tiveram seu genoma avaliado seguindo os passos descritos na seção 2.7. As métricas do sequenciamento, tais como: número de *reads* obtidas, tamanho das *reads*, cobertura de sequenciamento, número de *contigs*, N50, entre outros, de cada uma das 74 amostras estão listadas no APÊNDICE.

O tamanho do genoma obtido para as amostras, foi em média, de 2,76Mb \pm 0,15. Os 74 genomas de *E. faecium* apresentaram um percentual de conteúdo G+C médio de 37,75 \pm 0,11. A cobertura de sequenciamento obtida foi de 47x \pm 19. Os genomas foram montados, na média, em 205 \pm 48 *contigs*. A anotação do genoma identificou em média 2.715 \pm 159 CDS, 5 \pm 2 rRNA e 54 \pm 7 tRNA.

3.6 Determinação da Estrutura Populacional e Diversidade Genética

A análise da diversidade genética através da avaliação dos perfis alélicos de MLST foi realizada para as 74 amostras que tiveram o genoma sequenciado. No Quadro 2, estão listados todos os STs que compõe a estrutura populacional de *E. faecium* nesse estudo e algumas de suas principais características.

Dentre as 74 amostras avaliadas, foi possível observar 13 STs distintos circulando pelas instituições hospitalares do estado do Rio de Janeiro no período avaliado. Dois deles foram reportados pela primeira vez por nosso grupo (ST1626 e ST1628) e pertenciam ao CC17. Um dos STs ainda não nomeado (ST NOVO) contém um alelo do gene *ddl* ainda não catalogado no PubMLST. Dos 13 STs encontrados, seis deles foram observados em apenas uma amostra, são eles ST97, ST203, ST253, ST885, ST1626 e ST1628. Quatro destes, o ST97, ST253, ST885 e ST1628 não apresentaram determinantes de virulência e resistência à vancomicina e foram todos oriundos de pacientes atendidos em uma mesma instituição (hospital II). Também, ST97, ST253 e ST885 não pertencem ao CC17.

O ST78 foi o predominante, sendo encontrado em 21 amostras, particularmente no período de 2004 até 2007. Entretanto, ao observar as amostras do período de 2009 a 2016, o ST412 foi o mais frequente (19 amostras). Através deste estudo, foi possível observar a

emergência de outros dois STs distintos, não pertencentes ao CC17, ST963 e ST896 (Quadro 2). O ST963 teve sua primeira amostra isolada no ano de 2011, sendo observado em todos os anos posteriores totalizando 12 amostras. Já o ST896 foi encontrado em nove amostras a partir do ano de 2015. Todos esses STs apresentaram uma ampla disseminação em instituições hospitalares do estado do Rio de Janeiro. O ST78, ST412 e o ST963 foram identificados em dez instituições diferentes e o ST896 em seis.

Os resultados indicaram a possível ocorrência de uma nova sucessão de STs, com a predominância do ST412 no período de 2009 a 2012, com uma frequência de 66,7% (14/21) das amostras no período. Cerca da metade das amostras pertencentes ao ST963 foi encontrada nos anos de 2013 e 2014 (5/12). Posteriormente, o ano de 2015 concentrou a maioria das amostras do ST896 (8/9).

Comparado ao ST78, as amostras representativas de STs que emergiram no período proposto pelo estudo (2009 a 2016) apresentaram características fenotípicas e genotípicas marcantes. A presença de resistência a níveis elevados de aminoglicosídeos (HLAR) esteve presente em 58,3% (7/12) das amostras pertencentes ao ST963, enquanto naquelas do ST412, somente 31,5% (6/19) apresentaram essa característica fenotípica. No ST78, a presença do fenótipo HLAR foi de 90,5% (19/21). A presença de resistência à vancomicina também foi predominante no ST78, ST963 e ST896, sendo encontrada em 100% das amostras. Essa característica esteve presente em 73,7% (14/19) das amostras do ST412. Em relação aos perfis genotípicos, foi possível observar que a presença do gene *hyl* pode ser relacionada ao ST412, ST896 e ST963, estando presente em todas as amostras deste último (Quadro 2).

Quadro 2 – Principais características dos diferentes STs encontrados nos 74 genomas de *Enterococcus faecium* circulantes em instituições hospitalares do estado do Rio de Janeiro

ST	CC	Perfil Alélico							N	Distribuição		Perfil de Virulência	Marcadores Fenotípicos de Resistência
		<i>atpA</i>	<i>ddl</i>	<i>gdh</i>	<i>purK</i>	<i>gyd</i>	<i>pstS</i>	<i>adk</i>		Ano	Hospital		
78	CC17	15	1	1	1	1	1	1	21	2004 (4); 2005 (5); 2006 (6); 2007 (4); 2008 (1); 2009 (1)	I (6); IV (3); VI (1); VII (2); XVI (1); XVII (1); XVIII (1); XIX (1); XX (1); XXI (1); XXII (1); XXIII (1); NI (1)	<i>esp</i> ⁺	AMP ^R ; CIP ^R ; HLAR (19); VAN ^R
97	Não-CC17	4	5	1	3	1	20	1	1	2015	II	SD	AMP ^R ; CIP ^R
192	CC17	15	1	1	1	1	7	1	2	2010 (1); 2011 (1)	II	<i>esp</i> ⁺ , <i>hyl</i> ⁺	AMP ^R ; CIP ^R ; HLAR
203	CC17	15	1	1	1	1	20	1	1	2011	I	<i>esp</i> ⁺	AMP ^R ; CIP ^R ; VAN ^R
253	Não-CC17	3	7	3	35	1	1	1	1	2013	II	SD	AMP ^R ; CIP ^R
412	CC17	15	1	1	44	1	20	1	19	2008 (2); 2009 (6); 2010 (3); 2011 (3); 2012 (2); 2015 (2); 2016 (1)	I (3); II (3); III (5); IV (1); V (1); VI (1); IX (1); X (2); XII (1); XXVII (1)	<i>esp</i> ⁺ (10); <i>esp</i> ⁺ , <i>hyl</i> ⁺ (6); SD (3)	AMP ^R ; CIP ^R ; HLAR (6); VAN ^R (14)
844	CC17	15	2	1	1	1	11	1	3	2013	V (1); XI (2)	<i>esp</i> ⁺ , <i>hyl</i> ⁺	AMP ^R ; CIP ^R ; HLAR; VAN ^R
885	Não-CC17	4	3	1	6	1	20	1	1	2013	II	SD	AMP ^R ; CIP ^R ; HLAR
896	Não-CC17	15	2	1	44	1	20	1	9	2015 (8); 2016 (1)	VIII (1); XIV (2); XV (2); XXV (2); XXVI (1); XXIX (1)	<i>esp</i> ⁺ (5); <i>esp</i> ⁺ , <i>hyl</i> ⁺ (3); SD (1)	AMP ^R ; CIP ^R ; HLAR (4); VAN ^R
963	Não-CC17	15	2	1	44	1	11	1	12	2011 (1); 2013 (3); 2014 (2); 2015 (3); 2016 (3)	II (1); V (2); VI (1); IX (1); XII (1); XIII (2); XV (1); XXIV (1); XXX (1); XXXI (1)	<i>esp</i> ⁺ , <i>hyl</i> ⁺ (11) <i>hyl</i> ⁺ (1)	AMP ^R ; CIP ^R ; HLAR (7); VAN ^R
1626	CC17	15	87	1	1	21	1	1	1	2009	VII	<i>esp</i> ⁺	AMP ^R ; CIP ^R ; HLAR; VAN ^R
1628	CC17	14	3	1	2	1	1	1	1	2010	II	SD	CIP ^R
NOVO	-	15	-	1	44	1	11	1	2	2015	VIII	<i>esp</i> ⁺ , <i>hyl</i> ⁺	AMP ^R ; CIP ^R ; VAN ^R

Legenda: *Sequence type* (ST); Complexo clonal (CC); número de amostras (N); Sem determinantes (SD); Ampicilina (AMP); Ciprofloxacina (CIP); *High-level aminoglycoside resistance* (HLAR); Vancomicina (VAN); Não informado (NI); + (positivo); ^R(resistência).

Fonte: A autora, 2020

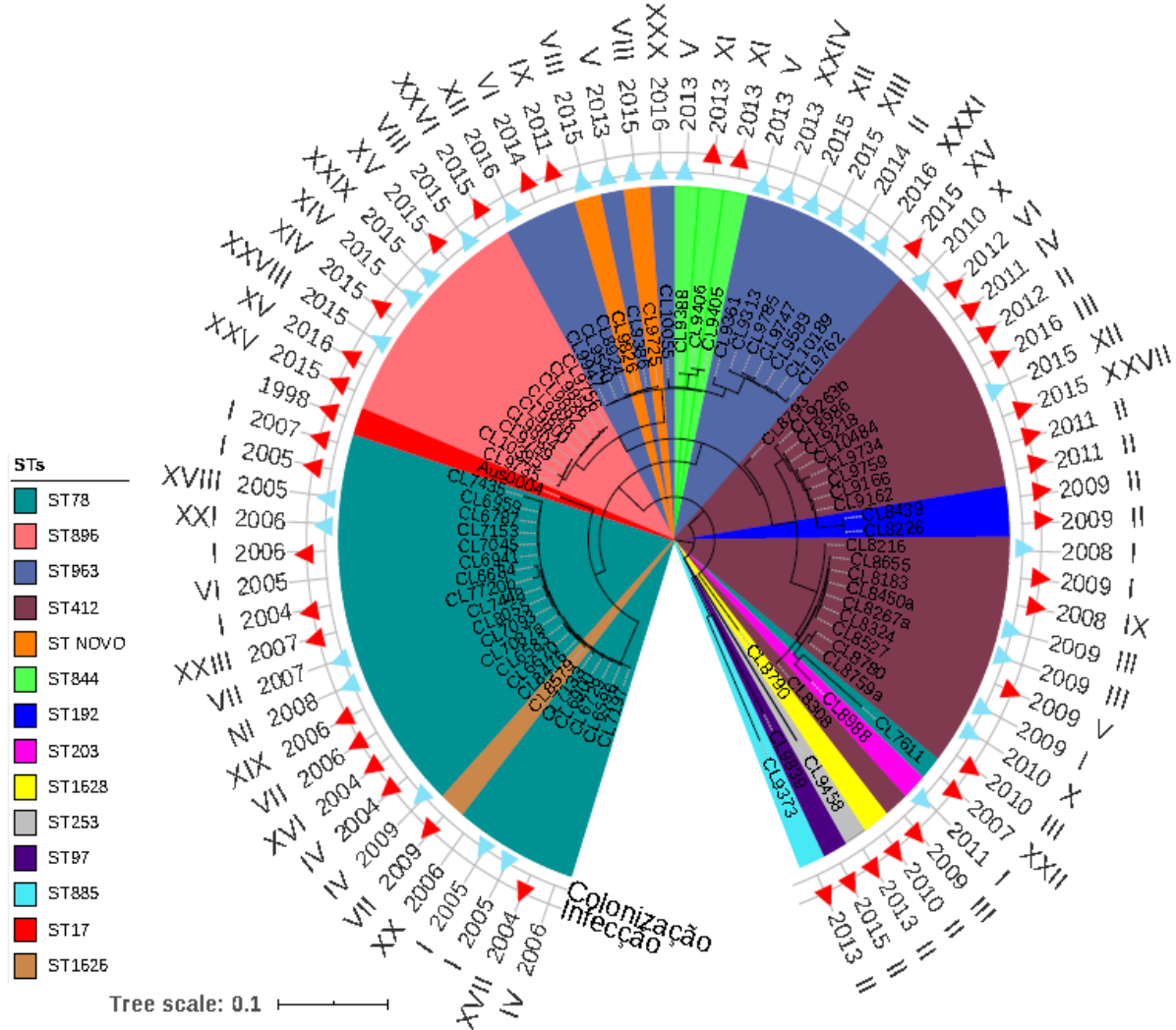
3.7 Reconstrução Filogenética

A reconstrução das relações filogenéticas utilizando Máxima Verossimilhança a partir da matriz de SNPs permitiu identificar ao menos dois clados distintos (Figura 2). Apenas quatro amostras não se agruparam nesses dois ramos. Um dos clados foi formada quase que exclusivamente pelas amostras do ST78 e incluiu também a amostra de referência AUS0004, além do ST1626, descrito pela primeira vez nesse trabalho. O grupo de amostras do ST78 e do ST896 teve origem em um único ancestral comum, sendo considerados como grupos de origem monofilética.

O outro clado apresentou duas ramificações distintas: uma onde foram agrupadas todas as amostras do ST412; e outra constituída das amostras do ST963 e ST896, evidenciando uma ancestralidade comum entre esses grupos de STs e distinta da que originou o ST78. Além disso, nossos resultados indicaram a emergência do ST412 em dois períodos diferentes por grupos filogeneticamente distintos. Um dos ramos tinha amostras isoladas no período de 2008 a 2010, enquanto que o outro grupo apresentava uma circulação mais recente, de 2011 a 2016.

A distribuição desses grupos de STs em ramos filogeneticamente distintos ratificou os dados preliminares mostrados anteriormente no Quadro 2, indicando uma sucessão de STs ao longo dos oito anos.

Figura 2 – Árvore filogenética baseada em SNP construída através de Máxima Verossimilhança e enraizamento pelo ponto médio (*midpoint rooting*), dos 74 genomas de *Enterococcus faecium*

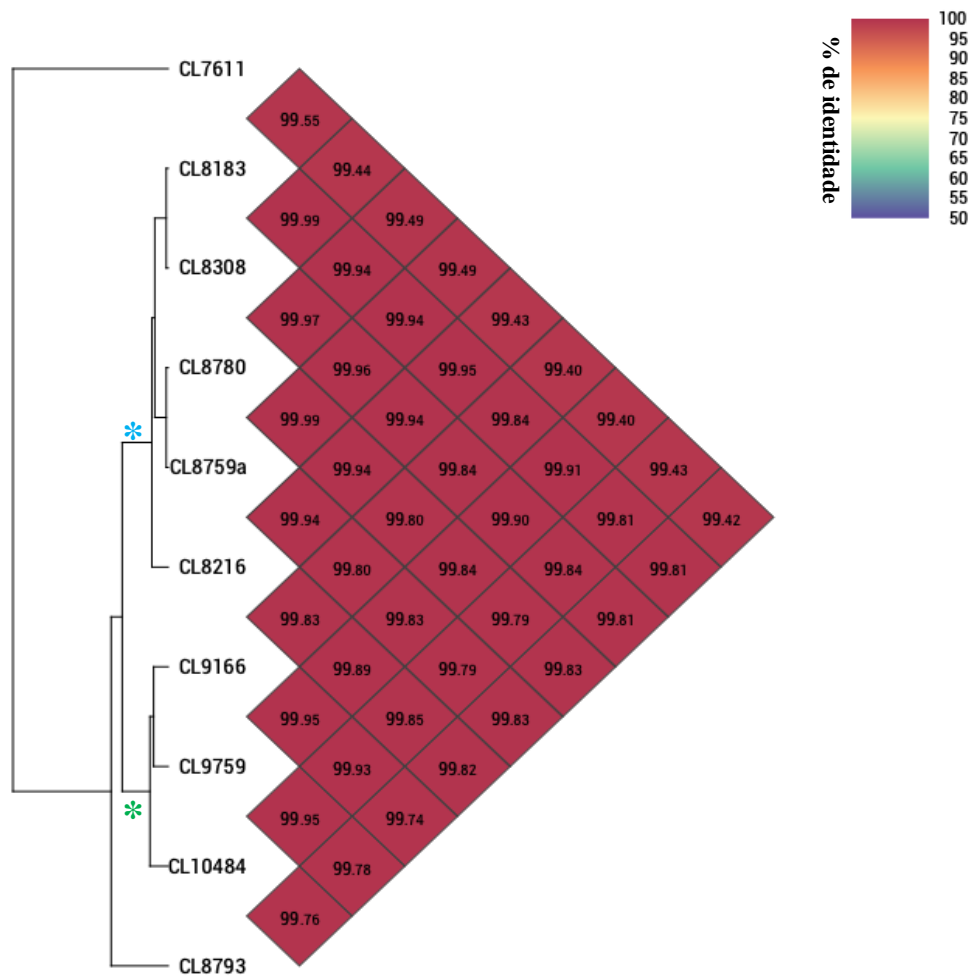


Legenda: As amostras foram coloridas de acordo com o ST ao qual pertencem; Círculo externo em algarismos romanos (I a XXXI) – hospital de origem; Círculo interno com data (1998 a 2016) - ano de isolamento; Triângulos vermelhos - espécimes clínicos representativos de infecção; Triângulos azuis - espécimes clínicos representativos de colonização. Aus0004 incluída como amostra de referência.

Fonte: A autora, 2020.

Através da árvore filogenética, foi possível observar também que uma amostra do ST78 (CL7611), estava filogeneticamente relacionada ao grupo do ST412 (Figura 2). Para investigar a similaridade desse genoma com as amostras do ST412, uma análise do percentual de ANI foi realizada pelo programa OrthoANI (Figura 3). A média de identidade nucleotídica entre a amostra CL7611 e nove genomas, de diferentes braços do ST412, foi de 99,45%. A árvore construída pelo método de UPGMA utilizando os valores de ANI evidenciou uma separação mais clara entre a amostra CL7611 e o grupo de amostras do ST412. Além disso, foi possível confirmar a observação quanto à divisão do ST412 em dois ramos distintos, demarcados em azul e verde na Figura 4. O percentual médio de identidade nucleotídica entre as amostras do ramo demarcado em azul foi de 99,95%, enquanto que para as amostras pertencentes ao ramo demarcado em verde foi de 99,94%.

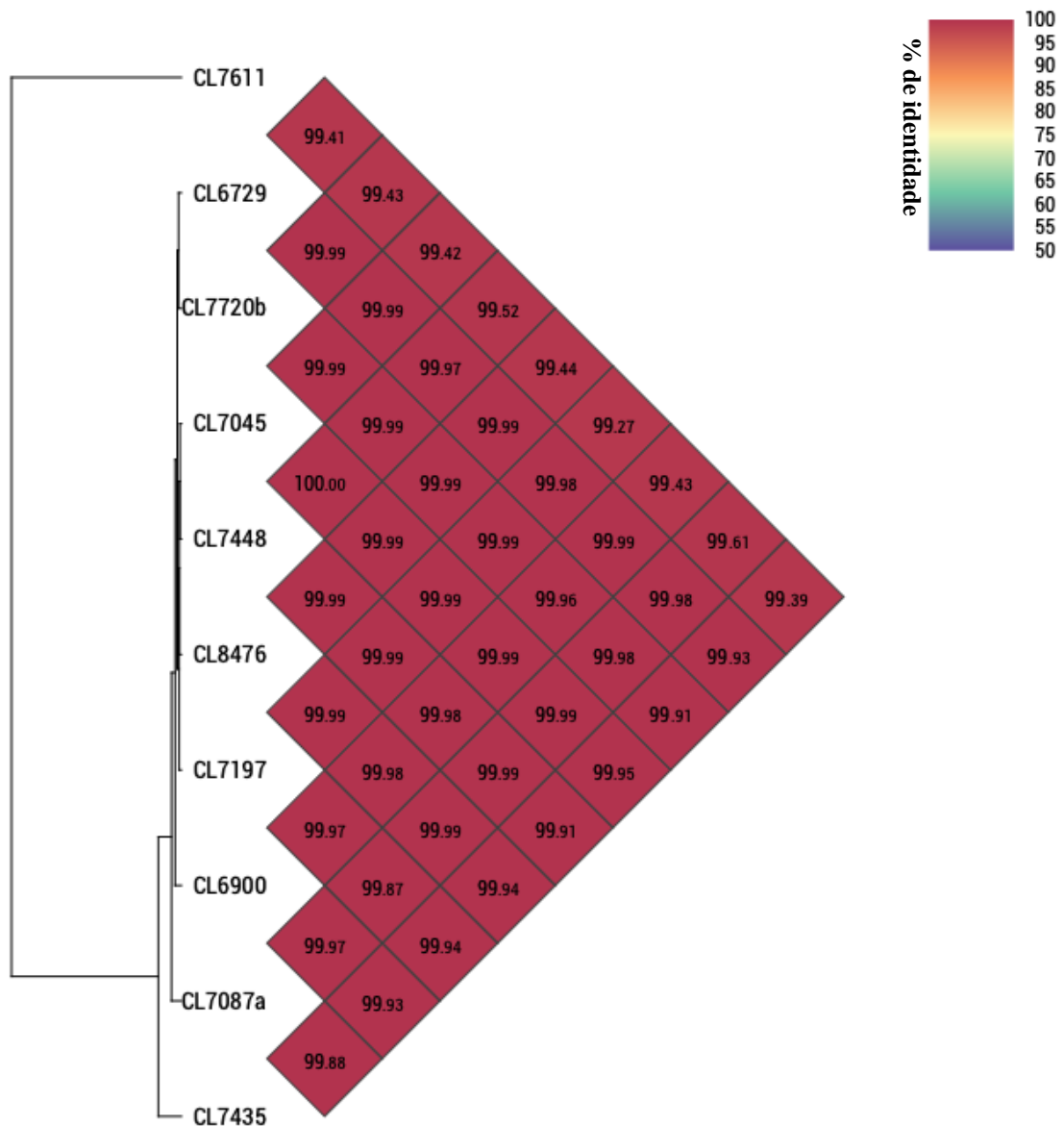
Figura 3 – Mapa de calor com arranjo hierárquico em *cluster* através do algoritmo UPGMA indicando os valores de ANI entre a amostra CL7611 e outras nove amostras pertencentes ao ST412



Legenda: * * indicam a demarcação dos ramos distintos do ST412.
Fonte: A autora, 2020.

A mesma análise foi realizada entre a amostra CL7611 e outras amostras identificadas do ST78. Os resultados estão demonstrados na Figura 4. A média de identidade nucleotídica obtida foi de 99,43%. O valor médio de ANI, entre as outras amostras do grupo do ST78, foi de 99,96%. Tais resultados confirmaram um distanciamento entre a amostra CL7611 e as demais também pertencentes ao ST78. Os resultados corroboraram a estreita identidade entre amostras pertencentes ao grupo do ST78, indicando uma elevada semelhança genética, representada por braços bem curtos na árvore filogenética.

Figura 4 – Mapa de calor com arranjo hierárquico em *cluster* através do algoritmo UPGMA indicando os valores de ANI entre a amostra CL7611 e outras nove amostras pertencentes ao ST78



Fonte: A autora, 2020.

3.8 Estudo do Conteúdo Genético

O estudo da composição do conteúdo genético teve como principal objetivo comparar possíveis diferenças na presença/ausência de genes, principalmente aqueles envolvendo características importantes, tais como a resistência a antimicrobianos e a virulência, que pudessem ser associadas ao surgimento e expansão de algumas linhagens no cenário hospitalar. Diante desta finalidade, após a observação de três linhagens circulantes em tempos distintos durante o período de 2009 a 2016 e a constatação da possível substituição da linhagem do ST78 nesse período, as amostras foram divididas em cinco grupos: ST78, ST412, ST963, ST896 e “Outros STs” (reunindo as amostras pertencentes aos demais STs identificados no período). Os resultados estão mostrados na Figura 5.

O resistoma dos 74 genomas analisados foi composto por 21 genes associados a diferentes classes de antimicrobianos, tais como: um gene associado à resistência aos glicopeptídeos (*vanA*), cinco genes associados à resistência a níveis elevados de aminoglicosídeos [*aac(6')-Ie-aph(2'')-Ia*; *ant(6')-Ia*; *ant(9')-Ia*; *aph(3')-IIIa*; *sat4*], sete genes associados à resistência aos macrolídeos, estreptograminas e lincosamidas [*efmA*; *erm(B)*; *ermA*; *msrC*; *lnuB*; *lsaE*; *eatAv*], quatro genes associados à resistência às tetraciclina [*tet(L)*; *tet(M)*; *tet(S)*; *tet(U)*], dois genes associados à resistência ao trimetoprim (*dfrG* e *dfrF*), um gene associado à resistência às oxazolidinonas [*cfr(B)*] e um gene associado à resistência ao cloranfenicol (*cat*).

De todos os genes identificados, apenas oito deles foram encontrados em mais de 80% (60/74) das amostras. São eles: *vanA*, *aac(6')-Ie-aph(2'')-Ia*; *ant(6')-Ia*, *aph(3')-IIIa*; *sat4*, *efmA*; *ermB*, *msrC*. O gene *cfr(B)*, característico em amostras resistentes à linezolida, foi encontrado em uma amostra isolada no ano de 2015 pertencente ao ST896; no entanto, essa amostra foi caracterizada fenotipicamente como sensível a esse antimicrobiano. Os genes *tet(L)* e *tet(M)* foram encontrados em 18,9% (14/74) e 23% (17/74) das amostras, respectivamente. Entretanto, foram identificados em todas as amostras pertencentes ao ST896. Além dos mencionados anteriormente, através da anotação do resistoma, foi possível identificar outros dois genes associados à resistência à tetraciclina. Os genes *tet(S)* e *tet(U)* foram identificados na amostra CL7611, pertencente ao ST78, mas filogeneticamente relacionada ao grupo do ST412 (Figura 5).

A associação da presença/ausência de determinantes de resistência com as complexas sucessões de linhagens observadas no estudo pode ser observada para o gene *dfrF*. Este gene

não foi detectado em amostras do grupo do ST78, porém foi encontrado em 17 das 19 amostras do ST412 sequenciadas. Além disso, o gene também foi identificado em 11 das 12 amostras do ST963 sequenciadas e em seis das nove amostras ao ST896.

Comparando os dados obtidos para a presença de genes de resistência através da técnica de PCR *multiplex* com o observado pela análise do sequenciamento do genoma completo, foi possível observar algumas discrepâncias. O gene *aph(2'')-Ic* foi identificado em três amostras apenas pela metodologia de PCR, não sendo detectado nas sequências avaliadas. Situações opostas também foram observadas: quatro amostras não apresentaram produtos de amplificação para o gene *aph(3')-IIIa*, mas tiveram o mesmo detectado por análise do genoma completo; uma amostra que apresentou em seu genoma a presença dos genes *ant(6')-Ia* e *ant(9')-Ia*, não apresentou produtos de amplificação compatíveis por PCR *multiplex*.

Cabe ressaltar que como a pesquisa através de PCR *multiplex*, utilizada neste estudo, foi apenas realizada para amostras fenotipicamente resistentes a níveis elevados de gentamicina ou estreptomicina, outros genes envolvidos nos fenótipos HLAR que estão presentes no protocolo utilizado mas são associados à resistência à outros aminoglicosídeos, como o caso do gene *aph(3')-IIIa*, acabam tendo sua real frequência não estimada. Das 29 amostras fenotipicamente sensíveis a gentamicina, 22 apresentaram em seu genoma a presença do gene *aph(3')-IIIa*.

Algumas situações de incongruência entre o fenótipo e o genótipo também foram observadas. Trinta e quatro amostras apresentaram em seu genoma o gene *ant(6')-Ia*, mas foram classificadas fenotipicamente como sensíveis a níveis elevados de estreptomicina (posteriormente detalhado no item 3.9). O mesmo ocorreu para 20 amostras classificadas como sensíveis a níveis elevados de gentamicina, mas que apresentaram no genoma o gene *aac(6')-Ie-aph(2'')-Ia*.

O viruloma foi composto por 13 genes distintos, associados às mais variadas estratégias de virulência. Dentre estes, oito genes são relacionados a aderência: *acm* (MSCRAMM de ligação de colágeno); *ebpaABC* operon (pili); *srtC* (sortase); *ecbA* (proteína de ancoragem de parede celular); *efaA* (antígeno de endocardite A); *sgrA* (adesina de ligação ao fibrinogênio); *scm* (MSCRAMM de ligação de colágeno) e *esp* (proteína de superfície de enterococos). Além desses, também foram identificados o locus *cps* associado à produção de uma cápsula antifagocitária, o gene *bopD* envolvido na formação de biofilme, uma glicosil hidrolase codificada pelo gene *hyl*, uma serina protease codificada pelo gene *htrA* e o gene *bsh* associado à resistência à bile (Figura 5).

Desses 13 genes, somente três deles não foram identificados em mais de 80% das amostras, são eles: *scm*, *hyl* e *bsh*. O gene *hyl* foi identificado em 29 amostras, sendo 41% (12/29) delas pertencentes ao ST963. Este gene não foi identificado em nenhuma amostra do ST78. Em contrapartida, o gene *scm*, identificado no genoma de 20 amostras, foi principalmente anotado em amostras pertencentes ao ST78 (9/20), seguido pelo ST412 (7/20), principalmente em amostras pertencentes à primeira emergência do grupo, isoladas do período de 2008 a 2011.

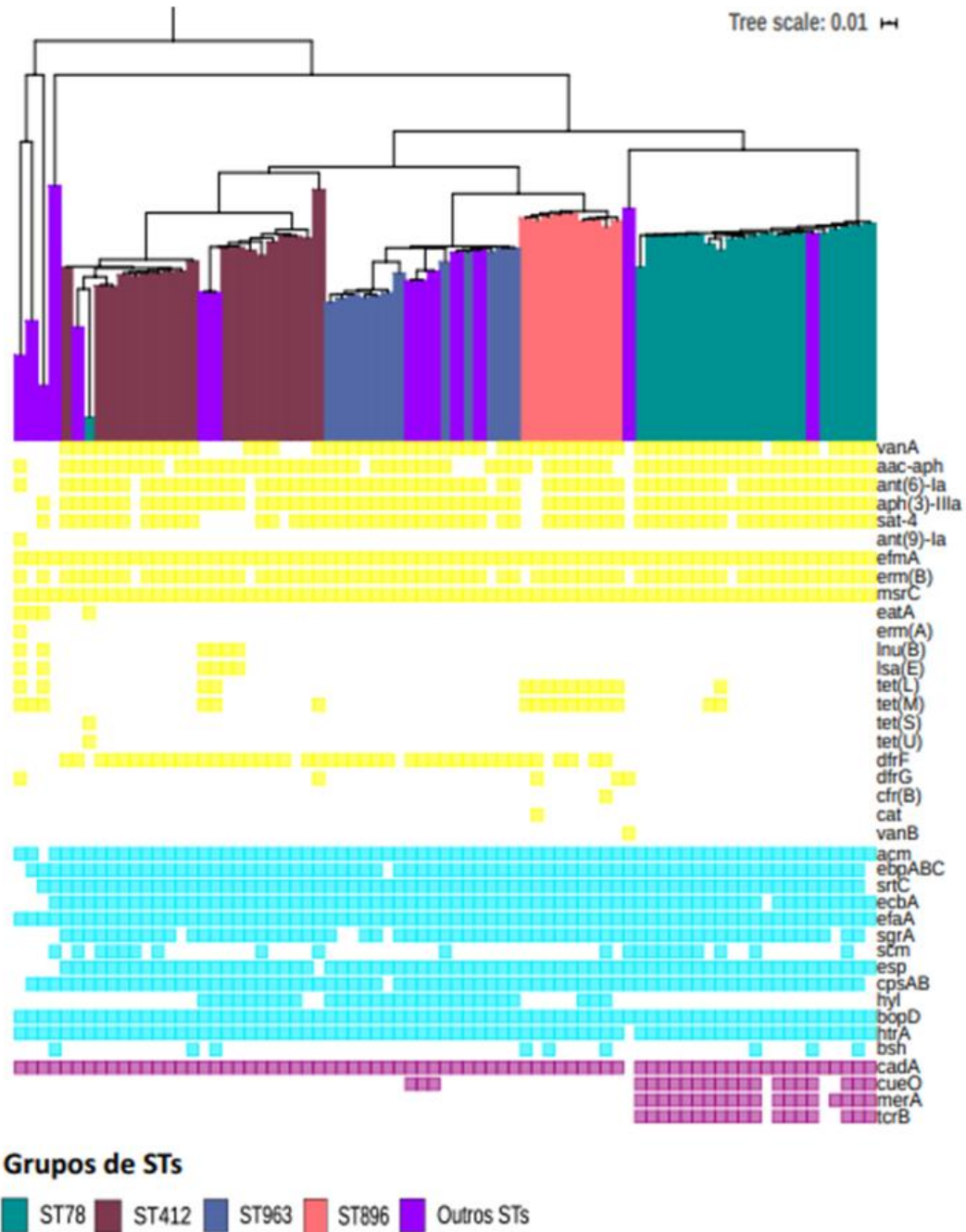
Os dados obtidos através da pesquisa do viruloma não apresentaram discrepâncias em relação ao que foi obtido pela técnica de PCR *multiplex*. Entretanto, a pesquisa de determinantes de virulência utilizando o genoma completo permitiu identificar um maior número de genes de virulência que não estão contemplados no protocolo de PCR utilizado.

Em relação aos genes que codificam resistência a metais pesados, dos quatro genes pesquisados, somente *cadA* que codifica resistência ao cádmio foi encontrado em todas as amostras. Os demais genes – *tcxB*, *merA* e *cueO* – foram encontrados apenas em amostras do ST78 (Figura 5).

Dentro da avaliação do conteúdo genético, também foi observada a presença dos genes associados ao mobiloma, tais como: profagos, ISs e plasmídeos. Os resultados obtidos estão demonstrados na Figura 6. Os genomas de *E. faecium* apresentaram uma extensa diversidade de plasmídeos, sendo identificados 15 tipos plasmidiais diferentes. As amostras apresentaram entre um e oito plasmídeos no genoma ($\bar{x} = 3,6$). De todos os tipos diferentes de plasmídeos encontrados, *repUS15* e *rep17* foram identificados em mais de 80% das amostras.

Além disso, foi possível observar a associação da presença de alguns tipos de plasmídeos com certos grupos de STs. Por exemplo, todas as amostras pertencentes ao ST896 (N=9) apresentaram o plasmídeo *repUS43*, representando 64,3% das amostras que apresentaram esse elemento (9/14). Da mesma forma, o plasmídeo *rep11a* foi prevalente entre amostras dos ST963 e ST896. Das 25 amostras em que esse plasmídeo foi detectado, 60% (15) pertenciam a esses dois STs. Além desses, *rep14b* foi identificado em 87,5% (21/24) de amostras pertencentes ao ST412 e ST963.

Figura 5 – Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs e enraizamento pelo ponto médio (*midpoint rooting*) demonstrando diferenças no conteúdo genético de 74 genomas de *Enterococcus faecium*



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás - “Outros STs”; Quadrados amarelos - presença de genes associados à resistência aos antimicrobianos; Quadrados azuis - presença de genes de virulência; Quadrados lilás - presença de genes associados à resistência aos metais pesados.

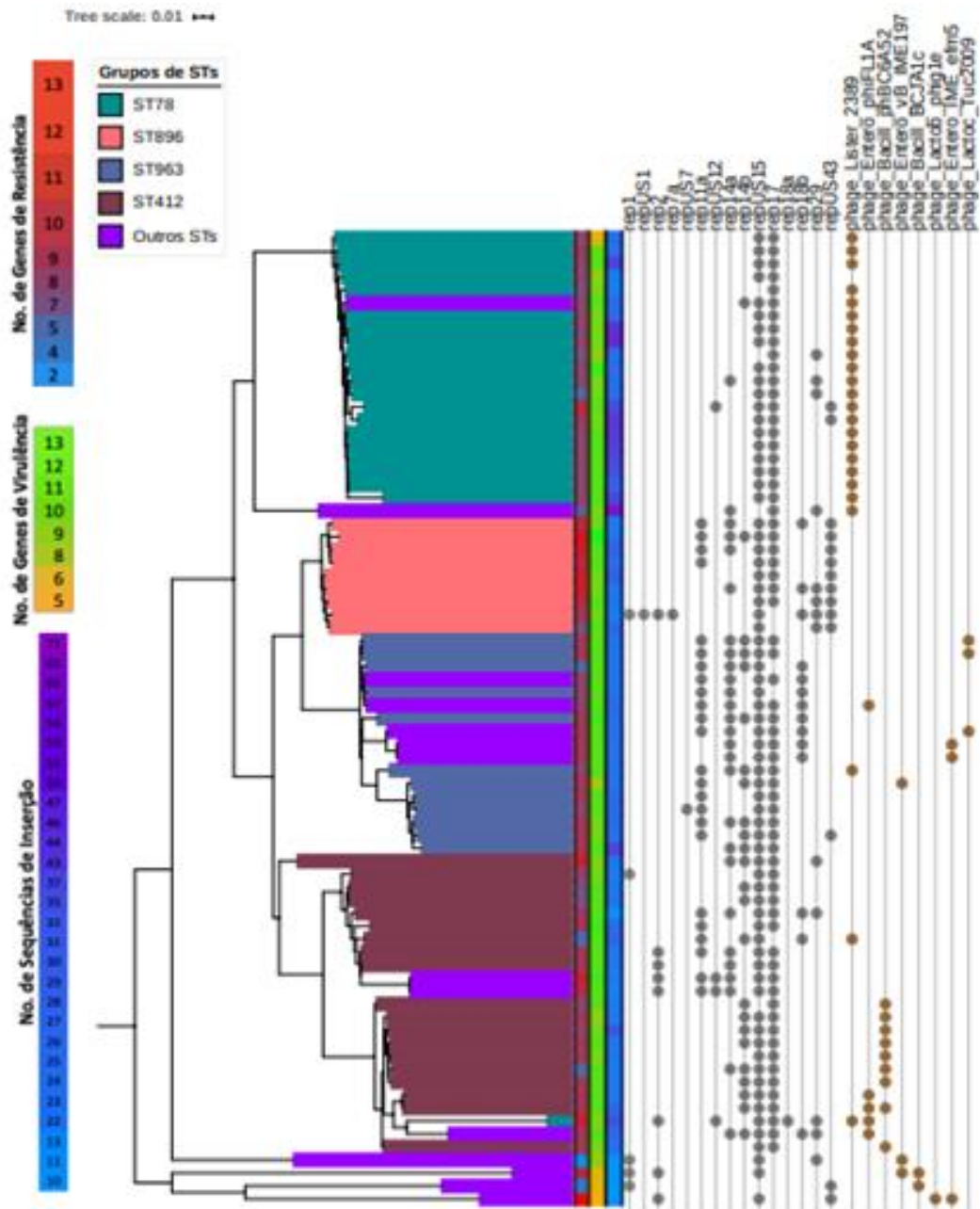
Fonte: A autora, 2020.

A Figura 6 reúne os resultados obtidos com as análises do conteúdo genético das amostras de *E. faecium*. Pode-se destacar que nas 74 amostras estudadas foram anotados entre dois e 13 genes de resistência ($\bar{x} = 8,6$). Já o número de fatores de virulência anotados variou de cinco a 13 genes por genoma ($\bar{x} = 10,4$). Das amostras que apresentaram genes de resistência em quantidades maiores que a média observada (N=40), 47,5% (19/40) carregaram cinco ou mais plasmídeos. Entre as 46 amostras que obtiveram valores acima da média para a presença de determinantes de virulência, 36,9% (14/46) apresentaram entre cinco e seis plasmídeos.

Foram anotados oito profagos distintos intactos (Figura 6). O número máximo de profagos por genoma foi de dois. Em 40% das amostras não foi detectado nenhum profago intacto. Todavia, nos outros 60% de amostras que apresentaram algum profago, foi observado um possível padrão de circulação desses elementos. Todas as amostras do ST78 tiveram em seu genoma a presença do profago PHAGE_Lister_2389, que foi observado apenas esporadicamente em outras três amostras pertencentes aos ST412, ST963 e ST1626. Por outro lado, o profago PHAGE_Bacill_phBC6A52 foi detectado apenas em amostras pertencentes ao ST412, todas elas oriundas da primeira emergência desse ST, que inclui amostras isoladas entre 2008 e 2009.

Em relação as IS identificadas, na Figura 6 é possível observar que as amostras apresentaram entre 30 e 74 IS no genoma ($\bar{x} = 31$ por amostra). Ao todo, 32 tipos de IS distintas foram constatados. Na Tabela 4 estão listadas todas as 32 IS encontradas e suas respectivas distribuições entre os cinco grupos de STs estudados. As amostras do grupo do ST78 foram as que apresentaram o maior número de IS, totalizando 879 desses elementos nos 21 genomas que compuseram esse grupo. O segundo grupo com mais IS foi o ST412, com 512 detectadas em 19 genomas. Os grupos de amostras do ST963 e ST896, compostos por 12 e nove genomas cada, tiveram 375 e 276 IS anotadas, respectivamente. O grupo nomeado de “Outros STs” foi o que obteve a menor quantidade de IS detectadas, 299 em 16 genomas.

Figura 6 – Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs e enraizamento pelo ponto médio (*midpoint rooting*) demonstrando o mobiloma associado ao genoma de *Enterococcus faecium*



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertenceram: verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás - “Outros STs”. Três mapas de calor são observados consecutivamente, o primeiro aponta o número de genes de resistência detectados, onde a cor azul indica o menor número e a cor vermelha o maior. Na linha sequencial, está demonstrado o número de genes de virulência detectados, onde a cor amarela indica o menor número e a cor verde o maior. Na terceira linha está demonstrado o número de ISs encontradas, onde a cor azul indica o menor número e a cor roxa o maior; Círculos cinzas – presença de plasmídeos; Círculos marrons - presença de profagos intactos.

Fonte: A autora, 2020.

Dentre as 32 diferentes sequências de inserção identificadas, *ISEfa11* foi a mais frequentemente anotada: 315 vezes no total. Em seguida, *ISEfa5* foi anotada 210 vezes. Outras seis IS foram detectadas mais de 100 vezes, entre elas: *ISEfa8* (174), *ISEfa10* (139), *ISEnfa3* (129), *IS1476* (120), *ISEfm1* (104) e *ISEf1* (102). Entre as IS que foram observadas apenas uma única vez estão: *ISSIN*, *IS1062*, *IS153*, *ISWci2*, *IS1310* e *ISEfa13* (Tabela 4).

A maioria das IS anotadas foi identificada pelo banco de dados como sendo de origem da própria espécie *E. faecium*, ou como oriundas de outros membros do gênero. Entretanto, identificamos nove IS de espécies distintas, sendo elas: *ISLgar5* de *Lactococcus garvieae*, identificada 34 vezes; *ISCco2* de *Campylobacter coli*, identificada 73 vezes; *IS1182* e *IS256* de *Staphylococcus aureus*, detectadas 82 e 93 vezes cada; *IS1297*, *ISSIE* e *ISSIN* de *Lactococcus lactis*, que foram detectadas em conjunto 73 vezes; *ISWci2* de *Weissella cibaria*, detectada uma única vez e *IS153* de *Lactobacillus sanfranciscensis*, também detectada uma única vez.

Dentre os cinco grupos de STs estudados, foi possível notar diferenças entre as IS predominantes. Todas as IS que obtiveram o maior número de cópias em um determinado grupo ST, estavam presentes em todas as amostras pertencentes ao grupo. Nas amostras do grupo do ST78, as *ISEfa11* e *ISfa5* foram as mais frequentes, representando 33,9% (298/279) das sequências de inserção presentes nesse grupo. Nos grupos ST412 e ST963, *ISEfa11* também foi a mais encontrada, seguida por *ISEfa8*. Já no grupo do ST896, depois de *ISEfa11*, a segunda mais frequente foi *ISEfa5* (24 vezes). As amostras pertencentes ao grupo denominado “Outros STs” tiveram um perfil bastante diverso, com um predomínio de *ISEfa10*, identificada 26 vezes e *ISEfa8*, identificada 23 vezes. Além disso, foi possível observar que as ISs detectadas menos de três vezes foram associadas preferencialmente em amostras dos grupos do ST896 ou do “Outros STs” (Tabela 4).

Ademais, também foi possível constatar que a presença de algumas sequências de inserção pode estar associada com a emergência de alguma linhagem. A *ISLgar5* não foi observada em amostras do ST78, sendo somente identificada a partir da expansão do ST412 e demais STs subsequentes. Já as *IS1216V* e *IS1252*, detectadas cada uma 20 vezes nas amostras do ST78, passaram a não ser identificadas a partir da emergência do ST412, sugerindo um declínio na sua circulação. A *IS1252*, não foi observada nas amostras pertencentes ao ST963 e ST896.

Tabela 4 – Listagem das 32 sequências de inserção detectadas no genoma das 74 amostras de *Enterococcus faecium* estudadas e sua respectiva distribuição entre cinco grupos de STs distintos

IS	Nº de IS (Nº de Amostras) / Grupo de ST				
	ST78 (N = 21)	ST412 (N = 19)	ST963 (N = 12)	ST896 (N = 9)	“Outros STs” (N = 13)
IS1216V	20 (17)	1 (1)	0	2 (2)	2 (2)
IS1216E	11 (10)	13 (11)	11 (9)	14 (6)	11 (10)
IS1251	18 (10)	18 (18)	15 (12)	10 (9)	8 (8)
IS1252	20 (20)	1 (1)	0	0	3 (3)
IS1476	38 (21)	29 (19)	22 (12)	13 (7)	18 (10)
IS1485	20 (20)	18 (18)	12 (12)	10 (9)	14 (13)
IS16	21 (21)	18 (18)	12 (12)	9 (9)	10 (10)
IS1678	21 (21)	20 (18)	12 (12)	10 (9)	9 (9)
IS6770	22 (21)	20 (18)	12 (12)	9 (9)	13 (12)
ISEf1	27 (21)	26 (18)	23 (12)	12 (8)	14 (12)
ISEfa4	21 (20)	18 (18)	12 (12)	8 (8)	9 (9)
ISEfa5	130 (21)	29 (18)	27 (12)	24 (9)	11 (9)
ISEfa7	23 (21)	18 (18)	12 (12)	9 (9)	9 (9)
ISEfa8	50 (21)	49 (19)	29 (12)	23 (9)	23 (9)
ISEfa10	40 (21)	34 (19)	23 (12)	16 (9)	26 (13)
ISEfa11	168 (21)	53 (19)	42 (12)	30 (9)	22 (9)
ISEfm1	46 (21)	21 (18)	13 (12)	9 (9)	15 (12)
ISEfm2	20 (18)	4 (4)	9 (7)	9 (8)	9 (8)
ISEnfa3	42 (21)	37 (19)	24 (12)	10 (9)	16 (9)
IS256	40 (21)	20 (19)	12 (12)	9 (9)	12 (11)
IS1182	40 (21)	16 (15)	11 (11)	7 (6)	8 (7)
IS1297	19 (19)	18 (18)	12 (12)	9 (9)	11 (11)
ISCco2	21 (21)	21 (18)	11 (11)	7 (7)	13 (10)
ISEfa13	1 (1)	0	0	0	0
ISLgar5	0	10 (10)	12 (12)	3 (3)	9 (7)
IS1542	0	2 (2)	7 (7)	1 (1)	0
IS1310	0	0	0	1 (1)	0
ISWci2	0	0	0	1 (1)	0
ISS1E	0	0	0	2 (2)	1 (1)
IS153	0	0	0	0	1 (1)
IS1062	0	0	0	0	1 (1)
ISS1N	0	0	0	0	1 (1)
Total	879	514	375	267	299

Legenda: *Insertion sequence* (IS); *Sequence type* (ST); Número de amostras (N); Escala de cores aplicada em cada coluna para indicar o menor (verde) e o maior valor (vermelho) obtido em cada grupo.

Fonte: A autora, 2020.

3.9 Incongruência Fenotípica e Genotípica

Através do estudo do conteúdo genético e da indicação dos genes presentes no resistoma, constatamos que algumas amostras apresentaram incongruências entre os resultados obtidos em testes fenotípicos de resistência para níveis elevados de estreptomicina e o que havia sido observado genotipicamente. Trinta e seis amostras que foram consideradas suscetíveis à estreptomicina tiveram o gene *ant(6')-Ia* anotado em seus genomas por diferentes ferramentas utilizadas. As amostras que apresentaram esse fenômeno foram isoladas no período de 2008 a 2016, e foram classificadas como pertencentes ao ST412 (17), ST963 (9), ST896 (5), ST NOVO (2), além de uma do ST203 e uma do ST844. Essas amostras representaram, respectivamente, 78,9% (15/19), 83,3% (10/12), 55,5% (5/9) e 7,7% (1/13) do total de amostras que compõem cada um dos cinco grupos de STs observados no estudo.

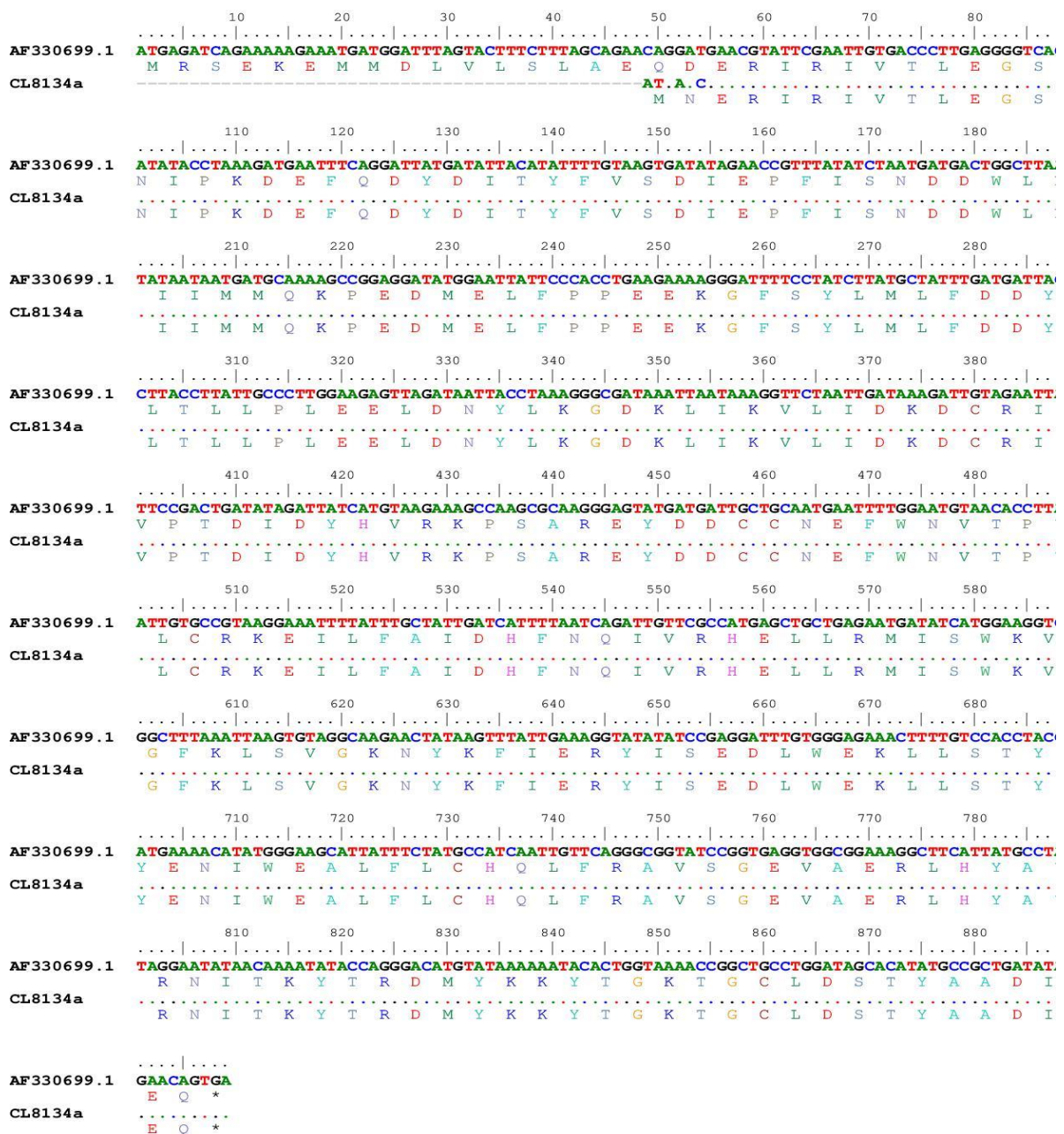
Os dados iniciais obtidos por testes de disco-difusão (amostras sensíveis), que demonstraram incongruência com o PCR convencional [presença de produto de amplificação compatível com o gene *ant(6')-Ia*], foram confirmados pela determinação da concentração inibitória mínima (metodologia padrão-ouro). Os valores de CIM obtidos variaram entre 16 µg/mL a 512 µg/mL, resultados que mantiveram as amostras na categoria sensível para níveis elevados de estreptomicina.

Diante de tais resultados, em uma fase seguinte, a presença do gene *ant(6')-Ia* foi investigada nas demais amostras que compuseram o estudo e se mostraram sensíveis para estreptomicina em concentrações elevadas. Das 124 amostras sensíveis a estreptomicina que não foram avaliadas pela metodologia de sequenciamento do genoma completo, 103 delas apresentaram o gene por técnica de PCR, totalizando 50,3% das amostras deste estudo que mostraram essa característica de incongruência fenotípica-genotípica.

Assim, as possíveis alterações na estrutura do gene *ant(6')-Ia* foram investigadas nos genomas de amostras sequenciadas. Através do alinhamento da sequência de nucleotídeos e aminoácidos com uma sequência de referência (AF330699.1), foi possível observar uma deleção dos 48 primeiros nucleotídeos, que levaram a alterações deletérias na sequência proteica (Figura 7). Além disso, foram identificadas quatro mutações não sinônimas nas posições 49C>A, 50A>T, 52G>A e 54T>C, que levaram a substituições pontuais de uma glutamina por uma metionina (G17M) e de um ácido aspártico por uma asparagina (D18N).

As análises através da ferramenta Primer-BLAST indicaram que o par de iniciadores proposto por Swenson et al. (1995), utilizados nas etapas de detecção do gene por PCR convencional, anelam em posições fora da área afetada pelas alterações observadas (sequência iniciadora direta na posição 179; e reversa na 775), o que proporcionou a identificação das alterações estudadas.

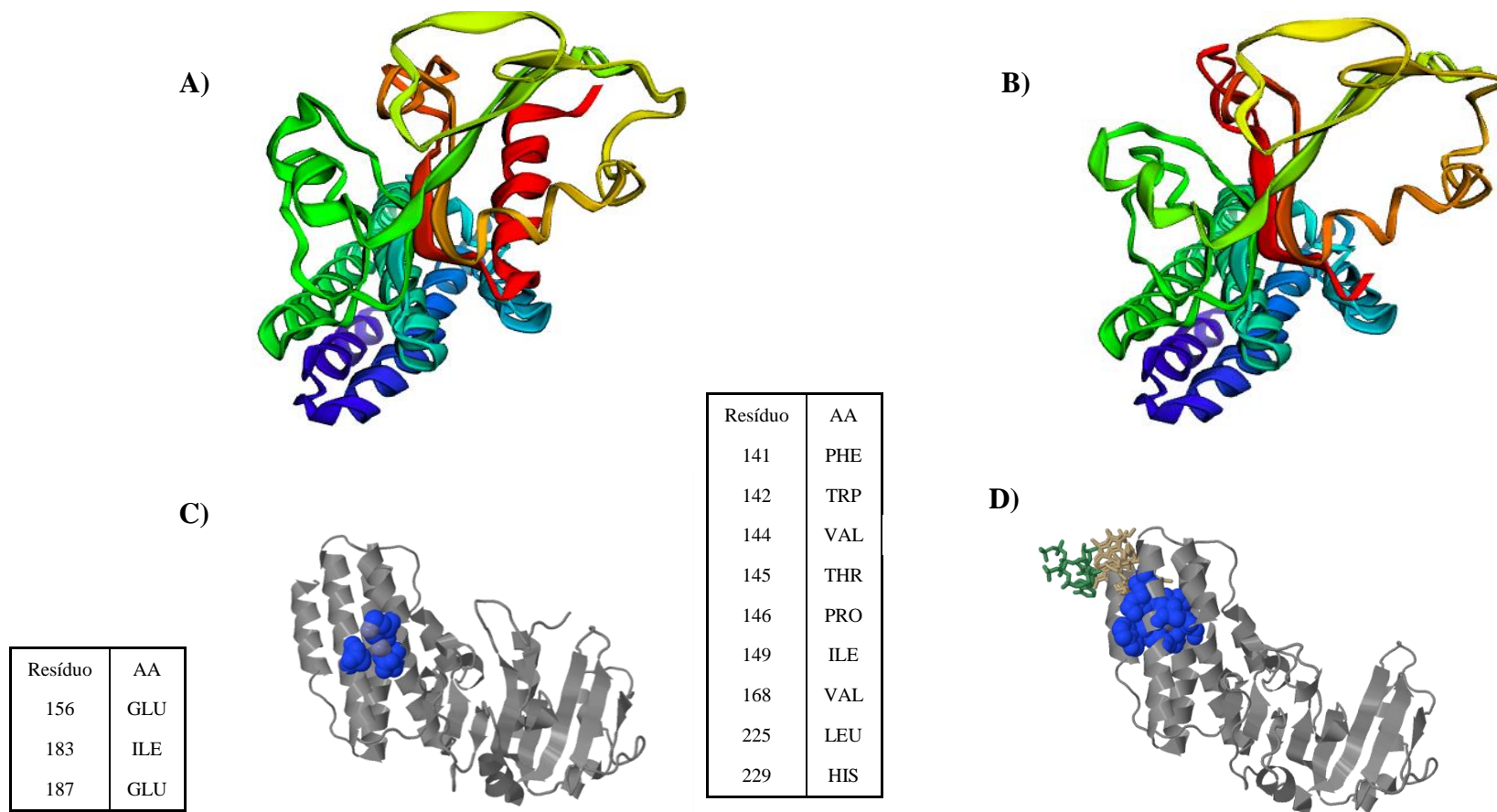
Figura 7 - Alinhamento para comparação das sequências nucleotídicas do gene *ant(6')-Ia*, e respectivas sequências de aminoácidos, entre a referência (AF330699.1) e uma das amostras de *Enterococcus faecium* apresentando resultados incongruentes para a susceptibilidade a estreptomicina, identificada no presente estudo



Fonte: A autora, 2020

A modelagem da proteína codificada pelo gene *ant(6')-Ia* conseguiu predizer uma estrutura em 3D com 100% de confiança e 85% de cobertura (Figura 8). O arcabouço proteico demonstrou alterações na região N-terminal quando comparadas ao modelo de referência (Figuras 8A e 8B). Além disso, essas mudanças levaram a uma alteração no sítio de ligação da molécula (Figura 8D). Em seu estado nativo, a proteína ANT(6')-IA teria seus sítios de ligação nos resíduos 156, 183 e 187, respectivamente, glutamato-isoleucina-glutamato. Com as alterações causadas pela deleção de 17 aminoácidos e consequentes alterações no arcabouço proteico, o programa utilizado fez uma predição de nove resíduos diferentes como representativos do sítio de ligação da molécula.

Figura 8 – Modelo estrutural 3D da proteína ANT(6')-Ia nativa (A) e mutada (B) e respectivos sítios de ligação previstos (C,D)



Legenda: Modelo colorido de vermelho (N-terminal) para azul (C-terminal) para a sequência de referência (AF330699.1) e a sequência mutada; Em azul – sítios de ligação previstos; Aminoácidos (AA); Glutamato (GLU); Isoleucina (ILE); Fenilalanina (PHE); Triptofano (TRP); Valina (VAL); Treonina (THR); Prolina (PRO); Leucina (LEU); Histidina (HIS).

Fonte: A autora, 2020.

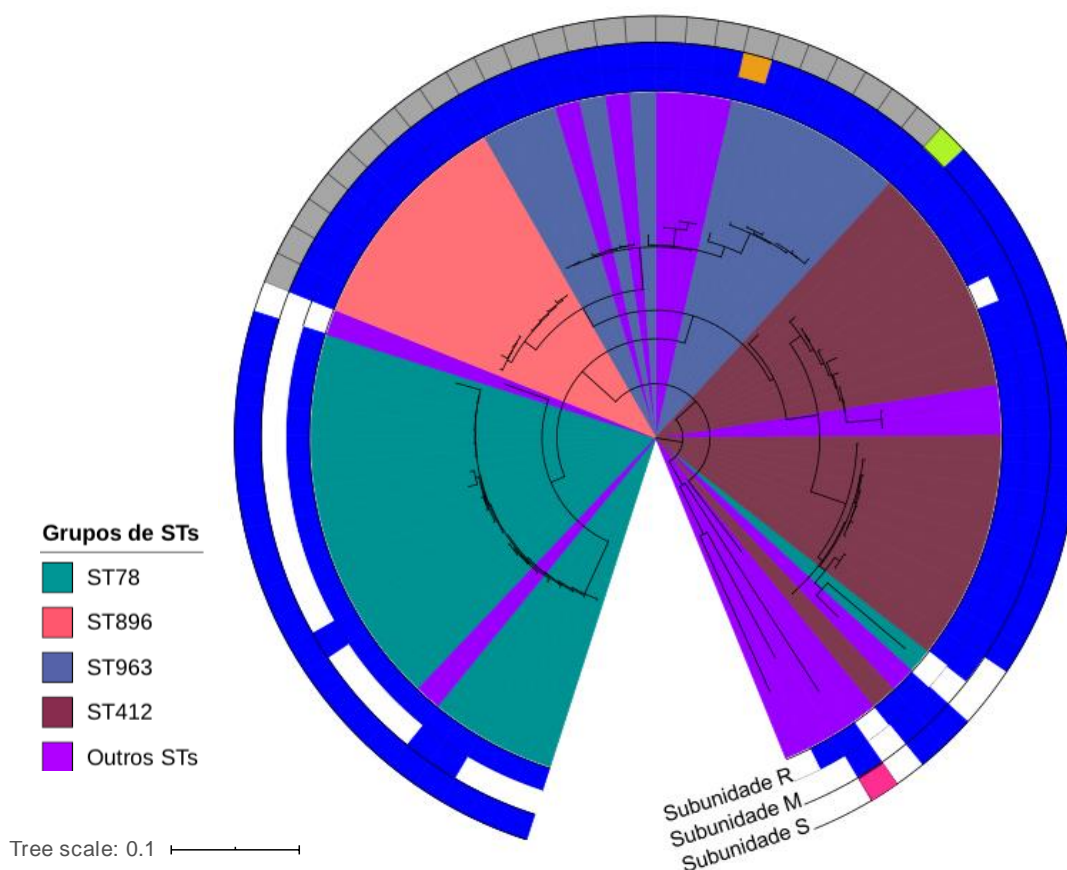
3.10 Análise dos Mecanismos de Defesa

Dentro da análise do conteúdo genético foi, também, investigada a presença de sistemas descritos como barreiras para aquisição de DNA externo, dentre eles o sistema CRISPR-Cas e o sistema de restrição/modificação (R-M). As análises realizadas não identificaram, com nível desejável de confiança, regiões associadas a elementos CRISPR-Cas, através do programa o utilizado (ITEM 2.8.2).

Por outro lado, análises utilizando as sequências descritas no trabalho de Huo et al. (2019), foi possível identificar o sistema R-M do tipo I em 71 dos 74 genomas analisados. Apenas nas amostras CL7611 (ST78), CL8790 (ST1628) e CL9373 (ST885) não foi possível a detecção do sistema completo. Entretanto, como observado na Figura 9, nem todas as amostras apresentaram uma composição homogênea da estrutura do sistema, tanto pela presença das subunidades, quanto na similaridade entre elas. A subunidade R, que possui uma atividade de endonuclease de restrição (REase), foi identificada em 70 amostras, sendo similar em todas elas. Já a subunidade M, que possui uma atividade de metiltransferase cognata (Mtase), foi encontrada em 52 genomas, sendo que em uma das amostras a sequência identificada era diferente das demais. Dentre as amostras do ST78, a subunidade M foi identificada em apenas três desses genomas.

A subunidade S, que tem a função de especificar a sequência de DNA a ser reconhecida, foi a que apresentou maior variabilidade. Sessenta amostras tiveram essa subunidade detectada; entretanto, quatro sequências distintas foram encontradas (demonstradas na Figura 9 em rosa, azul, verde e cinza). As amostras dos grupos do ST78 e ST412 apresentaram similaridade entre as suas subunidades S, enquanto que as amostras dos grupos do ST963 e ST896 apresentaram uma subunidade semelhante entre si, mas diferente da encontrada nos demais. Duas amostras apresentaram ainda, pontualmente, sequências distintas para a subunidade S.

Figura 9 – Árvore de filogenia construída através de Máxima Verossimilhança utilizando matriz de SNPs com enraizamento pelo ponto médio (*midpoint rooting*) demonstrando a presença/ausência das três subunidades que compõem o sistema R-M



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde – ST78, vermelho escuro – ST412, azul – ST963, rosa – ST896 e lilás – “Outros STs”; Na parte mais externa está indicada a presença/ausência da subunidade S, seguida pela subunidade M. A linha mais interna indica a presença/ausência da subunidade R. As seqüências de cada uma das subunidades que apresentaram similaridade entre si, estão pintadas na mesma cor.

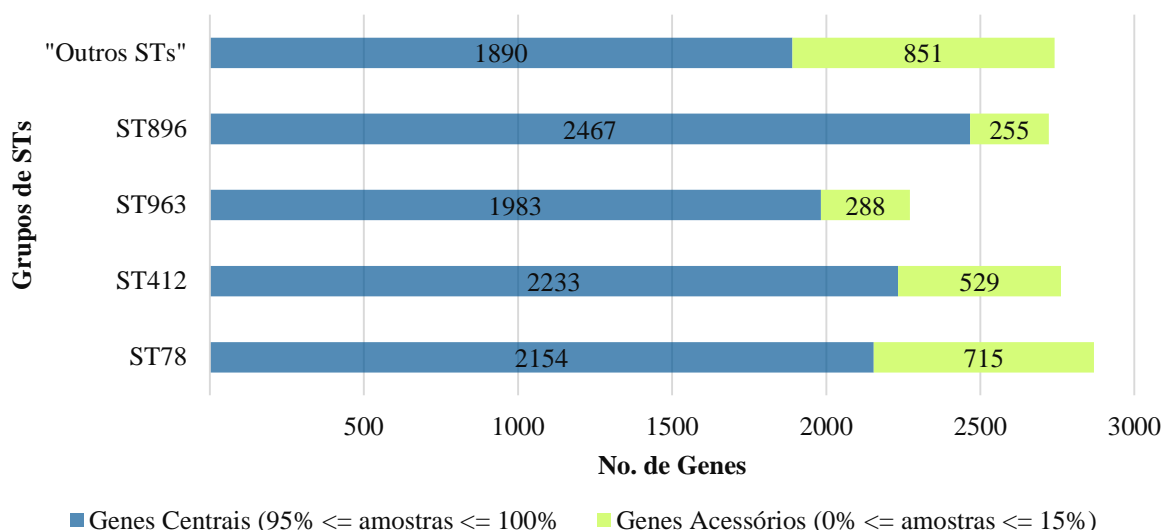
Fonte: A autora, 2020.

3.11 Análise do Pangenoma

O pangenoma das 74 amostras foi estimado em 5.598 genes, sendo 1.925 destes pertencentes ao genoma central (*core* genoma total). Em relação aos grupos de STs estudados, diferenças substanciais na composição do genoma central e do genoma acessório foram observadas (Gráfico 7). O grupo denominado “Outros STs” foi o que continha o maior número de genes considerados acessórios (851 genes), seguido pelo grupo do ST78 (715 genes). Este ST, juntamente com o ST412 e o ST896, continha o maior número de genes

considerados centrais, ou seja, que apareciam em mais de 95% das amostras do grupo. O número de genes centrais nestes grupos foi de 2.154, 2.233 e 2.467, respectivamente. O ST963 e o ST896 continham o menor número de genes acessórios.

Gráfico 7 – Composição do genoma central e do genoma acessório dos cinco grupos de STs de *Enterococcus faecium* observados no estudo



Legenda: *Sequence type* (ST).

Fonte: A autora, 2020.

Considerando os dados obtidos que apontaram para a presença de genes exclusivos em determinados grupos de ST, o passo seguinte foi caracterizá-los para investigar a relação entre a aquisição de uma dada característica e a emergência dos diferentes grupos de STs. Para essa análise, não foram incluídas as amostras pertencentes ao grupo nomeado “Outros STs”, uma vez que não era filogeneticamente único, sendo apenas uma reunião dos outros STs observados no estudo, que apareceram esporadicamente.

Os resultados obtidos demonstraram que 59 genes, no total, foram preditos como sendo exclusivos das amostras do ST78, enquanto que, 45 genes foram exclusivos do grupo de amostras do ST412. O ST963 e o ST896 tiveram o menor número de genes exclusivos, 15 e 12, respectivamente (Figura 10). Embora os genes que não possuíam função conhecida nos bancos de proteínas tenham sido os mais presentes, foi possível observar que o ganho de alguns tipos de funções estava mais associado a certos grupos de STs.

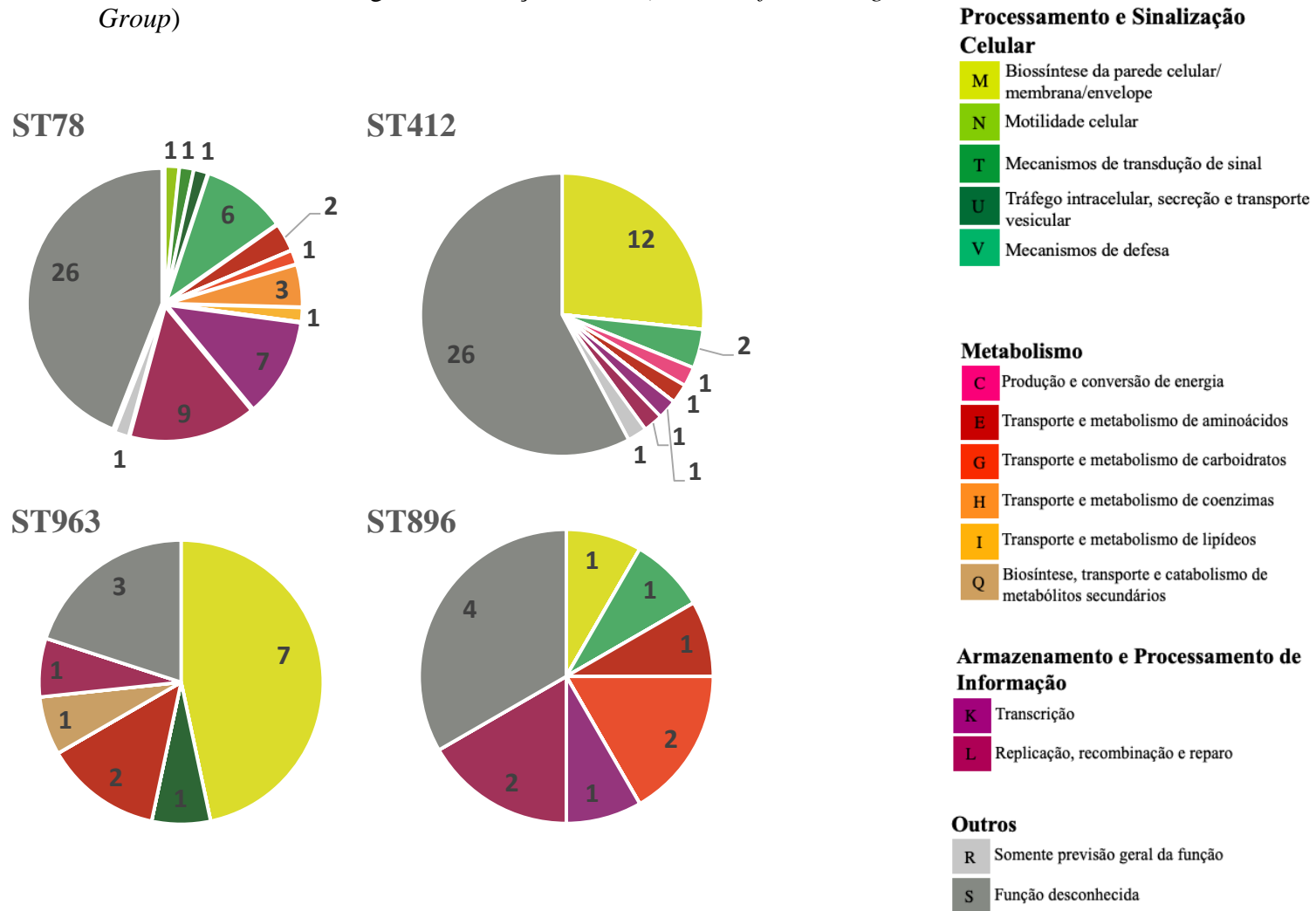
Sendo assim, excluindo os de função desconhecida, para o ST78, 27% (9/33) dos genes preditos como exclusivos estavam envolvidos em processos de replicação,

recombinação e reparo do DNA, seguidos por genes associados a mecanismos de transcrição (21% - 7/33) e mecanismos de defesa (18% - 6/33).

No entanto, para os grupos de amostras pertencentes ao ST412 e ST963, o perfil obtido foi completamente distinto. Nesses dois grupos, a maioria dos genes exclusivos que possuíam função identificada estava envolvida no processamento e sinalização celular, principalmente com genes associados à biossíntese de parede celular, membrana e envelope, representando 63% (12/19) e 58% (7/12) do total obtido para cada grupo. No ST896, 50% (4/8) dos genes exclusivos com função predita foram relacionados a processos de transcrição e também ao transporte e metabolismo de carboidratos (Figura 10).

Dentre os genes obtidos por esses grupos que mereceram destaque, estão os que codificam um sistema toxina (zeta ζ) / antitoxina (épsilon ϵ) nas amostras pertencentes ao grupo do ST78, além da presença em 18 dos 19 genomas do grupo do ST412 de um gene associado à produção de uma beta-lactamase de classe C. Nesse grupo, também foi considerada exclusiva a presença dos genes do operon *neuABCD*, envolvidos no metabolismo de ácido siálico. Nas amostras pertencentes ao ST963, foram identificados os genes *epsD*, *epsC* e *epsL* que estão associados à produção de exopolissacarídeos. No ST896, que teve um ganho de função bastante associado ao metabolismo e transporte de carboidratos, foi identificado o gene *dgoD*, que codifica o D-galactonato desidratase e está associado ao metabolismo do galactonato.

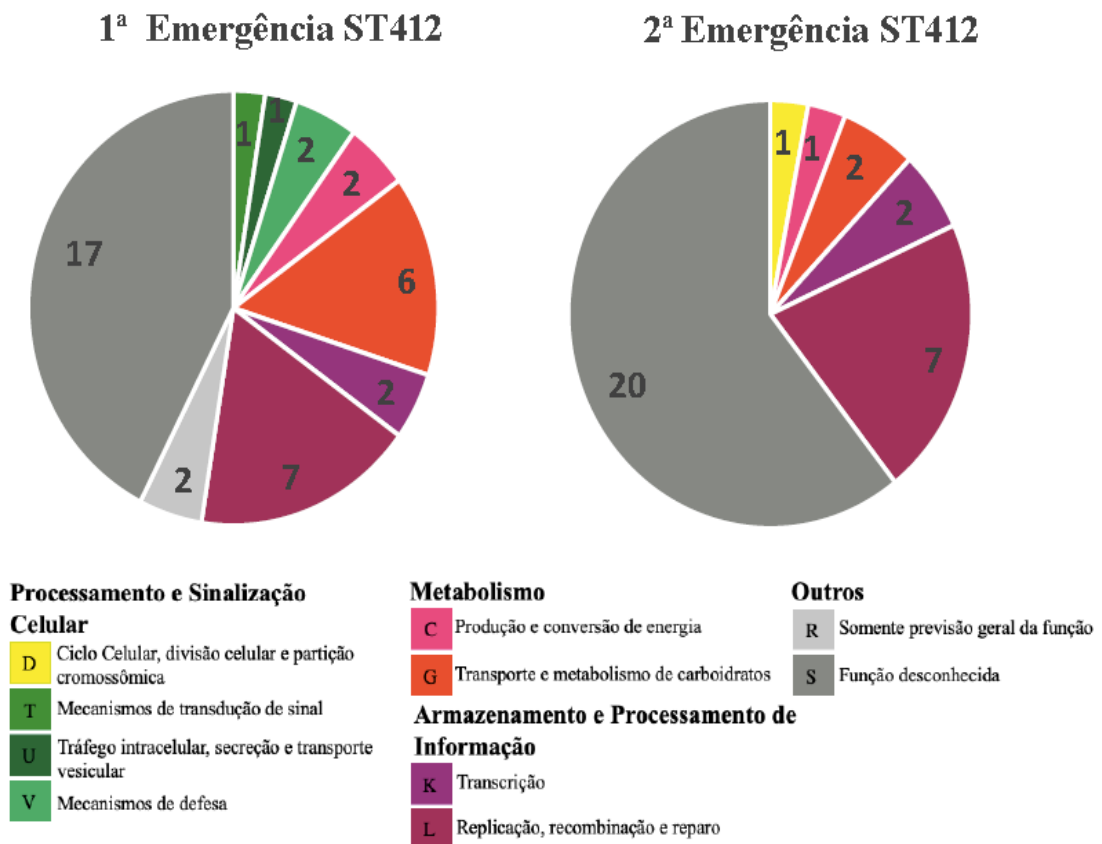
Figura 10 – Distribuição dos genes preditos como exclusivos para cada um dos grupos de STs de *E. faecium* estudados de acordo com a sua categoria de função COG (*Cluster of Orthologous Group*)



Fonte: A autora, 2020.

Além disso, como foi observado através da reconstrução da filogenia (Seção 3.7 - Figura 2), o ST412 teve dois momentos diferentes de emergência, representados por grupos filogeneticamente distintos. Sendo assim, foi também investigada a presença e função de genes exclusivos para cada um desses momentos, nomeados neste estudo como 1ª emergência e 2ª emergência. Foram considerados como genes exclusivos de cada momento aqueles que apresentaram uma especificidade de 100% e sensibilidade maior ou igual a 80% (Figura 11). As amostras pertencentes à 1ª emergência tiveram 40 genes preditos como exclusivos, enquanto que o grupo pertencente à 2ª emergência teve 33 genes exclusivos. Em ambos os grupos, a maior parte dos genes preditos era de função desconhecida, seguida por genes associados à replicação, recombinação e reparo de DNA. Contudo, uma maior diversidade nas amostras da 1ª emergência foi observada, principalmente quando relacionado aos genes associados ao processamento e sinalização celular.

Figura 11 – Distribuição dos genes preditos como exclusivos para cada um dos grupos pertencentes às duas emergências do ST412 de acordo com a sua categoria de função COG (*Cluster of Orthologous Group*)



Fonte: A autora, 2020.

A amostra CL7611, pertencente ao ST78 mas que filogeneticamente ficou relacionada com o ST412 (Seção 3.7 – Figura 2), não foi incluída nos resultados apresentados acima. Essa amostra, em decorrência da sua particularidade, foi analisada individualmente. Foi observado que das 2.843 sequências codificantes (CDS) previstas no genoma desta amostra, 68 também estavam presentes em amostras pertencentes a STs diferentes, exceto no ST412. Por outro lado, foram identificados apenas 44 CDS nessa amostra que apareceram em todos os outros grupos de ST, exceto no ST78. Entretanto, considerando apenas os genes que foram identificados anteriormente como sendo exclusivos de cada ST, a amostra CL7611 apresentou 16 genes previstos como exclusivos do ST412 e apenas quatro que foram exclusivos do ST78.

3.12 Estudo das Mutações

Para identificar as mutações presentes nos 74 genomas de *E. faecium* avaliados, as sequências foram alinhadas em relação ao genoma da amostra de referência AUS0004. A média de SNPs observadas foi de 5.962 ± 1.711 . Uma média de 92 ± 26 inserções e 82 ± 25 deleções em relação à referência foi observada. Os resultados obtidos entre os diferentes grupos de STs estão apresentados na Tabela 5. O grupo de “Outros STs” foi o que obteve a maior média de variações em relação à referência, tanto no número de SNPs (7.645) quanto na presença de inserções e deleções, 123 e 109, respectivamente. O alto valor de desvio-padrão observado indica a heterogeneidade do grupo.

Dentro dos demais grupos que possuem amostras relacionadas filogeneticamente, o ST412 foi o que obteve o maior valor de SNPs, 6.069 (Tabela 5). Em relação aos eventos de inserção e deleção, as amostras do ST963 apresentaram o maior número médio de inserções, com 94. Por outro lado, as amostras do ST896 foram as que atingiram a maior média de eventos de deleção, totalizando 92. O grupo do ST78 obteve os menores valores nos três tipos de variações avaliados, indicando uma similaridade maior do grupo em relação à referência AUS0004.

Considerando a capacidade dessas mutações em gerar efeitos de alto impacto no gene, como o ganho ou perda prematura de códons de parada, foi observado que nas 74 amostras estudadas, o ganho de códons de parada foi observado em média 16 ± 4 vezes em cada genoma; já a perda do códon de parada foi detectada em média 13 ± 2 vezes. Uma

comparação entre os cinco grupos de STs pesquisados também apontou uma variação da presença desses efeitos nos diferentes grupos (Tabela 5)

Amostras pertencentes ao ST412 foram as que apresentaram o maior número médio de eventos de ganho ou perda de códons de parada, 19 e 16 respectivamente. Destaque-se que os valores observados foram superiores aos encontrados nas amostras pertencentes ao grupo de “Outros STs”, que por tendência dos dados já demonstrados nas demais análises, sempre mostram uma grande variabilidade. Por outro lado, o ST963 obteve o menor número desses eventos impactando o genoma de suas amostras, foram 13 relacionados ao ganho de um códon de parada e nove à perda.

Tabela 5 – Valores obtidos para os diferentes eventos de mutação e seus efeitos nos cinco grupos de STs estudados

Variações (Média)	STs				
	ST78	ST412	ST963	ST896	“Outros STs”
SNPs	4.895 ± 374	6.069± 696	5.981± 569	5.773± 166	7.645± 3380
Inserções	75 ± 12	89 ± 8	94 ± 7	89 ± 4	123 ± 48
Deleções	59 ± 6	82 ± 6	89 ± 10	92 ± 6	109 ± 41
Efeitos de Alto Impacto					
Ganho do códon de parada	13 ± 4	19 ± 3	13 ± 2	15 ± 1	18 ± 5
Perda do códon de parada	13 ± 0.8	16 ± 1	9 ± 1	14 ± 0.4	12 ± 3

Legenda: *Single nucleotide polymorphisms* (SNPs); *Sequence type* (ST); ± (mais ou menos).

Fonte: A autora, 2020

De acordo com os dados obtidos, foram avaliados quais os genes e, conseqüentemente, quais proteínas foram impactados pelo ganho ou perda de um códon de parada, além da associação entre esses genes com o respectivo grupo ST.

Os resultados obtidos para o ganho de um códon de parada estão mostrados na Figura 12. Para a construção da figura, foram incluídos somente ganhos prematuros de códon de parada que tenham aparecido em mais de uma amostra. No total, 197 proteínas diferentes foram impactadas pelo ganho de um códon de parada, gerando, assim, uma proteína truncada. Dentre os cinco grupos estudados, as amostras de “Outros STs” foram as que tiveram o maior número de proteínas diferentes afetadas, 114 no total, sendo que 61 delas apareceram somente uma vez. Nas amostras pertencentes ao ST78, 63 proteínas diferentes foram impactadas, sendo que seis delas eram exclusivas desse grupo de amostra. As amostras do ST412 tiveram um número menor de proteínas afetadas, 48, porém sete delas eram exclusivas. Nos grupos do

ST963 e ST896, o número de proteínas afetadas foi de, respectivamente, 31 e 23, e as consideradas exclusivas desses grupos foram cinco e três, respectivamente.

Na Figura 12 pode também ser observado que os efeitos decorrentes do ganho de códons de parada em algumas proteínas foram identificados em mais de 90% das amostras estudadas foi o caso de uma acetiltransferase e de uma proteína determinante do formato de parede celular codificada pelo gene *mreC*.

Todas as amostras pertencentes aos ST78 e ST412 apresentaram uma proteína do sistema tipo I de restrição-modificação impactada. Já nas amostras pertencentes ao ST412, ST963 e ST896, foi possível notar a presença do ganho de códon de parada em algumas proteínas associadas à regulação e reparo de DNA, como a *M protein trans-acting positive regulator* (em tradução livre: proteína M reguladora positiva trans-ativação) e a *pyrimidine dimer DNA glycosylase* (em tradução livre: DNA glicosilase do dímero de pirimidina). Estas mesmas proteínas não se apresentaram modificadas nas amostras do ST78.

Nas amostras do ST412, foi também observada uma proteína reparadora de DNA *mismatch (DNA mismatch repair protein)* afetada pelo ganho de um códon de parada. Nas amostras associadas à 1ª emergência desse grupo, foi possível observar um transportador de múltiplas drogas (*Drug resistance MFS transporter*) e uma proteína envolvida na síntese de cobalamina (*Cobalamin synthesis protein*), ambas truncadas.

O segundo tipo de efeito de alto impacto analisado foi a perda do códon de parada, observada em 50 proteínas no total. Os diferentes grupos de STs tiveram um número diferente de proteínas contendo este erro, além de se observar de que alguns efeitos estavam presentes apenas em grupos de amostras de um mesmo ST (Figura 13). O grupo de “Outros STs”, por reunir amostras heterogêneas, foi o que apresentou o maior número de proteínas diferentes afetadas, 36 no total.

Dentro do grupo de amostras do ST78, o número de proteínas alteradas foi de 26, sendo que três destas foram consideradas exclusivas desse grupo. Assim, como já comentado anteriormente, as amostras do ST412 obtiveram um menor número de proteínas impactadas, porém com uma quantidade maior delas consideradas exclusivas, quatro no total. O ST896 obteve 16 proteínas diferentes modificadas; todavia, nenhuma delas era exclusiva desse grupo. Por fim, o ST963 foi o que obteve o menor número de proteínas diferentes que perderam o códon de parada, 11 no total, sendo uma delas identificadas apenas nesse grupo de amostras.

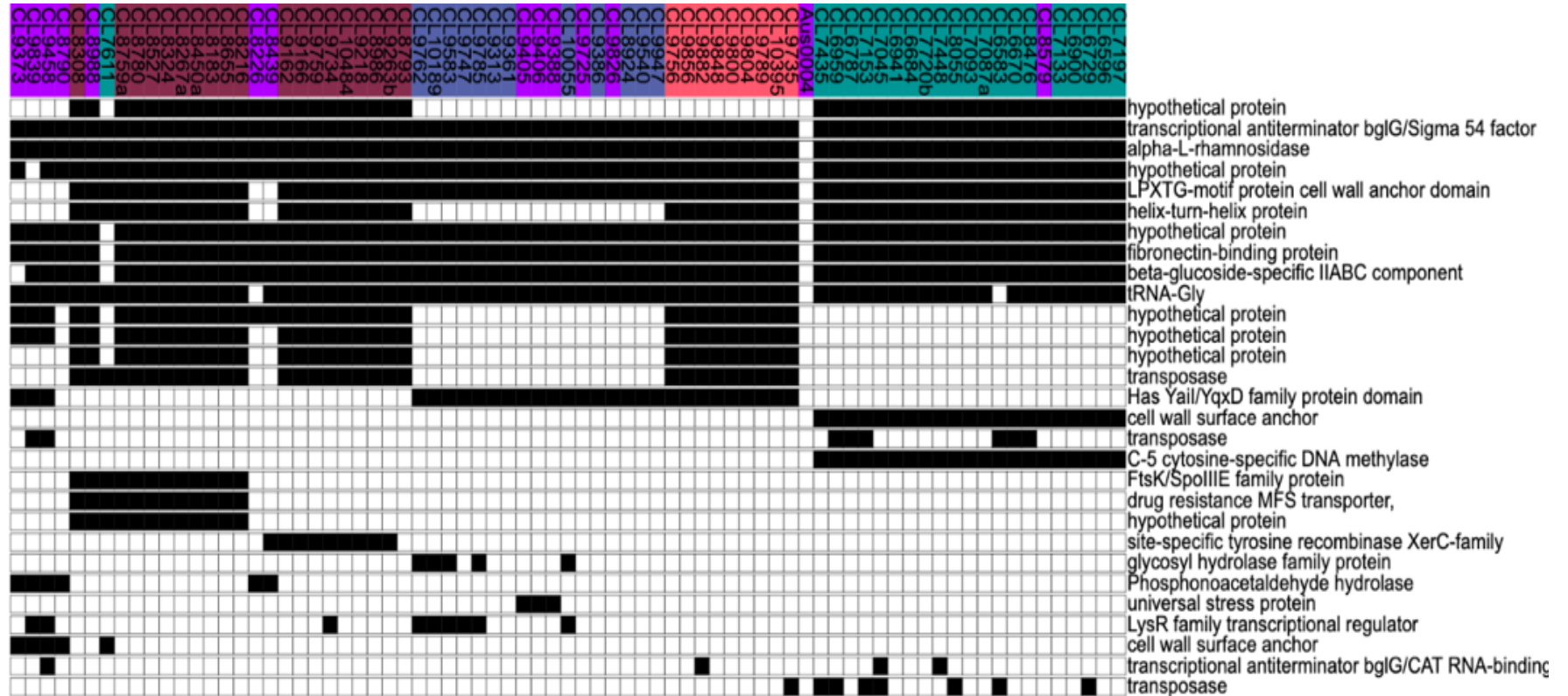
Figura 12 – Avaliação das proteínas anotadas a partir da amostra de referência Aus0004 que tiveram o ganho prematuro de códons de parada nos 74 genomas de *Enterococcus faecium*



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás: “Outros STs”; Quadrados pretos - presença da alteração.

Fonte: A autora, 2020.

Figura 13 – Avaliação das proteínas anotadas a partir da amostra de referência Aus0004 que tiveram a perda prematura do códon de parada nos 74 genomas de *Enterococcus faecium*



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás: “Outros STs”; Quadrados pretos - presença da alteração.

Fonte: A autora, 2020.

Dentre as proteínas que apresentaram alterações, oito delas estiveram presentes em mais de 80% das amostras (Figura 13). Algumas associadas, por exemplo, a estruturas de parede celular (*LPXTG-motif protein cell wall anchor domain, fibronectin binding protein*); a transcrição do DNA (*transcriptional antiterminator bglG/Sigma 54 factor*). Nas amostras do ST78, foi identificada especificamente uma proteína associada à parede celular (*cell wall surface anchor*) e outra associada à metilação de DNA (*C-5 cytosine-specific DNA methylase*). Já no ST412, variações entre as amostras relacionadas à 1ª e à 2ª emergências deste grupo foram observadas: nas amostras da 1ª emergência pode ser observada uma proteína relacionada à partição cromossômica afetada pela perda do códon de parada (*FtsK/SpoIIIE family protein*), enquanto que na 2ª emergência a proteína afetada estava relacionada a processos de recombinação (*site-specific tyrosine recombinase XerC-family*).

3.13 Análise das Mutações Associadas à Resistência às Quinolonas e β -lactâmicos

Os genes *gyrA*, *gyrB*, *parC* e *parE* tiveram sua sequência alinhada em relação a uma referência para observação de mutações que conferem resistência às quinolonas. Os resultados estão mostrados no Quadro 3. Foram observadas mutações tanto nas duas subunidades da DNA girase (*gyrA* e *gyrB*), quanto nas duas subunidades da DNA topoisomerase tipo IV (*parC* e *parE*)

No gene *gyrA*, todas as mutações encontradas foram na posição 84; porém, houve variação em relação ao aminoácido substituído. A substituição mais frequente foi a troca da serina pela arginina (Ser84Arg), observada em 33 (44,6%) amostras. Todas as amostras do ST78 tiveram essa alteração identificada, que também foi encontrada em dez amostras do ST412, em uma amostra do ST203 e na única amostra do ST1626. A segunda substituição mais frequente foi a troca da serina pela tirosina (Ser84Tyr), detectada em 27 (36,5%) amostras, incluindo todas as amostras do ST963 e do ST896. Esta variação foi encontrada em apenas uma amostra do ST412. O restante das amostras pertencentes ao ST412 apresentou a substituição da serina por isoleucina (Ser84Ileu). Apenas quatro amostras (5,4%), pertencentes ao ST97, ST253, ST885 e ST1628, não apresentaram nenhuma mutação no gene *gyrA*.

A maioria das amostras não apresentou alterações na subunidade B da DNA girase. Apenas uma amostra pertencente ao ST253 apresentou a substituição de uma treonina por uma serina na posição 197 (Thr197Ser) (Quadro 3).

No gene *parC*, foram observados dois padrões distintos de substituição, que se mostraram amplamente associados aos diferentes grupos de STs. Todas essas alterações ocorreram no aminoácido serina presente na posição 82. A substituição mais frequente, presente em 44 (61%) amostras, foi a troca da serina por arginina (Ser82Arg). Esta alteração foi encontrada em todas as amostras do ST412 (19) e do ST963 (12), e em oito das nove amostras do ST896. Além disso, também foi identificada em amostras do ST97, ST192 e nas duas amostras pertencentes a ST novos e ainda não nomeados. Todas as amostras do ST78 avaliadas (N=21) possuíram a substituição da serina por isoleucina (Ser82Ileu). Esta substituição foi detectada ainda, no ST203, no ST844 e no ST1626. Apenas três amostras (4,2%) não apresentaram padrão de mutação nesse gene. Além disso, não foi possível avaliar o gene *parC* em duas amostras incluídas no estudo, pois as sequências obtidas estavam parcialmente quebradas, impossibilitando a análise.

Quatro alterações distintas foram encontradas no gene *parE* em cinco amostras diferentes. Algumas dessas amostras apresentaram mais de uma alteração. A troca da isoleucina pela treonina na posição 144 (Ile144Thr) foi a mais frequentemente encontrada. Esta alteração foi identificada de forma exclusiva (1 amostra), ou concomitante com a substituição do glutamato na posição 226 pela valina (Glu226Val) (1 amostra) ou com a substituição na posição 685, de uma asparagina por uma lisina (Asn685Lys) (2 amostras). O quarto tipo de alteração correspondeu a troca de valina por leucina na posição 555 (Val55Leu), que foi observada em uma única amostra, e sem nenhum outro tipo de alteração associada (Quadro 3).

Apenas três amostras (4,05%) não apresentaram mutações concomitantemente em *gyrA* e *parC*, dentre elas, duas apresentaram alterações em *parE* e foram caracterizadas como não suscetíveis à ciprofloxacina e à norfloxacina. A outra amostra apresentou alterações em *gyrB* e *parE* concomitantemente e foi caracterizada como não suscetível às três quinolonas testadas.

Quadro 3 – Alterações observadas nas duas subunidades da DNA girase (*gyrA* e *gyrB*) e da topoisomerase tipo IV (*parC* e *parE*) e sua distribuição nos 74 genomas de *Enterococcus faecium* avaliados

Gene	Padrão de Substituição		Número Total de Amostras	Número Total por ST
	NT (Posição)	AA (Posição)		
<i>gyrA</i>	T252A	Ser84Arg	33	ST78 (21); ST203 (1); ST412 (10); ST1626 (1)
	G251T	Ser84Ile	10	ST192(2); ST412 (8)
	A250T / G251A	Ser84Tyr	27	ST412 (1); ST844 (3); ST896 (9); ST963 (12); ST NOVO (2)
	SEM MUTAÇÕES	SEM MUTAÇÕES	4	ST97 (1); ST253(1); ST885 (1); ST1628 (1)
<i>gyrB</i>	C590G	Thr197Ser	1	ST253
	SEM MUTAÇÕES	SEM MUTAÇÕES	73	TODOS OS OUTROS STs
<i>parC</i>	A244C	Ser82Arg	44	ST97 (1); ST192 (2); ST412 (19); ST896 (8); ST963 (12); ST NOVO (2)
	G245T	Ser82Ile	25	ST78 (20); ST203 (1); ST844 (3); ST1626 (1)
	SEM MUTAÇÕES	SEM MUTAÇÕES	3	ST253 (1); ST885 (1); ST1628 (1)
<i>parE</i>	T431C	Ile144Thr	4	ST97 (1); ST253 (1); ST885 (1); ST1628 (1)
	A677T	Glu226Val	1	ST885
	G1663T	Val555Leu	1	ST963
	C2055A	Asn685Lys	2	ST97 (1); ST253 (1)
	SEM MUTAÇÕES	SEM MUTAÇÕES	68	TODOS OS OUTROS STs

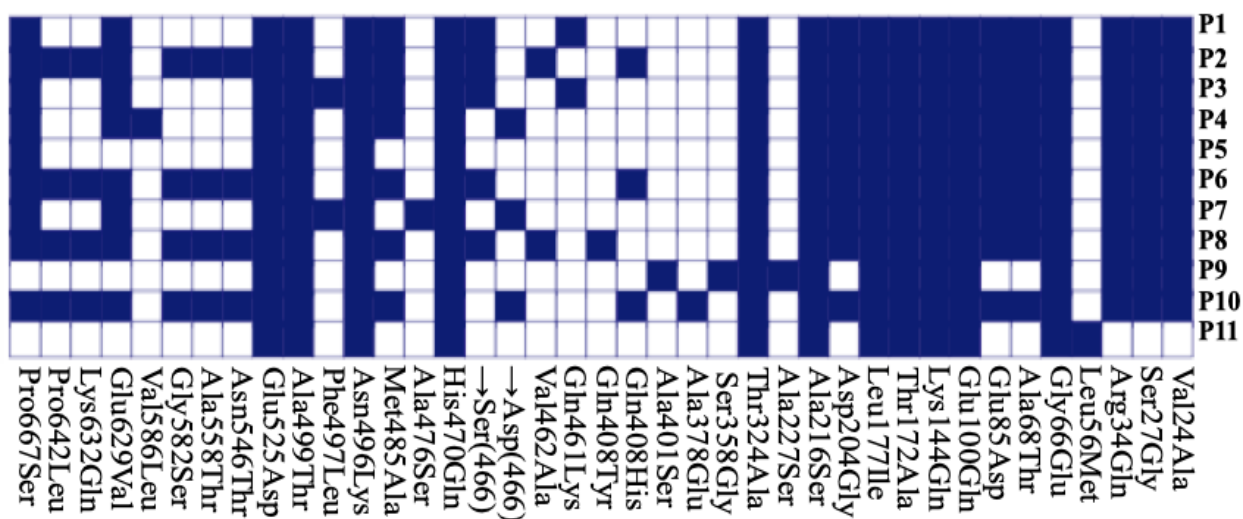
Legenda: Nucleotídeo (NT); Aminoácido (AA); *Sequence type* (ST); Adenina (A); Citosina (C); Guanina (G); Timina (T); Arginina (Arg); Asparagina (Asn); Glutamato (Glu); Isoleucina (Ileu); Lisina (Lys); Leucina (Leu); Serina (Ser); Tirosina (Tyr); Treonina (Thr); Valina (Val).

Fonte: A autora, 2020.

A avaliação das mutações presentes no gene que codifica a proteína ligadora de penicilina 5 (*pbp5*), associada à resistência aos β -lactâmicos, foi realizada em 72 das 74 amostras estudadas. Em dois genomas, a sequência estava quebrada e, por isso, não foi possível ser avaliada. Um grande acúmulo de mutações ao longo do gene pode ser observado. No total 36 aminoácidos diferentes sofreram alterações, além da detecção da inserção de uma serina ou ácido aspártico na posição 467 (Figura 14).

Ao todo, as amostras puderam ser divididas em 11 perfis distintos de acordo com o padrão de mutação obtido. Algumas alterações puderam ser observadas em todos os 11 perfis, enquanto que algumas substituições estiveram presentes em alguns grupos de perfis, ou mesmo isoladamente. Dentre os perfis mais encontrados, estão P1 e P2, que juntos representaram 87,5% (63/72) dos genomas avaliados. Encontrado em 44 (61%) amostras, P2 apresentou alteração em 28 posições diferentes e foi identificado em sete STs distintos, incluindo 89,5% das amostras do ST412 (17/19), 83,3% (10/12) do ST963 e todas do ST896. Por outro lado, P1 acumulou alterações em 22 pontos distintos e foi identificado em apenas 19 (26,4%) amostras, sendo a maioria delas do ST78, além de uma amostra do ST1626. Este perfil correspondeu a 85,7% das amostras do ST78.

Figura 14 – Distribuição das alterações observadas no gene que codifica a proteína ligadora de penicilina 5 (*pbp5*) em onze perfis distintos de amostras de *Enterococcus faecium*



Legenda: Quadrados azuis - presença da substituição; Alanina (Ala); Arginina (Arg); Asparagina (Asn); Ácido aspártico (Asp); Glutamato (Glu); Glicina (Gly); Glutamina (Gln); Histidina (His); Isoleucina (Ile); Leucina (Leu); Lisina (Lys); Metionina (Met); Fenilalanina (Phe); Prolina (Pro); Serina (Ser); Tirosina (Tyr); Treonina (Thr); Valina (Val).

Fonte: A autora, 2020.

3.14 Avaliação das Unidades de Novidade Genética

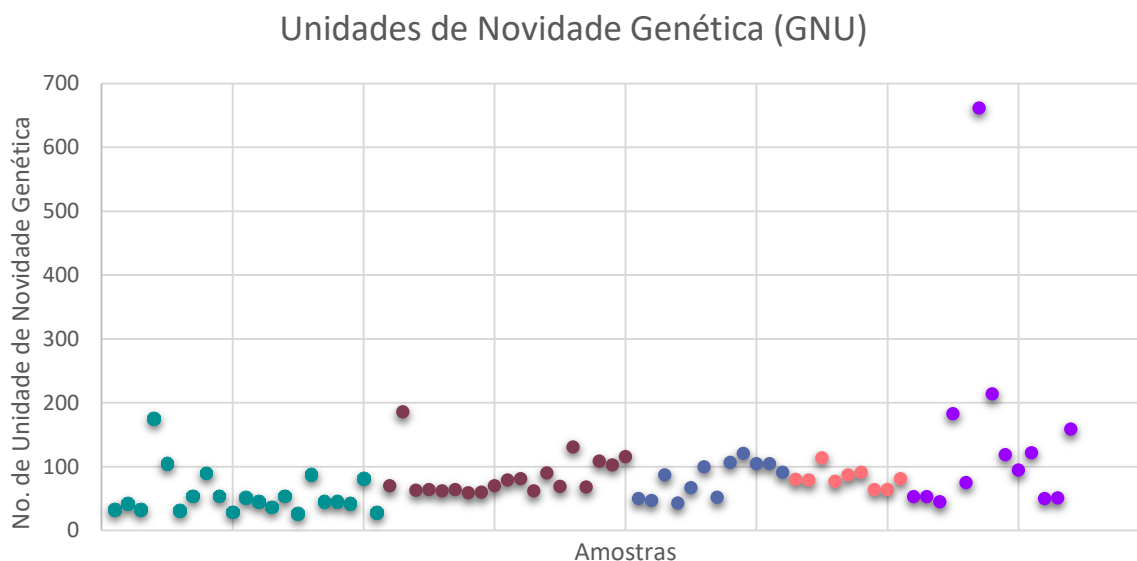
Os dados obtidos foram também avaliados utilizando a ferramenta WhatsGNU, objetivando um estudo mais específico sobre mutações nas amostras incluídas neste estudo. Com essa ferramenta, foi possível observar quais mutações, quando comparadas a um banco de dados extenso de genomas de *E. faecium*, eram identificadas como um novo alelo daquela proteína, sendo categorizada então como uma unidade de novidade genética, ou GNU. Nas 74 amostras estudadas, a média de GNUs foi de $\bar{x} = 86 \pm 77$ por genoma. Quando observados os números de unidades de novidade genética em amostras do mesmo grupo de ST, os valores foram, em geral, semelhantes (Figura 15).

As amostras do grupo “Outros STs” foram as que obtiveram o maior número de GNU identificadas ($\bar{x} = 102 \pm 58$ por genoma, excluindo a amostra CL9458 que apresentou valores extremos, 662 GNUs), seguida pelo ST412 ($\bar{x} = 85 \pm 32$). O ST963 e ST896 tiveram o mesmo número médio de unidades de novidade genética identificados, $\bar{x} = 81 \pm 27$ por genoma e, $\bar{x} = 81 \pm 15$, respectivamente. As amostras do ST78 foram as que obtiveram o menor número de GNU, $\bar{x} = 56 \pm 34$.

A amostra CL9458 pertencente ao ST253 e filogeneticamente distinta do restante das amostras (Seção 3.7 - Figura 2) obteve um total de 662 GNUs identificados. Dessas, 299 eram anotadas em no genoma como proteínas hipotéticas. Das 363 GNUs que apresentavam algum tipo de função, 41% (149/363) estavam presentes em todas as outras amostras incluídas no estudo.

Observando quais as proteínas obtiveram um escore GNU igual a zero, sendo então, consideradas como unidades de novidade genética, e os dados relacionados ao grupo de ST ao qual pertencia a amostra, foi possível observar um compartilhamento dessas novidades genéticas entre os diferentes grupos ST estudados. No total, 5.937 unidades novidades genéticas foram identificadas, sendo que 818 delas foram correlacionadas aos diferentes grupos de STs (Figura 16).

Figura 15 – Número de Unidades de Novidade Genética (GNU) nos 74 genomas de *Enterococcus faecium* estudados



Legenda: Cada ponto representa uma das 74 amostras; As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás: “Outros STs”.

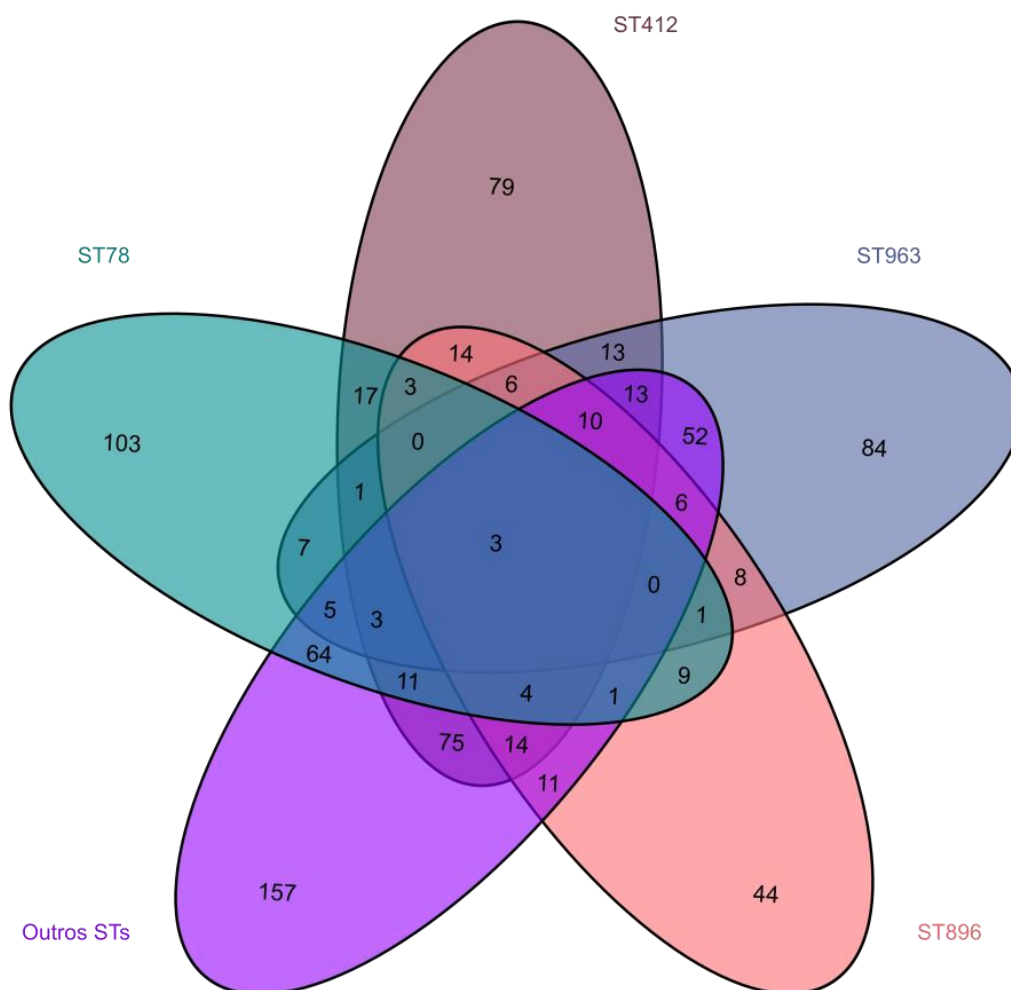
Fonte: A autora, 2020.

As amostras do grupo de “Outros STs” foram, como esperado, as que apresentaram o maior número de GNUs consideradas exclusivas, ou seja, que apareciam em mais de uma amostra do grupo e que não apareciam em nenhuma amostra dos outros quatro grupos (Figura 16). Foram identificadas 157 no total. O segundo maior grupo em frequência foi o do ST78, com 103 GNU identificadas, seguido pelo ST963 (84 GNU exclusivas) e ST412 (79 GNU exclusivas). O ST896 foi o que obteve o menor número de unidades de novidade exclusivas, 44 no total.

Além disso, foi observada uma variação no número de GNUs compartilhadas entre os grupos de STs de acordo com a maior ou menor proximidade filogenética. Entre as amostras do ST78 o número de GNU compartilhadas com grupos filogeneticamente mais distantes como o ST963 e ST896, foi de sete e nove, respectivamente, e somente uma GNU foi simultaneamente compartilhada entre esses três grupos. Já para o grupo ST412, foram observadas 13 e 14 GNUs compartilhadas com ST963 e ST896, respectivamente, considerando que esses grupos apresentaram uma mesma origem ancestral comum, de acordo com dados apresentados anteriormente. Além disso, seis GNUs foram compartilhadas

simultaneamente entre esses três grupos. Apenas três unidades de novidade genética foram compartilhadas entre os cinco grupos avaliados (Figura 16).

Figura 16 – Diagrama de Venn representando o compartilhamento de Unidades de Novidade Genética (GNU) entre os cinco grupos de STs estudados



Fonte: A autora, 2020.

Através da ferramenta WhatsGNU, também foi possível observar com quais genomas presentes no banco de dados de *E. faecium* construído para essa análise, as amostras incluídas no estudo obtiveram um maior número de correspondências (*matches*). Como foi possível identificar o ST ao qual pertenciam todas as amostras do banco de dados de referência, foi avaliada a relação entre as amostras deste estudo com outros genomas de *E. faecium* de diferentes origens. No Quadro 4, estão listados os dez genomas que cada grupo de ST obteve

maior correspondência, o ST ao qual pertencia esse genoma e número de amostras dentro do grupo que apresentavam essa correspondência.

Os genomas C68 (ST16) e D34RRF (ST25) foram os que apresentaram mais *matches* com amostras incluídas nesse estudo, independentemente do grupo ao qual pertenciam. As amostras do ST78 foram relacionadas com outras amostras do mesmo ST e do ST17 presentes no banco de dados. Já o grupo do ST412 deste estudo teve entre os dez genomas com mais correspondência, três do mesmo ST (VRE_13, 43834 e 2.1) e um do ST203. O ST963 apresentou relações extremamente próximas com genomas do ST736 presentes no banco: oito dos dez genomas com que as amostras do grupo obtiveram *match* pertenciam a este ST. As amostras do grupo do ST896 também apresentaram relação com o ST736, embora também tenham sido encontrados *hits* com outras amostras do ST896 (HBSHRP_10, HBSHRP_9).

Essa ferramenta também foi utilizada para avaliar as relações ancestrais da amostra CL7611, que demonstrou discrepâncias entre as amostras incluídas neste estudo. Conforme já apresentado, esta amostra pertence ao ST78 esteve mais relacionada filogeneticamente ao grupo ST412. Nessas análises, CL7611 apresentou correspondência com três genomas relacionados ao grupo do ST78 (UAA909, UAA1484, UAA952), pertencentes ao ST17. Além disso, CL7611 também apresentou correspondência com outros três genomas que não foram identificados nem no grupo do ST78 e nem no grupo do ST412, sendo dois desses genomas também pertencentes ao ST17 e um pertencente ao ST78. A amostra CL7611 não apresentou entre seu *top hits* de genomas presentes no bando de referência, nenhum genoma que tivesse sido exclusivo ao grupo do ST412.

Quadro 4 – Lista com os dez genomas presentes no banco do programa WhatsGNU que obtiveram mais *hits* com as amostras de *Enterococcus faecium* dos diferentes grupos de STs estudados

Top hits	Grupos														
	ST78 (N = 21)			ST412 (N = 19)			ST963 (N = 12)			ST896 (N = 9)			“Outros STs” (N = 13)		
	Genoma	ST	Nº	Genoma	ST	Nº	Genoma	ST	Nº	Genoma	ST	Nº	Genoma	ST	Nº
1º	C68	16	21	C68	16	19	C68	16	11	C68	16	9	C68	16	13
2º	D344RRF	25	21	D344RRF	25	19	D344RRF	25	11	D344RRF	25	9	D344RRF	25	13
3º	EnGen0376	736	21	EnGen0376	736	19	EnGen0312	736	11	EnGen0312	736	9	EnGen0376	736	13
4º	EnGen0377	736	21	EnGen0377	736	19	EnGen0376	736	11	EnGen0376	736	9	EnGen0377	736	13
5º	UAA909	17	21	VRE_13	412	19	EnGen0377	736	11	EnGen0377	736	9	ER04526.3A	736	12
6º	E1644	78	20	E2560	78	16	EnGen0323	736	10	HBSJRP_10	896	9	EnGen0312	736	5
7º	E2369	78	20	VRE_108	192	15	EnGen0375	736	10	HBSJRP_9	896	9	EnGen0323	736	5
8º	UAA1484	17	19	43834	412	11	ER04526.3A	736	9	ER04526.3A	736	7	E2560	78	3
9º	UAA952	17	19	2.1	412	8	EnGen0319	736	7	EnGen0323	736	7	EnGen0375	736	3
10º	UAA911	17	18	6600e06	203	8	EnGen0314	736	5	EnGen0375	736	6	VRE_108	192	3

Legenda: Número de amostras do grupo (N); *Sequence Type* (ST); Nº - número de amostras que tiveram *match* com o genoma em questão.

Fonte: A autora, 2020.

3.15 Estudo dos Eventos de Recombinação Homóloga

A taxa estimada de recombinação relativa à mutação (R / θ) para todo o genoma das 74 amostras de *E. faecium* incluídas no estudo foi de $0,0680 \pm 0,0027$, e o impacto relativo da recombinação para a mutação foi de 1.87. Esse resultado indica que ao longo do processo evolutivo, os eventos de recombinação homóloga introduziram duas vezes mais substituições nucleotídicas no genoma do que mutações pontuais.

A árvore reconstruída após a investigação dos eventos de recombinação (Figura 17B) apresentou braços mais curtos que os observados na árvore que não levava em conta os efeitos da recombinação (Figura 17A). Ao identificar e reconstruir os eventos de recombinação, foi possível observar que muitos se deram em nodos ancestrais da árvore de filogenia. Na Tabela 6, encontram-se os resultados em relação ao número de eventos de recombinação observados nesses nodos e à extensão desses eventos em tamanho e quantidade de pares de base que foram afetados. Estes resultados sugeriram que a quantidade de mudanças observadas anteriormente, demonstradas pelo tamanho dos braços na árvore, ocorreram, em geral, por efeitos de recombinação, principalmente, nos nodos que formam a estrutura da árvore.

O nodo 145 foi o que obteve o maior número de eventos de recombinação detectado, 12 no total. Entretanto, em termos da extensão desses eventos, o nodo 91 foi o que obteve a maior região atingida, com 100.427 pares de bases. Este nodo representou o último ancestral que deu origem ao grupo de amostras do ST963. O nodo 144, ancestral comum do ramo onde estava alocada a amostra CL7611 e do ramo que deu origem ao grupo do ST412/ST963/ST896, foi o que obteve o menor número de eventos de recombinação detectado, apenas um. Mesmo com quatro eventos detectados, o nodo 147 foi o que obteve a menor extensão de pares atingidos por recombinação (1.166 pb).

Outros nodos importantes para a explicação do surgimento de alguns STs que foram prevalentes no estudo também tiveram eventos de recombinação detectados, como, por exemplo, o nodo 146, último ancestral entre o braço onde estão as amostras do ST78 e as ramificações que deram origem ao ST412 e seus relacionados. Este nodo teve seis eventos de recombinação identificados, com um total de 16.665 pares de bases atingidos (Tabela 6).

Por outro lado, não foram encontrados eventos de recombinação no nodo ancestral, que dividiu o grupo onde estavam as amostras do ST412 do braço que, posteriormente, deu origem aos grupos ST963 e ao ST896 (Figura 17). Eventos de recombinação também foram detectados no nodo 120, ancestral que deu origem aos diferentes grupos do ST412, que emergiram no período avaliado. Nesse nodo, foram identificados três eventos recombinatórios que afetaram um total de 42.840 pares de base

Tabela 6 – Relação do número de eventos de recombinação e sua extensão nos diferentes nodos ancestrais da árvore de filogenia de *Enterococcus faecium*

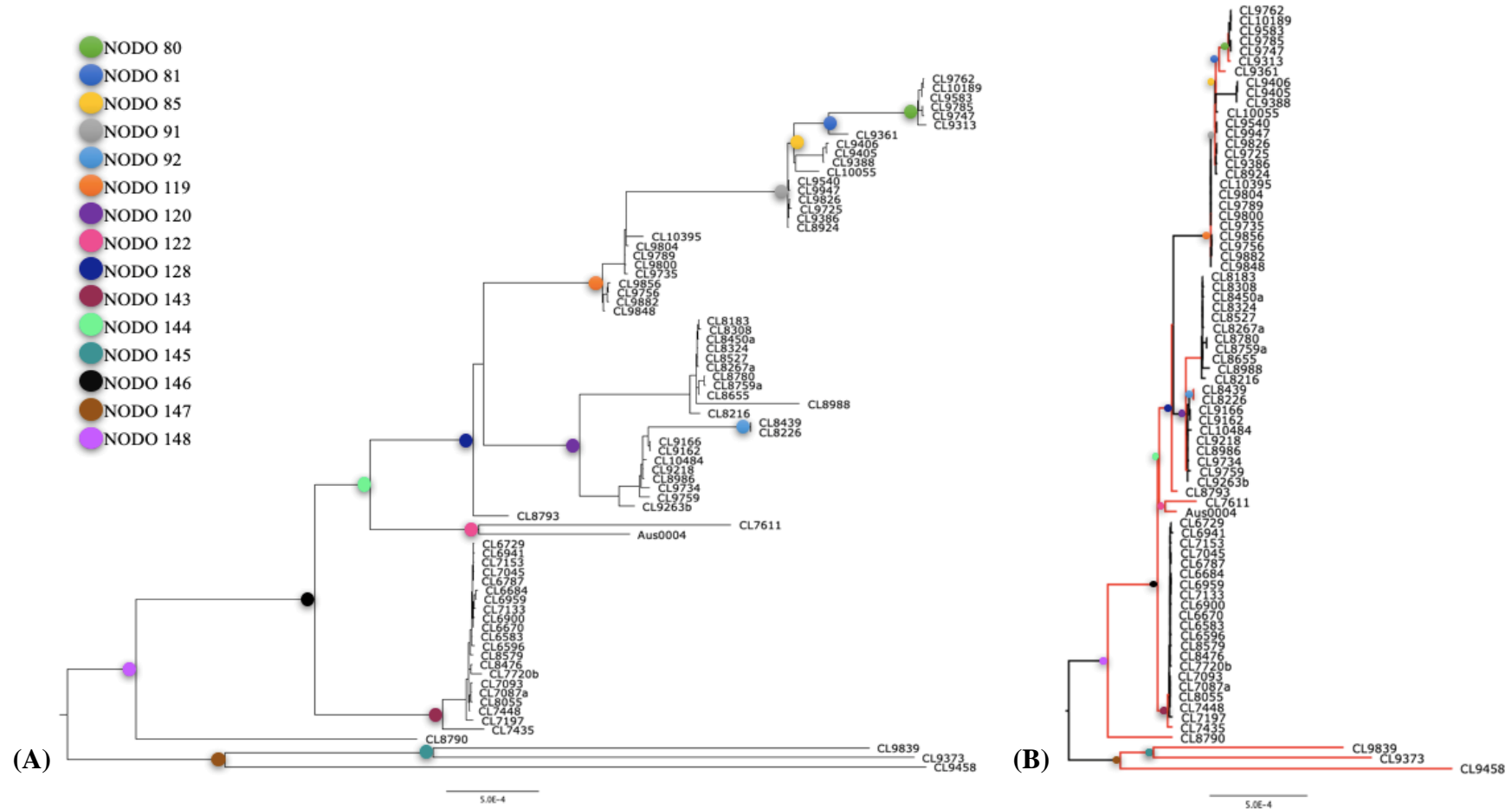
Nodos	No. de Eventos de Recombinação	Extensão dos Eventos de Recombinação (pb)
NODO 80	6	64.409
NODO 81	3	51.238
NODO 85	2	1.987
NODO 91	3	100.427
NODO 92	7	73.249
NODO 119	8	47.934
NODO 120	3	42.840
NODO 122	4	48.141
NODO 128	4	10.656
NODO 143	5	17.107
NODO 144	1	8.975
NODO 145	12	27.901
NODO 146	6	16.665
NODO 147	4	1.166
NODO 148	3	3.054

Legenda: pares de base (pb); Número (No)

Fonte: A autora, 2020.

Em relação ao número de eventos de recombinação observados no genoma das amostras, apenas CL9458, pertencente ao ST253, obteve mais de 30 eventos detectados em seu genoma. Outras duas amostras, CL9838 (ST97) e CL9373 (ST885), apresentaram entre 20 e 30 eventos de recombinação em seu genoma. A amostra CL7611, que até o momento apresentou diversas características conflitantes em relação ao ST ao qual pertence e às suas relações filogenéticas, obteve oito eventos de recombinação detectados em seu genoma, com um total de 64.558 pares de bases atingidos (Figura 18).

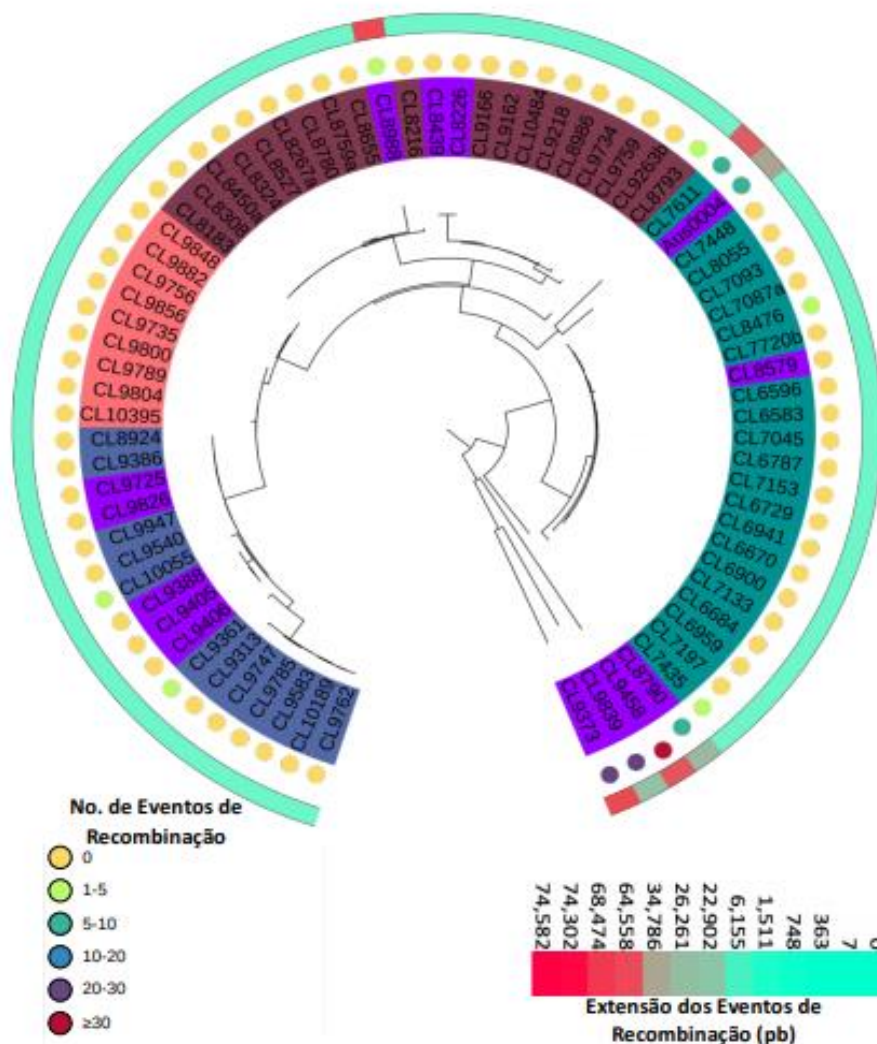
Figura 17 – Avaliação dos impactos de eventos de recombinação homóloga na reconstrução da filogenia de *Enterococcus faecium*. (A) Árvore de filogenia construída através Máxima Verossimilhança baseada no alinhamento SNPs e enraizamento pelo ponto médio (*midpoint rooting*) anterior a detecção das regiões de recombinação homóloga. (B) Árvore de filogenia construída através de Máxima Verossimilhança após serem identificadas as áreas de recombinação



Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás: “Outros STs; As bolinhas coloridas representam cada um dos nodos ancestrais onde foram detectados eventos de recombinação.

Fonte: A autora, 2020.

Figura 18 – Árvore de filogenia construída através Máxima Verossimilhança baseada no alinhamento de SNPs e enraizamento pelo ponto médio (*midpoint rooting*) demonstrando os eventos de recombinação homóloga em amostras de *Enterococcus faecium*



Legenda: Número de eventos de recombinação representado por círculos coloridos; Linha colorida mais externa representa a extensão dos eventos de recombinação em quantidade de pares de bases (pb), seguindo uma escala de cores do menor (verde) ao maior (vermelho) valor; As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, sendo verde – ST78; vermelho escuro – ST412, azul – ST963, rosa – ST896 e lilás – “Outros STs”.

Fonte: A autora, 2020.

Cabe ressaltar que, dentre as amostras estudadas, o maior número de pares de bases atingidos não foi diretamente relacionado ao maior número de eventos de recombinação identificados em um dado genoma. A amostra CL8988 apresentou um total de 74.582 pares de bases afetados em apenas três eventos de recombinação detectados (Figura 18). As amostras do ST963 e do ST78, que foram atingidas por eventos de recombinação, não

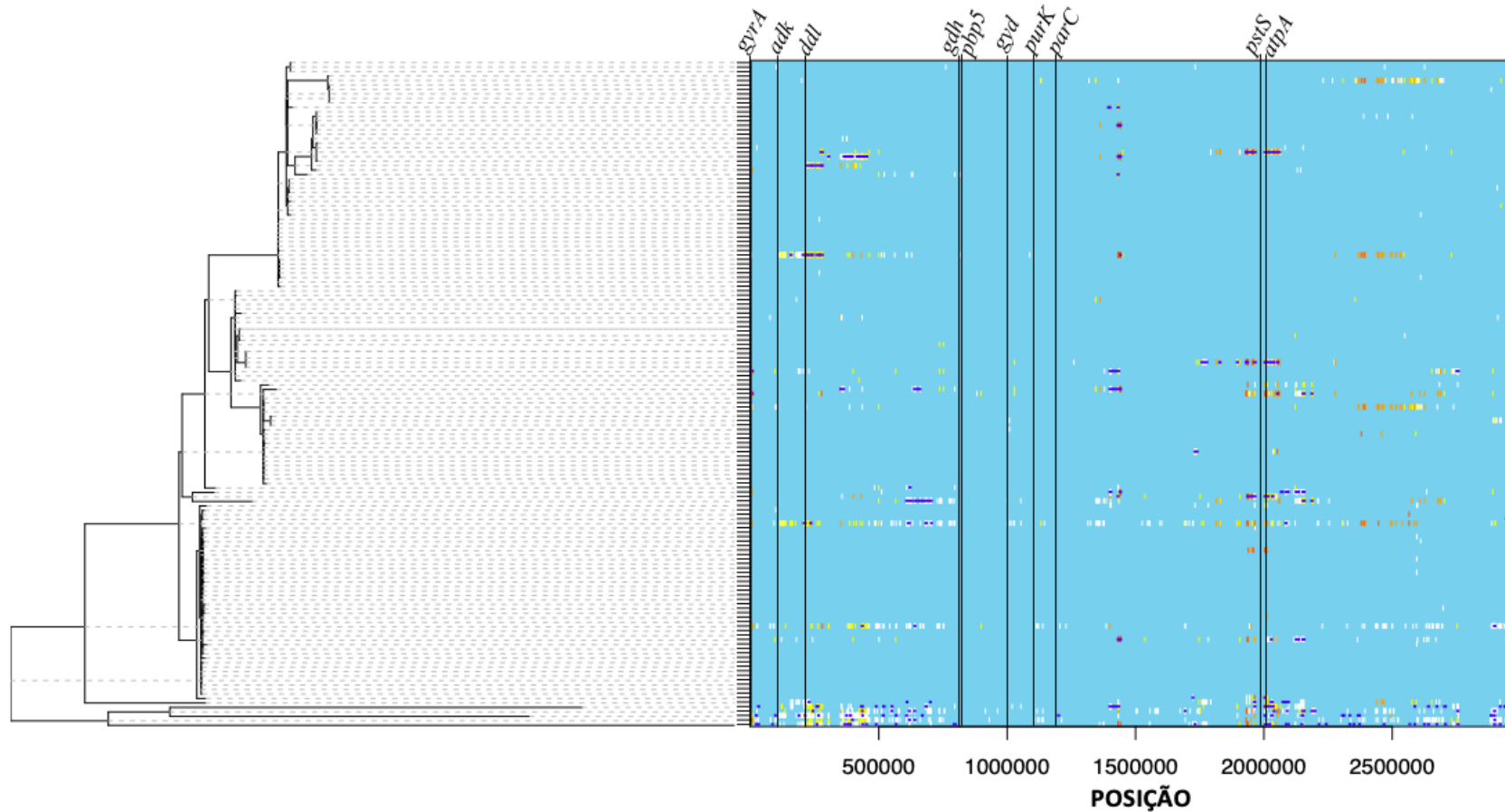
apresentaram comprometimento de grande parte do seu genoma. Outras variações incluem as duas amostras do ST963, em que uma apresentou um total de 363 pb atingidos por recombinação (CL9361), enquanto a outra o tamanho afetado foi de 6.155 pb (CL10055). Também, na outra amostra do ST78 incluída neste estudo, além da CL7611, foram identificados apenas sete pares de bases sob processo de recombinação.

Por fim, foi possível identificar que algumas regiões do genoma foram mais especificamente atingidas por processos de recombinação homóloga (Figura 19). Além disso, alguns genes de interesse nos objetivos propostos por este estudo, como os sete genes que compõem o esquema de MLST e os genes associados à resistência aos antimicrobianos como o *gyrA/parC* (resistência às quinolonas) e *pbp5* (resistência aos β -lactâmicos), foram também avaliados quanto a presença de regiões afetadas por eventos de recombinação.

Os resultados obtidos demonstraram que as regiões do genoma onde foi identificada a maior parte dos eventos de recombinação estavam entre as posições 50 milhões e 100 milhões, e na região acima da posição dois milhões. A maioria dessas regiões possuía sítios não polimórficos. Entretanto, quando polimórficos, apresentaram algum grau de semelhança estrutural, decorrente de paralelismo ou evolução convergente, sendo considerados sítios homoplásicos (Figura 19). Dentre os genes avaliados, apenas dois estiveram sob sítios de recombinação, *ddl* e *atpA*.

Em relação ao gene *ddl*, apenas os nodos 119 e 146 apresentaram eventos de recombinação associados. Todavia, cabe ressaltar que o nodo 119, como observado na Figura 17, representou o ancestral que deu origem ao grupo do ST896 e do ST963. Esses dois STs foram portadores do alelo 2 para o gene *ddl*, enquanto que o ST78 e o ST412 apresentaram o alelo 1. Eventos de recombinação na região onde se encontravam o gene *atpA* foram detectados em duas amostras (CL8790 e CL9458) e no nodo 146.

Figura 19 – Árvore de filogenia construída através de Máxima Verossimilhança pelo software ClonalFrameML e enraizamento pelo ponto médio (*midpoint rooting*) detalhando as regiões de recombinação inferidas para os ramos ancestrais e terminais dos 74 genomas de *Enterococcus faecium*



Legendas: Linhas verticais pretas representam o posicionamento dos sete genes do MLST (*adk*, *ddl*, *gdh*, *gyd*, *purK*, *pstS*, *atpA*) e três genes associados à resistência aos antimicrobianos (*gyrA*, *parC*, *pbp5*) em relação a referência Aus0004; Eventos de recombinação estão marcados em azul escuro; Sítios não polimórficos estão mostrados em azul claro; Sítios polimórficos apresentam uma escala de cor indicando o nível de homoplásia. Branco: não homoplásico, do amarelo ao vermelho: grau de homoplásia.

Fonte: A autora, 2020.

3.16 Evolução Molecular

Para este teste, foram escolhidos os genes que no esquema de MLST são os que diferem os grupos do ST78, ST412, ST896 e ST963, sendo eles *ddl*, *pstS* e *purK*. A avaliação foi realizada utilizando-se as sequências nucleotídicas de todas as amostras pertencentes a esses STs (61 ao todo). Os resultados demonstrados na Tabela 7 indicam que os três genes do esquema de MLST se apresentam sobre seleção neutra. Embora tenham obtido valores de D maiores que zero, os testes estatísticos indicaram que os resultados não foram significativos ($p > 0,05$), não sendo possível rejeitar a hipótese nula de estarem sob a influência de processos estocásticos.

Os genes associados à resistência às quinolonas (*gyrA* e *parC*) e o gene responsável pela resistência aos β -lactâmicos (*pbp5*) também tiveram sua evolução molecular avaliada. Para os genes *gyrA* e *parC*, foram incluídas as sequências de todas as amostras pertencentes ao ST78, ST412, ST963 e ST896, além de três amostras que não apresentaram mutações como referência para comparação. Para *pbp5*, bem como todas as amostras do estudo apresentaram mutações no gene, a sequência de referência X84860.1 foi incluída na análise. Os resultados obtidos indicaram que tanto o gene *gyrA* quanto o gene *pbp5* não estavam sobre processos de evolução neutra. Os resultados foram suportados estatisticamente, revelando que os dois genes tiveram valores de D estimados em 2,54 (*gyrA*) e -2,43 (*pbp5*).

Tabela 7 – Parâmetros estimados na determinação do valor de D de Tajima para cinco genes de *Enterococcus faecium* avaliados

Genes	S	η	K	Pi	Θ (por sequência)	Θ (por sítio)	Valor de D	p-valor
<i>ddl</i>	1	1	0.45	0.0009	0.21	0.0004	1.46	< 0.05
<i>pstS</i>	18	18	5.69	0.01	3.84	0.006	1.68	> 0.05
<i>purK</i>	1	1	0.45	0.0009	0.21	0.0004	1.46	> 0.05
<i>gyrA</i>	24	27	10.25	0.0041	5.69	0.0023	2.54	< 0.05
<i>parC</i>	11	11	1.33	0.0005	2.34	0.0009	- 1.21	> 0.05
<i>pbp5</i>	108	109	7.03	0.0034	23.37	0.011	- 2.43	< 0.05

Legenda: S – número de sítios polimórficos; η (Eta) – número total de mutações; K – número médio de diferenças nucleotídicas; Pi – diversidade nucleotídica; Θ – Theta.

Fonte: A autora, 2020.

3.17 Estimativa do Tempo de Divergência Clonal

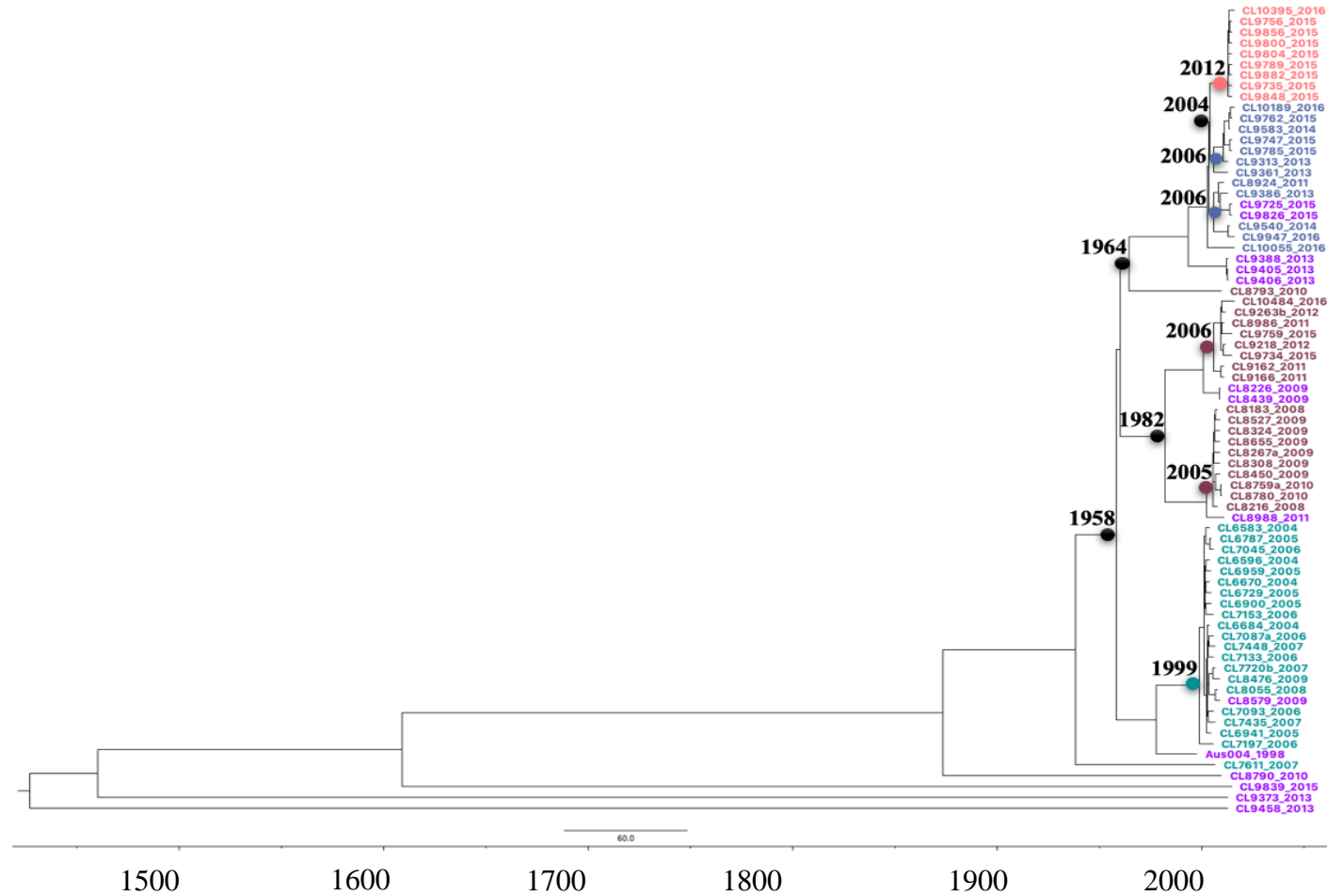
Através da reconstrução da filogenia pelo método bayesiano, foi possível fazer uma análise temporal dos processos de divergências clonal observados e estimar o relógio molecular. A taxa de mutação foi estimada em 2.34×10^{-7} SNPs por sítio por ano, utilizando-se 95% de HPD (do inglês, *highest posterior density*).

A árvore filogenética construída confirmou a divisão das amostras em diferentes grupos correlacionados ao ST ao qual pertenciam (Figura 20). O ancestral comum mais recente (MRCA – do inglês *Most recent common ancestor*) foi estimado ter ocorrido no ano de 2012. As amostras do ST963, que em outras reconstruções não se mostraram como um grupo homogêneo, apresentaram dois MRCA distintos, porém, datados, ambos, do ano de 2006. A divergência entre esses dois grupos teria ocorrido no ano de 2004.

O ST412 teve sua emergência confirmada em dois períodos distintos, sendo a estimativa de divergência datada do ano de 1982 (Figura 20). O ancestral comum mais recente que levou à 1ª emergência surgiu no ano de 2005; já o responsável pela 2ª emergência surgiu em 2006. A divergência entre o ST412 e os outros dois ST mais recentes, mencionados anteriormente, ST963 e ST896, ocorreu no ano de 1964. Por fim, a divergência entre o ST78 e o ST412 foi estimada quatro anos antes, em 1958. O ST78, cujos resultados anteriores já haviam demonstrado se tratar de um grupo bastante homogêneo, possuindo uma origem monofilética, teve seu MRCA estimado em 1999.

Os resultados obtidos pela análise bayesiana demonstraram claramente ondas de emergência de diferentes linhagens. Na Figura 21, reconstruímos essa sucessão de STs observadas. A emergência do ST78 se iniciou no ano de 1999, perdurando por dez anos, até 2009. Este foi o último ano de registro de amostras pertencentes a esse ST neste estudo. Os dados obtidos com as amostras avaliadas demonstraram que a emergência do ST412 ocorreu ainda durante o período de circulação do ST78. Contudo, o pico de observação da primeira emergência se deu em 2009, quando então chegou ao seu fim. A segunda emergência ocorreu principalmente entre 2011 e 2012, com as amostras mais recentes isoladas no estudo datadas nos anos de 2015 e 2016.

Figura 20 – Árvore filogenética reconstruída através de análise bayesiana estimando o tempo de divergência entre as diferentes linhagens de *Enterococcus faecium* isolados no período de 2004 a 2016

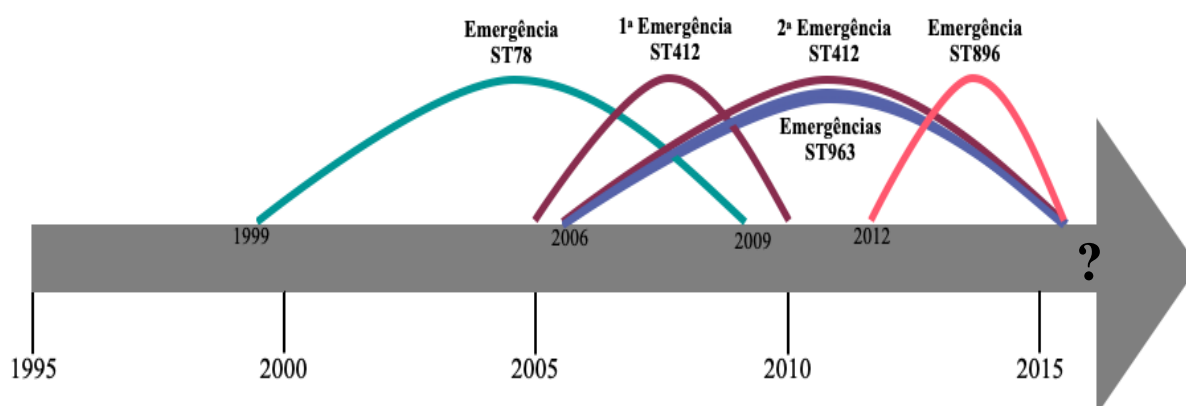


Legenda: As amostras foram coloridas de acordo com o grupo de ST ao qual pertencem, verde - ST78, vermelho escuro - ST412, azul - ST963, rosa - ST896 e lilás: “Outros STs”; Círculos coloridos demarcam o nodo onde o ano de divergência foi estimado.

Fonte: A autora, 2020.

A emergência do ST963 data de 2006 e amostras pertencentes a este ST foram identificadas até 2016. Durante todo o período, esse grupo de ST foi observado principalmente entre 2011 e 2015, sendo o único ST a circular entre os anos de 2013 e 2014. No período de 2015 a 2016, obtivemos três STs circulando concomitantemente, com expansão e pico máximo de observação do ST896. O surgimento deste ST ocorreu em 2012.

Figura 21 – Linha do tempo reconstruindo os eventos de emergência de linhagens de *Enterococcus faecium* em instituições hospitalares do estado do Rio de Janeiro



Fonte: A autora, 2020.

Observando dados compilados ao longo do estudo e usando testes estatísticos, foi possível associar algumas características aos diferentes grupos de STs, o que seria sugestivo de que alguns tipos de eventos ocorridos no genoma das amostras de *E. faecium* possam explicar a emergência e sucessão dessas linhagens (Tabela 8).

O grupo do ST78 obteve a menor média de SNPs observada, $\bar{x} = 4.895$ SNPs por amostra, sendo estatisticamente correlacionado a uma menor presença de variação. Por outro lado, este grupo obteve o maior número médio de IS preditas ($\bar{x} = 41,61$) em relação aos outros grupos. O valor de p confirmou que esta característica esteve associada as amostras desse ST. Uma outra característica relacionada a este grupo foi a presença de profagos, sendo este o grupo com o maior número médio de profagos observados. Entretanto, quando avaliado o número de plasmídeos presentes em cada amostra, o ST78 obteve a menor média estimada, $\bar{x} = 2,42$ plasmídeos por amostra.

Já as amostras do ST412 obtiveram o maior número de CDS preditas (2.792); porém, correlações significativas só foram encontradas em relação ao ST963. A presença de genes de resistência foi fortemente associada ao ST896, que obteve a maior média. Este grupo também obteve valores significativos em relação à presença de plasmídeos, exceto em relação ao ST963.

Tabela 8 – Métricas obtidas para diversos parâmetros estudados no genoma de *Enterococcus faecium* e suas correlações entre os diferentes grupos de ST predominantes

Parâmetros	Grupos (Média ± DV)			
	ST78	ST412	ST963	ST896
SNPs	4.895* ± 374	6.069 ± 696	5.981 ± 569	5.773 ± 166
CDS preditos	2.684 ± 236	2.792* ± 41	2.664 ± 162	2.765 ± 30
Genes de Resistência	8,09 ± 1,13	8,31 ± 1,45	8,41 ± 1,64	10,11* ± 1,45
Genes de Virulência	10,23 ± 1,22	10,68 ± 0,67	10,66 ± 0,98	10,77 ± 0,97
Presença de IS	41,61* ± 13,5	26,84 ± 6,24	30,91 ± 6,65	29,66 ± 5,5
Presença de Profagos	1* ± 0,31	0,63 ± 0,59	0,33 ± 0,49	0
Presença de Plasmídeos	2,42* ± 1,07	3,42 ± 1,12	4,66 ± 0,77	5* ± 1,65

Legenda: *Single nucleotide polymorphism* (SNP); *Coding sequence* (CDS); *Insertion sequence* (IS); Desvio padrão (DV); ± (mais ou menos); * indicam valores de $p < 0,05$.

Fonte: A autora, 2020.

4 DISCUSSÃO

A espécie *Enterococcus faecium* constitui um dos membros do gênero *Enterococcus* que pertence a microbiota do trato gastrointestinal humano e que nas últimas décadas ganhou notório destaque como um dos mais importantes agentes de infecções relacionadas à assistência à saúde (IRAS), não apenas dentro do gênero, como em relação ao grupo de bactérias Gram-positivas. Sua escalada de comensal a patógeno altamente adaptado ao ambiente hospitalar está associado a um processo evolutivo dinâmico, envolvendo sua grande plasticidade genômica, que permitiu com que gradualmente adquirisse diversos mecanismos de resistência aos antimicrobianos e de virulência, junto com fenômenos de recombinação homóloga e expansão populacional.

Este estudo teve por objetivo abordar sob uma perspectiva genômica, com a utilização de sequenciamento do genoma completo e ferramentas de biologia computacional, os processos evolutivos que levaram ao surgimento e expansão de populações clonais que circularam em instituições hospitalares do estado do Rio de Janeiro no período de 2009 a 2016. E se deu, principalmente, após as primeiras observações em estudos de nosso grupo (FARIA, 2012), que apontou a substituição da linhagem epidêmica circulante, até então o ST78, pelo ST412.

O estudo foi composto por 205 amostras oriundas de 31 instituições hospitalares. A caracterização do perfil de suscetibilidade, realizada pelo método de disco-difusão, revelou elevados percentuais de amostras não suscetíveis às mais variadas classes de antimicrobianos. Foi constatado que na amostragem avaliada, para ampicilina, ciprofloxacina, eritromicina e penicilina, as frequências de não suscetibilidade chegaram a ser superiores a 95%. Cabe ressaltar que elevados percentuais de resistência para *E. faecium* se mostram, via de regra, extremamente comuns. Uma meta-análise realizada por Jabbari et al. (2019) demonstrou que a prevalência de amostras de *E. faecium* isoladas de infecção de corrente sanguínea apresentando taxas de resistência a ampicilina era superior a 80% em quase todas as regiões do globo, exceto no Sudeste Asiático (50,5%) e no Oriente Médio (64,5%).

Por outro lado, para antimicrobianos como a gentamicina, o percentual de amostras não suscetíveis a níveis elevados foi de 39,5% neste estudo, enquanto que no de Jabbari et al. (citado anteriormente), os autores observaram percentuais variando entre 28% no continente americano até 65,5% na região do Pacífico Ocidental, que inclui países como Austrália, Nova Zelândia, Japão, China e Vietnã. Considerando-se a resistência à vancomicina, sabe-se que a

frequência de amostras VREfm circulantes é bastante variável de acordo com a região geográfica. Na Europa, a taxa de amostras VREfm no período de 2009 a 2014 variou entre 15% a 30%, enquanto nos Estados Unidos, no mesmo período, essa taxa ficou entre 70% e 80% (MENDES et al., 2016).

A resistência à vancomicina é uma característica que contribui para o destaque da espécie *E. faecium* como importante agente de infecções humanas. Em nosso estudo, visto que o conjunto de amostras avaliadas foi selecionado a partir de uma coleção de culturas, não seria possível extrapolar as taxas de resistência inferidas para o cenário local. Entretanto, as amostras foram provenientes de diferentes instituições, apontando para a ampla dispersão de VREfm no estado do Rio de Janeiro. Foram observados percentuais de resistência à vancomicina $\geq 80\%$ no período avaliado, excetuando-se apenas os anos de 2011 e 2012 (36,4% e 42,8%, respectivamente). Cabe ressaltar que nesse período foi quando observamos importantes mudanças em termos de estrutura populacional, com o pico de circulação da segunda onda do ST412 e do ST963.

A correlação da resistência à vancomicina com a presença de resistência a outras classes de antimicrobianos, como β -lactâmicos e quinolonas, observada no estudo, deve ser discutida e entendida por uma lógica reversa. Na realidade, a presença da resistência aos β -lactâmicos e quinolonas foi fundamental para o surgimento de amostras VREfm, pois permitiu o estabelecimento desse microrganismo no ambiente hospitalar. Alguns trabalhos na década de 1990 já alertavam que seria impossível evitar pressões seletivas em favor do surgimento de amostras multirresistentes, uma vez que paralelamente ao crescente aumento da frequência de cepas carregando resistência à vancomicina, linhagens endêmicas de *E. faecium* resistentes à ampicilina e fluoroquinolonas já eram comuns no período (ELIOPOULOS, 1997; HANRAHAN; HOYEN; RICE, 2000; SUPPOLA et al., 1999).

Globalmente, amostras VREfm carregando o gene *vanA* são mais comuns, enquanto que a circulação de *vanB* tem se mostrado mais restrita a algumas regiões, como a Austrália e, mais recentemente, com relatos de expansão na Alemanha (BENDER et al., 2016b; COOMBS et al., 2014; FALGENHAUER et al., 2019; FREITAS et al., 2016; MAHONY et al., 2018; XANTHOPOULOU et al., 2020). No Brasil, a predominância de amostras VREfm carregando o genótipo *vanA* já foi amplamente reportada (CAMARGO et al., 2004; CONCEICAO et al., 2011; KERBAUY et al., 2011; MERLO; DABUL; CAMARGO, 2015; PALAZZO et al., 2011; RESENDE et al., 2014; RUZON et al., 2010; SACRAMENTO et al., 2017; SILVA et al., 2012). Neste estudo, também não foram identificadas amostras portadoras do genótipo *vanB*, reforçando a hegemonia de *vanA* em nosso meio.

A disseminação de determinantes genéticos de resistência à vancomicina está relacionada à disseminação de elementos genéticos associados ao fenótipo HLAR. O megaplasmídeo pMGI tem sido apontado como responsável pela cotransferência do operon *vanSRHAXYZ*, contido no Tn1546 juntamente com o gene *aac(6')-Ie-aph(2'')*, que codifica a resistência a níveis elevados de gentamicina (HLGR) e está contido no Tn5281 (BEHNOOD et al., 2013; HEGSTAD et al., 2010; TOMITA et al., 2002; 2003). O percentual de amostras HLAR no presente estudo foi de 45,5%, com uma predominância maior de amostras do fenótipo HLGR. Outros estudos já indicaram que este é o fenótipo predominante entre amostras de *E. faecium* HLAR (HAGHI; LOHRASBI; ZEIGHAMI, 2019; PADMASINI; PADMARAJ; SRIVANI-RAMESH, 2014).

De fato, a estreptomicina foi amplamente utilizada em esquemas terapêuticos combinada à penicilina no período de 1945 a 1970 (NOSKIN et al., 1991), quando então, foi substituída por outros aminoglicosídeos como a amicacina, gentamicina e tobramicina, que surgiram como alternativas ao tratamento de cepas que já apresentavam resistência contra alguns aminoglicosídeos anteriores (MINGEOT-LECLERCQ; GLYPCZYNSKI; TULKENS, 1999). Dados do período de 2002 a 2009 relataram que em hospitais dos EUA a gentamicina representou 63% do uso de aminoglicosídeos, seguida pela tobramicina (28%) e amicacina (9%), indicando uma maior pressão seletiva em favor desses antimicrobianos (ABABNEH et al., 2012). Estes dados podem explicar a maior observação do fenótipo HLGR entre amostras de *E. faecium*, uma vez que a enzima bifuncional AAC(6')-APH(2'') confere resistência a níveis elevados para todos antimicrobianos da classe, exceto à estreptomicina (RAMIREZ; TOLMASKY, 2010).

O gene *aac(6')-Ie-aph(2'')* é descrito como prevalente entre amostras portadoras do fenótipo HLGR em diversas partes do mundo, tanto em amostras clínicas humanas quanto em amostras de origem animal e ambiental, da mesma forma que *ant(6')-Ia* em amostras HLSR (EL-MAHDY; MOSTAFA; EL-KANNISHY, 2018; KLIBI et al., 2006; NIU et al., 2016; NOVAIS et al., 2005; NOWAKIEWICZ et al., 2017; PETSARIS et al., 2005; PEYVASTI et al., 2020; RATHNAYAKE; HARGREAVES; HUYGENS, 2011; WOODFORD et al., 1993). Neste estudo, foi observado que o gene que codifica para a enzima bifuncional foi identificado em todas as amostras HLGR; e *ant(6')-Ia* foi identificado na frequência de 97,4%.

Um elevado percentual de amostras de *E. faecium* resistentes à ciprofloxacina também foi identificado neste estudo. Esta característica de resistência também é comumente descrita como estreitamente relacionada a amostras mais adaptadas ao ambiente hospitalar (LEAVIS

et al., 2006; WERNER et al., 2010). Alguns autores já relataram que amostras de *E. faecium* resistentes a ciprofloxacina mesmo que de origens diversas (“não-humana”) foram pertencentes a clones prevalentes no ambiente hospitalar (FREITAS, A.A et al., 2018; KIM et al., 2018).

Cabe, também, destacar a identificação de amostras não suscetíveis a antimicrobianos considerados como opções terapêuticas mais recentes para o tratamento de infecções graves (como endocardites) causadas por amostras de *E. faecium* resistentes à vancomicina. Neste contexto, foi possível observar que 8,8% das amostras incluídas neste estudo apresentaram esta característica para a associação quinupristina/dalfopristina. Estes antimicrobianos são combinados na proporção de 30% de quinupristina, uma estreptogramina B, e 70% de dalfopristina, uma estreptogramina A. Para o tratamento das infecções por *E. faecium*, quinupristina/dalfopristina é comumente administrada juntamente com altas doses de ampicilina (BETHEA; WALKO; TARGOS, 2004; NIGO, 2014; THOMPSON; LAVIN; TALBOT, 2003). Dados de outros países já apontaram por exemplo que, em 2004, o percentual de amostras resistentes à Q/D na Coreia do Sul foi de 10% (OH et al., 2005). Na Bélgica, um estudo realizado em 2007, cinco anos após as estreptograminas serem banidas do país como promotores de crescimento animal, o percentual de amostras humanas resistentes à Q/D foi de 23% e de amostras suínas, 34% (DE GRAEF et al., 2007). Adicionalmente, foi relatado que o gene *erm(B)*, um determinante de resistência amplamente disseminado em amostras de enterococos, bem como na amostragem que compôs este estudo, possa estar associado a eliminação do efeito bactericida desse antimicrobiano (DONABEDIAN et al., 2006; LÓPEZ et al., 2010; WANG et al., 2016). Esses relatos são preocupantes por indicar uma maior limitação ao tratamento de infecções causadas por VRE.

Ao todo, observamos que 91,7% de nossas amostras apresentavam resistência a pelo menos um agente de três ou mais classes de antimicrobianos distintas, sendo caracterizadas de como MDRs de acordo com os critérios estabelecidos neste estudo e baseados nas recomendações de Magiorakos et al. (2012). Amostras *E. faecium* MDR estão amplamente disseminadas pelo mundo e são a razão pela qual essa espécie faz parte das principais listas criadas por organizações e centros de controle de doenças, em relação ao problema da resistência aos antimicrobianos. Em um estudo de Freitas et al. (2018b) em Portugal, foi observado que todas as amostras de *E. faecium* isoladas em clínicas de cuidado prolongado apresentaram resistência a pelo menos três antimicrobianos de classes distintas.

No Brasil, um estudo de Santos et al. (2017), conduzido em pacientes portadores de neoplasias, o percentual de amostras da espécie apresentando o fenótipo MDR foi de 88,9%.

A circulação de amostras de *E. faecium* multirresistentes não é restrita ao ambiente hospitalar, pois estas já foram identificadas a partir de diferentes origens como, porcos e aves, dentre outros (BERTELLONI et al., 2015; CHAJĘCKA-WIERZCHOWSKA; ZADERNOWSKA; ŁANIEWSKA-TROKENHEIM, 2016; JOHNSTON; JAYKUS, 2004; NOVAIS et al., 2013).

O acompanhamento contínuo das variações dos percentuais de amostras resistentes se faz necessário para uma melhor avaliação global do cenário momentâneo e prospecção futura dos efeitos relacionados as infecções por essa espécie. Todavia, o que permeia o entendimento e a discussão da evolução de *E. faecium*, e de qualquer outro microrganismo multirresistente, passa pela reflexão sobre o impacto da utilização de antimicrobianos. Em um estudo realizado com dados obtidos da venda de antibióticos em farmácias da rede privada no Brasil, entre o período de 2013 a 2016, os autores indicaram que houve um crescimento do consumo de antimicrobianos no país em torno de 18%. No estado do Rio de Janeiro, a média foi superior ao observado nacionalmente, com um crescimento de 20%. Por outro lado, em países da Europa a taxa de aumento foi de 2%. No ano de 2016, somente em nosso estado, mais de 5 milhões de unidades, definidas como caixas ou ampolas, foram vendidas em farmácias da rede privada. Ao todo, na região sudeste 30,5 milhões de unidades foram vendidas no ano. Dentre as classes mais consumidas, estavam os β -lactâmicos e macrolídeos (NEVES CASTRO et al., 2020). Alguns dados mundiais também demonstraram um preocupante aumento no consumo de antimicrobianos em países considerados, de acordo com o Banco Mundial, como de renda média-alta (KLEIN et al., 2020).

É amplamente reconhecido que a pressão seletiva exercida pelo uso de antimicrobianos favorece a circulação de amostras multirresistentes e adaptadas ao ambiente hospitalar. Para a espécie *E. faecium*, essas amostras também apresentam o gene *esp*, que codifica a proteína de superfície enterocócica. O surgimento de populações de *E. faecium* adaptadas ao ambiente hospitalar pode ser descrito como um processo de múltiplos passos, que já foi chamado de “capitalismo genético” ou efeito de Methew, no qual a aquisição de mecanismos que traz vantagens seletivas (como por exemplo, a presença do gene *esp* e a resistência à ampicilina) permite a aquisição de mais elementos adaptativos (como por exemplo, o gene *hyl*) (BAQUERO; COQUE; CANTON, 2003; KLARE et al., 2005; RICE et al., 2003).

Entretanto, como também foi observado neste estudo, não parece existir uma correlação direta entre a presença de genes de virulência (como *esp* e *hyl*) e a resistência à vancomicina. Como já mencionado, no estabelecimento inicial desse microrganismo no ambiente hospitalar predominaram as amostras portadoras de *esp* e resistentes à ampicilina e

ciprofloxacina (BALDASSARRI et al., 2001; LEAVIS et al., 2003; WILLEMS et al., 2001). Dessa forma, a resistência à vancomicina não parece ter sido a condição *sine qua non* para o ganho do ambiente hospitalar por *E. faecium*. Entretanto, a identificação inicial de clones albergando essa característica de resistência foi indicativa de extrema preocupação, uma vez a pressão seletiva exercida pelo uso de vancomicina resultaria, como foi de fato observado, na sua disseminação e ampla circulação de amostras resistentes em curto período de tempo (COQUE et al., 2005).

A utilização de técnicas de sequenciamento do genoma completo fornece uma alta capacidade discriminatória para o entendimento de aspectos evolutivos. Atualmente, essa metodologia já ultrapassou barreiras inicialmente apontadas, como o custo e tempo de execução elevados. Por exemplo, o sequenciamento de um genoma bacteriano pela plataforma Illumina MiSeq (Illumina Inc.) custa em torno de \$25 dólares, o que para muitos laboratórios do mundo representa uma boa relação de custo-benefício. Além disso, nas novas plataformas o tempo médio necessário para sequenciar um genoma é de 24 horas, sendo que, em alguns dispositivos, como Nanopore (Oxford), esse processo pode ser feito em até duas horas (DIDELOT et al., 2012; MEDINI et al., 2008).

A partir do primeiro genoma completamente sequenciado de *E. faecium*, publicado por Lam et al. (2012), referente a amostra Aus0004, alguns trabalhos utilizando WGS investigaram aspectos relacionados à evolução da espécie (GOULIOURIS et al., 2019; LEBRETON et al., 2013), bem como a diversidade de amostras hospitalares em relação às amostras comensais (GALLOWAY-PEÑA et al., 2012; RAVEN et al., 2016), além de aspectos epidemiológicos relacionados a surtos e eventos de transmissão hospitalar (BRODRICK et al., 2016; LEONG et al., 2018; RAVEN et al., 2017). Em um estudo de 2016 realizado na Austrália, o sequenciamento do genoma completo foi fundamental para caracterizar o surgimento e disseminação de amostras não-tipificáveis pela técnica de MLST, devido a perda do gene *pstS* (CARTER et al., 2016).

Neste estudo, foram selecionadas 52 amostras a partir dos dados obtidos por métodos de caracterização fenotípica e genotípica. Além destas, outras 22 amostras pertencentes ao estudo de Faria (2012), referentes ao período de 2002 a 2008, foram incluídas para fins comparativos, uma vez que representavam a linhagem epidêmica anteriormente circulante em nossas instituições hospitalares, ST78, e as duas primeiras amostras observadas como pertencentes ao ST412.

Considerando-se os dados obtidos com as análises iniciais a partir de WGS, foi determinado que o tamanho médio do genoma para as 74 amostras foi estimado em 2,76Mb,

com identificação média de 2.715 CDS por genoma. Dados disponíveis na literatura especializada indicaram que no gênero *Enterococcus*, o tamanho do genoma é variável de acordo com a espécie. Assim, já foram registrados resultados que variaram de 2,31Mb para amostras da espécie *Enterococcus sulfureus* até 5,27Mb para *Enterococcus pallens*, e os genes preditos no genoma entre 2.154 até 5.107 (ZHONG et al., 2017).

Entretanto, o tamanho do genoma e a quantidade de CDS identificadas até mesmo dentro de uma mesma espécie como *E. faecium* também podem variar, particularmente, entre amostras de origens distintas. No estudo realizado por Zhong et al. (2019), com amostras de *E. faecium* de diversas origens, foram encontradas diferenças estatisticamente significativas em relação ao tamanho do genoma e a quantidade de genes. As amostras de *E. faecium* isoladas de humanos apresentaram um tamanho médio do genoma de 3Mb; enquanto nas obtidas de porcos e frangos foi de 2,70 Mb e 2,77 Mb, respectivamente, e de laticínios de 2,69Mb. O mesmo ocorreu em relação a quantidade de CDS anotadas; os autores relataram que as amostras humanas tiveram uma média de 2.866 CDS preditos e para as amostras oriundas de frangos e porcos, esse valor foi de 2.574 e 2.519, respectivamente (ZHONG et al., 2019).

Um outro parâmetro importante relacionado ao genoma bacteriano é o seu conteúdo de G+C. No presente estudo, o conteúdo G+C médio foi de 37,75% entre as amostras analisadas, o que está totalmente de acordo com a descrição para a espécie na plataforma NCBI (37,8%). Mais do que uma métrica de comparação e análise em estudos de sequenciamento do genoma, o percentual de conteúdo G+C bacteriano tem impactos importantes nos eventos de transferência horizontal de genes. Em geral, o DNA transferido horizontalmente, como DNA plasmidial e de fagos, apresenta um baixo conteúdo G+C, com diferenças menores do que 10% em relação ao genoma hospedeiro (NISHIDA, 2012; ROCHA; DANCHIN, 2002). Estes achados sugerem que, usualmente, bactérias não adquirem DNAs com conteúdo G+C superior ao seu próprio, uma vez que não conseguem manter e nem regular elementos genéticos com conteúdo muito diferente dos seus (NISHIDA, 2013).

A mudança na epidemiologia das subpopulações de *E. faecium* circulantes no ambiente hospitalar, com a emergência de diferentes STs, tem sido amplamente relatada. Neste estudo, as análises da estrutura populacional de amostras de *E. faecium* circulantes no estado do Rio de Janeiro, obtidas através da identificação dos STs no genoma completo, demonstraram uma evolução dinâmica com a emergência e expansão de diversas linhagens. Foi constatado que depois do período de circulação do ST78, sendo 2009 o ano de isolamento da última amostra analisada, observou-se a emergência e circulação da linhagem do ST412.

Esses dados corroboraram achados anteriores de trabalhos de nosso grupo (FARIA, 2012). Ademais, no período avaliado foram também observadas a emergência e a expansão de outras duas linhagens, a do ST963 e a do ST896.

Em diversos países existem relatos que atestam o dinamismo da espécie em relação ao aparecimento e rápida circulação de linhagens clonais específicas, levando inclusive ao desaparecimento de subpopulações anteriormente dominantes. Na Dinamarca, estudos epidemiológicos utilizando WGS indicaram que a partir de 2014 houve a substituição das linhagens até então circulantes, pertencentes ao ST117 e ST192, pelas associadas aos ST203 e ST80 (PINHOLT et al., 2019). Na Austrália e na Suíça, a linhagem associada ao ST795 tem sido considerada, nos últimos anos, endêmica em diversas instituições hospitalares e associadas ao acontecimento de surtos (MAHONY et al., 2018; WASSILEW et al., 2018). Em Taiwan, a vigilância epidemiológica nacional revelou a emergência e expansão do ST414 a partir do ano de 2012 (KUO et al., 2014).

Em nosso estudo, o ST412 foi responsável pela dominação do ambiente hospitalar a partir do ano de 2009. De acordo com o que pode ser avaliado na literatura, o primeiro relato de troca populacional associada à emergência desta linhagem envolveu um estudo com amostras de *E. faecium* isoladas de instituições hospitalares do Canadá (MCCRACKEN et al., 2013). Os autores observaram que a partir de 2006 houve a substituição das linhagens até então circulantes (ST16, ST17, ST154 e ST80), por aquelas pertencentes ao ST412, ST203 e ST584. A partir de então, outros estudos surgiram identificando alterações na prevalência de tipos ST, com a emergência do ST412 em diferentes regiões do mundo.

Em hospitais terciários da cidade de Nova York, o ST17 representou 76,7% das amostras de *E. faecium* entre 1994 e 1995. Entretanto, no período de 2009 a 2012 foi observada a emergência do ST736 (60% das amostras), seguido pelo ST412 (13,9%). Em 2013, o percentual de amostras pertencentes ao ST412 já era de 20,7% (WANG et al., 2018). Em um estudo multicêntrico com diferentes países da América Latina, o ST412 foi considerado predominante em amostras isoladas entre 2006 a 2008 (PANESSO et al., 2010). No período entre 2009 a 2014, 44,4% das amostras de *E. faecium* de diferentes países do Caribe pertenciam ao ST412 (AKPAKA et al., 2017). Em países como Grécia, Hungria, Japão, México e Suíça, a circulação de amostras pertencentes a este ST também já foi reportada (ABDELBARY et al., 2019; DAMANI et al., 2010; FRANYÓ et al., 2018; MAYORAL-TERÁN et al., 2020). Além disso, ST412 já foi descrito relacionado a amostras oriundas do trato gastrointestinal de corvos e também em superfícies inanimadas em instituições hospitalares (MICHAEL; NO; ROBERTS, 2017; ROBERTS et al., 2016).

No Brasil, em um trabalho de 2012, Silva et al., avaliando amostras de *E. faecium* isoladas no período de 2008 a 2010, em um hospital da cidade de São Paulo, demonstraram que 38,7% pertenceram ao ST412 e que o ST78 não foi observado. Um estudo temporal publicado por Sacramento et al. (2017), demonstrou que na cidade de São Paulo ocorreu uma mudança generalizada na epidemiologia das infecções enterocócicas. Os autores, avaliando amostras de *Enterococcus* spp. isoladas entre 1998 e 2015, observaram que a partir de 2006 as infecções de etiologia enterocócica, que até então eram predominantemente causadas pela espécie *E. faecalis*, passaram a ter uma relevante participação da espécie *E. faecium*, com a disseminação do ST412. A emergência e circulação do ST412 também já foram relatadas em diferentes locais da região sudeste do país (ALMEIDA; BREDA; SILVA, 2014; ALVES et al., 2017; MARCHI et al., 2018; PRADO et al., 2016; SANTOS et al., 2017).

Uma das preocupações em relação a emergência de diferentes subpopulações de *E. faecium* reside no fato que, em geral, as novas linhagens circulantes estão associadas a resistência a diversos antimicrobianos de última linha para o tratamento de infecções causadas por esse microrganismo. Diferentes trabalhos já relataram a associação do ST412 com a resistência à linezolida (ALMEIDA; BREDA; SILVA, 2014; JASPAN et al., 2010). No trabalho realizado em uma instituição hospitalar de referência no tratamento de pacientes oncológicos da Cidade do México, dentre as 12 amostras analisadas no período de 2009 a 2011, 50% pertenceram ao ST412 (seis amostras), sendo que cinco delas apresentaram resistência à doxiciclina e uma à linezolida (OCHOA et al., 2013). O ST736, relatado como clone epidêmico atualmente circulante nos Estados Unidos, também vem sendo associado a resistência à daptomicina (EL HADDAD et al., 2018; WANG et al., 2018). Em nosso estudo, apenas uma amostra não foi suscetível à linezolida. Entretanto, como esse dado ainda não estava disponível por ocasião da realização das etapas de sequenciamento neste estudo, essa amostra não foi selecionada e não teve o ST determinado.

O primeiro relato da circulação do ST963 foi feito por Santos et al. (2017), a partir de amostras isoladas nos anos de 2013 e 2014, de pacientes hospitalizados em um instituto de referência para o tratamento de pacientes oncológicos na cidade do Rio de Janeiro. Posteriormente, um outro trabalho, dessa vez com amostras oriundas de um hospital universitário da cidade de São Paulo, identificou uma amostra isolada em 2013 pertencente ao ST963 (MARCHI et al., 2018). Neste estudo, no período entre 2009 a 2016, o ST963 esteve presente em 23% (12/52) das amostras sequenciadas, sendo o segundo ST mais frequente no período. Observado primeiramente em uma amostra datada de 2011, os dados obtidos indicaram que esta linhagem já circulava no Brasil antes dos dados publicados por outros

autores e descritos acima. Considerando os estudos divulgados até então na literatura especializada, consideramos supor que a emergência do ST963 esteja associada às amostras brasileiras.

Em relação ao ST896, encontrado em nove amostras deste estudo a partir do ano de 2015, sendo o terceiro mais frequente, sua primeira descrição foi reportada em uma amostra clínica humana, isolada num hospital da cidade de Nova York no ano de 2013 (WANG et al., 2018). Posteriormente, apenas relatos em amostras brasileiras puderam ser observados. No mesmo estudo mencionado anteriormente na cidade de São Paulo, Marchi et al. (2018) encontraram oito amostras do ST896 no período de 2012 a 2014. No presente trabalho, tanto a linhagem do ST896 quanto a do ST963 foram observadas em diversas instituições hospitalares de nosso estado, concomitantemente e por um longo período. Estes achados em conjunto com o já divulgado pela literatura indicam uma ampla disseminação dessas linhagens em nosso meio.

Entre as características comuns das linhagens que emergiram a partir do ano de 2009, está a presença do alelo *purK44* e a presença do gene *hyl*. Em relação à presença do determinante de virulência *hyl*, como mencionado em parágrafos anteriores, este gene está relacionado às amostras circulantes no ambiente hospitalar. Embora diversos estudos tenham relatado a presença de *hyl* em amostras do ST78 (BONORA et al., 2007; PARK et al., 2011; SONG et al., 2006; YANG et al., 2016), este fato não foi observado neste estudo. Todavia, em linhagens recentes, como o ST963, por exemplo, todas as amostras apresentaram o gene. A presença desse determinante de virulência já foi descrita em outros estudos, em amostras pertencentes as linhagens que emergiram mais recentemente (FREITAS et al., 2010; MARCHI et al., 2018; OCHOA et al., 2013; SANTOS et al., 2017).

O alelo 1 do gene *purK* (incluído no protocolo de análise por MLST) foi considerado um marcador de linhagens circulantes no ambiente hospitalar (BONORA et al., 2004; COQUE et al., 2005; LEAVIS et al., 2003; WILLEMS et al., 2001). Entretanto, os STs que passaram a predominar mais recentemente, ST412, ST963 e ST896, apresentam o alelo 44 desse gene de manutenção celular, indicando mudanças na estrutura populacional das amostras circulantes no ambiente hospitalar, resultantes da emergência de diferentes linhagens. A observação de que as linhagens que estão emergindo e circulando recentemente, incluindo o ST736, clone epidêmico nos Estados Unidos (WANG et al., 2018), apresentam o alelo 44, pode sugerir que esse seja um novo marcador de amostras hospitalares.

Um outro aspecto importante em relação aos STs que emergiram recentemente, reside no fato de não pertencerem ao CC17. Como anteriormente descrito, o complexo clonal 17 de

E. faecium foi amplamente caracterizado por albergar linhagens adaptadas ao ambiente hospitalar (LEE et al., 2019a; TOP; WILLEMS; BONTEN, 2008). As linhagens pertencentes ao CC17 apresentam características genotípicas e fenotípicas marcantes, consideradas fundamentais para ascensão dessa espécie, que já foi considerada quase que exclusivamente comensal do TGI humano, ao *status* de um destacado patógeno oportunista circulante em instituições hospitalares de diversos países (BILLSTRÖM, et al., 2010; HSIEH et al., 2010; LEAVIS; BONTEN; WILLEMS, 2006; TITZE-DE-ALMEIDA et al., 2006; TOP et al., 2008; XU et al., 2011). As novas linhagens emergentes observadas no presente trabalho (ST963 e ST896), embora não pertençam ao CC17, demonstram atributos desse grupo, além de outras características importantes, como a presença de genes de virulência responsáveis pela produção de hidrolases (*hyl*) e adesinas (*scm*). Estes achados podem indicar o surgimento de um novo complexo clonal, considerando que essas amostras carregam diferentes atributos adaptativos e, portanto, estariam aptas a se destacarem no ambiente hospitalar.

Além dos STs já mencionados, neste estudo também foram identificadas amostras pertencentes a outros seis STs já reportados na literatura (ST97, ST192, ST203, ST253, ST844 e ST885), dois STs novos que foram primeiramente descritos neste trabalho (ST1626 e ST1628), e um ST ainda não nomeado (ST NOVO). Quando comparados aos dados presentes na literatura, a estrutura populacional em relação aos STs encontrados neste estudo, circulando em nosso estado, está muito mais próximo do observado em países da Europa e da Ásia, bem como na Austrália; e diferente do observado em outros países do continente americano, como os Estados Unidos (ABDELBARY et al., 2019; COOMBS et al., 2014; FREITAS et al., 2016; MICHAEL; ROBERTS, 2017, PINHOLT et al., 2015; ROBERTS et al., 2016; WARDENBURG et al., 2019).

O ST192 tem sido a linhagem associada a infecções invasivas na Alemanha, relacionadas, principalmente, ao fenótipo VanB (BENDER et al., 2016b). Na Suécia, o gene *vanB2* também esteve associado a este ST, através da inserção do Tn1549/Tn5382 no plasmídeo pRUM (rep17). Além disso, as amostras do ST192 foram descritas por possuírem um número maior de determinantes de virulência (>10) (SIVERTSEN et al., 2014). O ST844 também foi descrito associado à circulação no ambiente hospitalar, sendo isolado inicialmente em infecções de corrente sanguínea em pacientes oncológicos na Espanha (GUDIOL et al., 2013). De maneira curiosa, dentre as três amostras pertencentes ao ST844 identificadas neste estudo, duas delas foram isoladas de sangue. Por outro lado, o ST253 reportado como não pertencente a complexos clonais adaptados ao ambiente hospitalar, foi identificado em

amostras comensais humanas isoladas na Alemanha e em amostras oriundas de sistemas de tratamento de esgoto na França (LASCH et al., 2014; LECLERCQ et al., 2013).

O ST203 já foi descrito em países como: Coréia do Sul, China, Dinamarca, Grécia e Malásia, neste último, em amostras do trato gastrointestinal de aves (LEE; LEE; KIM, 2006; DAMANI et al., 2010; GETACHEW et al., 2013; HOFFMANN et al., 2018; ZHENG et al., 2007; ZHU et al., 2009). Na Austrália, esta linhagem foi considerada a responsável por substituir o ST17, anteriormente predominante nesse país. Os dados obtidos através de genômica comparativa, indicaram que as amostras do ST203 possuíam uma região de quase 500 kb exclusivas. Porém, essas regiões não foram associadas à presença de fatores de virulência, o que talvez pudesse ser responsáveis por sua emergência e circulação (JOHNSON et al., 2010). Na Alemanha, onde este ST é considerado epidêmico, sua circulação foi associada à presença dos determinantes *esp* e *hyl* (KLARE et al., 2005; XANTHOPOULOU et al., 2020).

O ST885, que foi encontrado em uma amostra isolada de sangue em nosso estudo, teve sua primeira descrição a partir de uma amostra isolada no solo (BEN SAID et al., 2016). Posteriormente, outro estudo também identificou sua presença em uma amostra clínica humana, que apresentou o gene *optrA*, relacionado à resistência as oxazolidinonas, como a linezolida (CAI et al., 2015).

Foi objetivo deste estudo determinar a filogenia de *E. faecium* através da construção de matrizes de SNPs. Os resultados obtidos permitiram identificar que os ST78, ST412, ST963 e ST896 constituíram braços diferentes da árvore de filogenia, confirmando que múltiplas linhagens circularam em nosso meio durante o período avaliado. Em um trabalho recente descrevendo a estrutura populacional de *E. faecium* na América Latina, Rios et al. (2020) também demonstraram a divisão em dois clados distintos, onde as amostras do ST412 foram separadas do clado que continha a linhagem do ST17. Cabe ressaltar que esses autores, mesmo utilizando abordagens para a reconstrução da filogenia de *E. faecium* distintas das que foram empregadas em nosso estudo, também observaram que esses novos STs emergiram e se expandiram como linhagens separadas.

Na verdade, a distância no relacionamento filogenético dessas linhagens foi observada até mesmo com o emprego de metodologia de PFGE. Suzuki et al. (2014) avaliando amostras pertencentes ao ST412 reportaram que estas foram agrupadas nos perfis de PFGE nomeados pelos autores como “A” e “B”; enquanto amostras do ST78 se agruparam, primordialmente, no perfil “E”. Em outro estudo, utilizando dendrogramas também construídos a partir das análises pela técnica de PFGE, os autores demonstraram que as amostras pertencentes ao

ST412 se agruparam em perfis separados daqueles constituídos por amostras do ST78 (FRANYÓ et al., 2018). Em um estudo brasileiro, as amostras pertencentes ao ST78, ST412, ST963 e ST896 foram, da mesma forma, alocadas em perfis de PFGE completamente distintos (MARCHI et al., 2018).

Através da reconstrução da filogenia, foi possível observar o posicionamento de alguns STs isolados dentro de clados robustos. As duas amostras do ST192, um SLV do ST78, ficaram inseridas dentro do grupo do ST412; o mesmo ocorreu para a amostra pertencente ao ST203, que também é um SLV do ST78. Alguns estudos já alertaram para a baixa resolução do MLST particularmente para inferir relações filogenéticas, considerando suas inerentes limitações quanto à reduzida proporção do genoma que é analisada (sete segmentos gênicos com tamanho em torno de 450 pb) e natureza das variações desses genes. Por serem considerados genes de manutenção celular, evolutivamente apresentam uma tendência de não acumular muitas variações (DIDELOT; MAIDEN, 2010).

Além disso, van Hal et al. (2016) demonstraram, em um estudo avaliando a dinâmica evolutiva de 132 amostras de *E. faecium* através de sequenciamento de genoma completo, que ao menos 28 eventos de recombinação ocorreram em genes que fazem parte do esquema de MLST da espécie (*ddl*, *atpA*, *gyd*, *pstS*), o que interfere nas reconstruções de filogenia e invalida o MLST como única metodologia para avaliação de expansão clonal.

Este parece ser o caso observado em nosso estudo para a amostra CL7611, pertencente ao ST78, que se mostrou agrupada dentro do clado do ST412. Quando foi utilizada uma metodologia para a reconstrução da filogenia, pela análise do conjunto de SNPs presentes em todos os genomas, essa amostra apresentou um distanciamento maior do clado ST412, mesmo compartilhando relações de ancestralidade com este grupo que não foram observadas para o ST78. Embora as duas árvores tenham sido reconstruídas a partir de SNPs, elas diferem substancialmente no conteúdo analisado. Uma árvore reconstruída através de uma matriz de SNPs é feita com base nas variações que cada amostra individualmente possui, em relação a referência escolhida. Já a outra metodologia empregada, consiste na reconstrução da filogenia a partir dos SNPs em relação à amostra de referência que estão presentes simultaneamente em todas as amostras.

Ainda em relação a amostra CL7611, os resultados obtidos, considerando a presença de genes exclusivos para grupos de STs e comparação com de genomas disponíveis no NCBI com os quais esta amostra obteve mais *hits*, indicaram que possivelmente trata-se de um híbrido. Além disso, a avaliação dos eventos de recombinação homóloga identificou ao menos oito eventos no genoma desta amostra, com 64.558 pb afetados. Dados semelhantes também

foram relatados por Galloway-Peña et al. (2012), que identificaram uma amostra, nomeada de “1.231.408” e isolada do sangue de um paciente hospitalizado, que apresentou partes do genoma mais similares com o clado de amostras comensais. Os autores consideraram que se tratava de uma amostra híbrida de *E. faecium*, cujo escore de divergência para o clado de amostras comensais variou entre 3,3 e 3,5, enquanto para o de amostras hospitalares variou entre 1,1 e 1,6. De forma semelhante, os dados relativos à amostra CL761 deste estudo apresentaram um escore de divergência para as amostras do ST78, entre 0,39 e 0,73; e para aquele que agrupa as amostras do ST412, esse escore variou entre 0,45 e 0,50.

As técnicas de biologia molecular, como o PCR e o sequenciamento de larga escala do genoma, são importantes para a compreensão sobre os mecanismos responsáveis pela resistência aos antimicrobianos e sua disseminação em diversas populações bacterianas. Em geral, os resultados obtidos em nosso estudo, através da caracterização genotípica baseada em PCR, apresentaram uma boa correlação com o que foi encontrado nas análises por WGS. Apenas em relação aos aminoglicosídeos, algumas discrepâncias foram observadas. Poucos relatos são descritos na literatura em relação a inconsistência entre as duas técnicas (MASON et al., 2018). A não predição de um gene através de WGS que tenha sido anteriormente amplificado por PCR pode ser explicada pela sua ausência nos bancos de dados utilizados, ou quando ocorre algum tipo de interrupção na sequência do gene, devido ao processo de montagem do genoma ou inserção de elementos (MORAN et al., 2017).

Por outro lado, a utilização de WGS permite uma visão completa, ou parcialmente completa, do todo, ao invés de pequenas peças que compõem um quebra-cabeça. Dentro desse contexto, foi possível identificar o resistoma, ou seja, os genes presentes no genoma ou na população que contribuem para a resistência aos antimicrobianos. Esse conjunto de genes de resistência não é apenas necessário para a emergência de subpopulações mais bem adaptadas, como também reflete a história evolutiva bacteriana.

A resistência aos antimicrobianos, embora seja frequentemente avaliada sob uma ótica antropogênica, deve ser considerada de fato um problema de magnitude ambiental e evolutiva. Muitos dos genes que hoje compõem o resistoma de bactérias patogênicas têm sua origem no ambiente, com um caminho evolutivo anterior à utilização dos antimicrobianos na medicina humana (WRIGHT; POINAR, 2012). Em um estudo de Marshall et al. (1998) os autores identificaram em bactérias produtoras de glicopeptídeos, da espécie *Amycolatopsis orientalis*, genes que apresentavam similaridade superior a 60% com aqueles que compõem o operon de resistência à vancomicina em enterococos. Em um estudo de 2011, os autores utilizando ferramentas de metagenômica identificaram variantes ancestrais do gene *vanA* e de

genes responsáveis pela resistência aos β -lactâmicos e tetraciclina em sedimentos de solo permanentemente congelado (*permafrost*) datados de 30.000 anos (D' COSTA et al., 2011).

Dentro desse resistoma, estão também genes que evoluíram através de protogenes ou genes precursores, que anteriormente estavam envolvidos em funções metabólicas e de manutenção celular e não possuíam nenhum efeito relacionado aos antibióticos. Muitas das enzimas de modificação de aminoglicosídeos, amplamente disseminadas em importantes patógenos humanos, têm estruturas e funções semelhantes a enzimas metabólicas, como por exemplo proteínas quinases (WRIGHT, 2010).

As pressões seletivas exercidas pelo homem têm contribuído para moldar o arranjo atual desses resistomas, tanto em diversidade dos genes quanto em abundância. A importância desse impacto decorre do fato de que esse conjunto de genes fornece a diversidade genética necessária para a ocorrência de mecanismos adaptativos que levam a emergência de novas linhagens, principalmente no ambiente hospitalar. Assim, o entendimento da composição do resistoma é fundamental para o combate de amostras MDR bem adaptadas a esse ambiente.

Neste estudo, ao todo 20 genes relacionados com a resistência às mais diversas classes de antimicrobianos foram encontrados na população analisada. Alguns apresentaram uma considerável abundância nessa população, enquanto outros se mostraram de forma mais esporádica e, em algumas situações, relacionados a apenas uma amostra. Dentre os genes que foram identificados em mais de 80% das amostras estudadas, está *sat4*. Este gene codifica para resistência à estreptotricina, um antibiótico produzido por *Streptomyces* spp. e que apresenta atividade similar aos aminoglicosídeos, e efeito, principalmente, em bactérias Gram-negativas. Durante muito tempo, este antimicrobiano foi utilizado como promotor de crescimento animal (SMALLA et al., 1993; WAKSMAN; WOODRUFF, 1942; WITTE et al., 2000).

O gene *sat4* foi descrito inicialmente em amostras de *Campylobacter coli*. Em amostras de *E. faecium*, está inserido em cluster gênico juntamente com *ant(6')-Ia* (também reconhecido como *aadE*), e *aph(3')* (ARIAS; MURRAY, 2012; JACOB et al., 1994; WERNER; HILDEBRANDT; WITTEM, 2003). No gênero *Enterococcus*, *sat4* foi identificado inserido no transposon Tn5405, bem como no plasmídeo pRE25 de *E. faecalis* (REHMAN et al., 2018; TEUBER; SCHWARZ; PERRETEN, 2003). Possui uma ampla distribuição entre amostras de origens diversas e com o advento da utilização do WGS vem sendo cada vez mais encontrado em amostras de origem humana (ASLAM; DIARRA; MASSON, 2012; CHEN et al., 2019; EISENBERGER et al., 2020).

Neste estudo, o gene *aph(3')-IIIa*, responsável pela característica de resistência a níveis elevados apenas de canamicina, também foi bastante frequente entre as amostras estudadas. Apesar de ter sido pesquisado através da técnica de PCR somente nas amostras que exibiram o fenótipo HLGR, esse gene se mostrou amplamente disseminado na amostragem. Dessa forma, cabe ressaltar que os testes fenotípicos preconizados para a identificação de HLAR (CLINICAL LABORATORY STANDARDS INSTITUTE 2019), utilizando como critério o uso de gentamicina e estreptomicina, não são suficientes para inferir a resistência a canamicina, quando resultante da presença desse determinante. Além disso, não se mostram como um excelente parâmetro para a pesquisa genotípica dos determinantes de HLAR.

Os resultados obtidos em relação a prevalência de genes de resistência aos macrolídeos, estreptograminas e lincosamidas, estão de acordo com outros estudos envolvendo o sequenciamento de genoma completo de um número grande amostras de *E. faecium*, em que se observa uma predominância de *efmA*, *erm(B)*, *msr(C)* (EICHEL et al., 2020; EISENBERGER et al., 2020; KERSCHNER et al., 2019; LEE et al., 2019b; 2020; REHMAN et al., 2018; RIOS et al., 2020).

Os genes associados a resistência às tetraciclinas, *tet(L)* e *tet(M)*, estiveram mais relacionadas com amostras do ST896, porém estes genes constituem os mais disseminados em *E. faecium*, devido à sua associação ao Tn916 que é um elemento genético amplamente distribuído (AGERSØ; PEDERSEN; AARESTRUP, 2006; FRAZZON et al., 2010; NISHIMOTO et al., 2005; REHMAN et al., 2018; RIZZOTTI et al., 2009; RATHNAYAKE; HARGREAVES; HUYGENS, 2012).

Os genes *tet(S)* e *tet(U)* foram observados apenas na amostra híbrida, CL7611. O gene *tet(U)* foi descrito inicialmente por Ridenhour et al. (1996), em amostras de *E. faecium* que apresentavam o plasmídeo pKQ10. Também associado a uma disseminação por elemento móvel, o gene *tet(S)* foi primariamente encontrado em amostras de *E. casseliflavus* e *E. gallinarum* (BROUWER; MULLANY; ROBERTS, 2010; FARD; HEUZENROEDER; BARTON, 2011; Novais et al., 2012; ROBERTS et al., 2006). Entretanto, assim como no presente trabalho, com a utilização de estratégia de sequenciamento do genoma completo para identificação e caracterização do resistoma de um grande número de amostras bacterianas, esses dois genes vêm sendo cada vez mais reportados em *E. faecium* (LEE et al., 2020; REHMAN et al., 2018; RIOS et al., 2020).

É reconhecido que os membros do gênero *Enterococcus* são intrinsecamente resistentes aos efeitos de trimetoprim e de sulfametoxazol, por serem capazes de obter ácido fólico do meio externo (HOLLENBECK; RICE, 2012). Por isso, testes *in vitro* já

demonstraram que os enterococos são suscetíveis ao trimetoprim-sulfametoxazol em meios sem a presença de folato (ZERVOS; SCHABERG, 1985). Sendo assim, a utilização dessa combinação de antimicrobianos como opção terapêutica é bastante controversa (GRAYSON et al., 1990; RAMBALDI et al., 1997; WISELL; KAHLMETER; GISKE, 2008; WOJCIECHOWSKI; CHANDRAN, 2013).

Neste estudo, dentre os achados relativos ao resistoma das amostras de *E. faecium*, ressaltamos o encontro do gene *dfrF*, responsável por produzir uma di-hidrofolato redutase responsável pela característica de resistência adquirida a níveis elevados de trimetoprim-sulfametoxazol (COQUE et al., 1999). Este determinante de resistência foi identificado na quase totalidade de amostras relacionadas aos STs que emergiram mais recentemente (ST412, ST963 e ST896). Entretanto, em um estudo avaliando o genoma de amostras de *E. faecium* responsáveis por surtos em hospitais da Alemanha, os autores identificaram esse gene disseminado em diversas outras linhagens, incluindo no ST78 (EISENBERGER et al., 2020). Sendo assim, só é possível afirmar que localmente essa característica esteve associada com os STs que emergiram mais recentemente.

Diferente do observado no presente trabalho, onde foi identificado em apenas quatro amostras, o gene *dfrG* tem sido reportado na literatura por ser amplamente disseminado em amostras de *E. faecium* (CHEN et al., 2019; RIOS et al., 2020). Ainda em 2009, Cattoir et al., avaliando amostras de *E. faecium* resistentes à trimetoprim-sulfametoxazol isoladas de hospitais franceses, identificaram o gene em uma frequência de 59,2%. No mesmo estudo alemão mencionado anteriormente, 77% dos isolados apresentaram *dfrG* em seu genoma (EISENBERGER et al., 2020).

Os genes *eatAv* e *cat*, encontrados esporadicamente em algumas amostras, estão associados ao fenótipo LS_AP e a resistência ao cloranfenicol, respectivamente. O fenótipo LS_AP é um mecanismo de resistência cruzada às lincosamidas, estreptograminas A e pleuromutilinas, caracterizado primeiramente em *E. faecium* no ano de 2013 (ISNARD et al., 2013). Desde sua descrição, apenas um estudo relatou a ocorrência de *eatAv* em cinco amostras de *E. faecium* que apresentaram uma característica marcante, a ausência do gene *pstS*, sendo não-tipificáveis pela técnica de MLST (LEMONIDIS et al., 2019). Já o gene *cat*, associado a produção da enzima cloranfenicol acetiltransferase, possui ampla descrição, embora sua presença em amostras de *E. faecium* das mais diversas origens seja apenas esporádica (CHEN et al., 2019; KANG et al., 2019; LEE et al., 2020; RIOS et al., 2020; YAHIA et al., 2018).

Por outro lado, diante do crescente acesso às tecnologias de sequenciamento do genoma completo e dos potenciais benefícios que esta metodologia trouxe para o rastreamento de resistência aos antimicrobianos, diversas discussões têm surgido sobre a sua ampla aplicação em substituição às técnicas fenotípicas clássicas (KÖSER; MATTHEW; PEACOCK, 2014; HENDRIKSEN et al., 2019; SU; SATOLA; READ, 2019). De fato, deve ser levado em consideração um conceito básico da biologia molecular, de que a predição de um gene não significa sua posterior expressão. Nesse sentido, destaca-se o reporte do comitê EUROPEAN COMMITTEE ON ANTIMICROBIAL SUSCEPTIBILITY TESTING, publicado em 2017, que elenca uma série de recomendações e conclusões quanto aos padrões que devem ser observados para a utilização da metodologia de WGS como preditora de resistência aos antimicrobianos. Uma importante reflexão feita pelos autores ressalta que a medida do diâmetro de uma zona de inibição, ou mesmo os valores de CIM, refletem vários parâmetros complexos, como, por exemplo, a biodisponibilidade da droga, e não apenas a presença ou ausência de um gene (ELLINGTON et al., 2017).

Entretanto, neste estudo foram identificados resultados incongruentes entre a metodologia de WGS e análises fenotípicas. Foram caracterizadas 36 amostras fenotipicamente sensíveis a níveis elevados de estreptomicina que carregavam o gene *ant(6')-Ia*. As análises complementares confirmaram os resultados obtidos e indicaram que deleções no gene acarretaram mudanças na estrutura conformacional da proteína. Além disso, a investigação por PCR nas amostras que não tiveram o genoma sequenciado identificou mais 124 amostras que apresentaram incongruência fenotípica e genotípica para estreptomicina. A observação de um número significativo de amostras apresentando esse fenômeno sugere uma disseminação clonal, através de transferência horizontal no ambiente hospitalar. Neste caso, os dados obtidos foram fundamentados exclusivamente por análises genotípicas.

O mesmo fenômeno pode ser observado em um estudo com amostras obtidas de carcaças de aves (JACKSON et al., 2004). Os autores demonstraram que duas amostras de *E. faecium* e uma de *E. faecalis* apresentaram essas incongruências entre fenótipo-genótipo. Uma explicação para a perda tão acentuada de partes da sequência de nucleotídeos dentro do gene, pode residir na observação de rearranjos nesse *cluster* genético (*aadE-sat4-aph3*); além da presença de ISs dentro do *Tn5405-like*, responsável por carrear esse *cluster* gênico, como já apontado por Werner et al. (2003). Em outro estudo, Leelaporn et al. (2008) identificaram 30 amostras que apresentaram o gene *aadE*, mas não demonstraram o fenótipo HLSR. De fato, esse fenômeno já foi identificado também em outros patógenos bacterianos, como *Escherichia coli* e *Salmonella* spp., em que os achados por WGS demonstraram divergência

daqueles indicados por testes fenotípicos para detecção da resistência aos aminoglicosídeos (MCDERMOTT et al., 2016; THOMAS et al., 2017; TYSON et al., 2015).

Além disso, também foi observado, de forma similar, a presença do gene *cfr*(B) em uma amostra que caracterizada fenotipicamente como sensível a linezolida. O gene *cfr*(B), descrito inicialmente em 2015, é uma variante do gene *cfr* que está associado a resistência à linezolida mediada por plasmídeo em *Enterococcus* spp. (DESHPANDE et al., 2015). Assim como todos os genes *cfr*, o mecanismo exercido consiste na metilação de alguns resíduos funcionais do RNA ribossômico 23S, resultando em efeitos protetivos contra a ação de diversos antimicrobianos (LONG et al., 2006). Embora as mutações no rRNA 23S sejam o principal mecanismo de resistência as oxazolidinonas, a circulação desse gene associada à uma disseminação plasmidial configura uma situação preocupante (BENDER et al., 2016a; DESHPANDE et al., 2018).

Contudo, uma das discussões que devem ser feitas consiste no entendimento da manutenção e perpetuação de um elemento que não apresenta funcionalidades. Em geral, as bactérias mantêm um genoma conciso e funcional, com poucos elementos não codificantes, e que são a base para uma reprodução rápida. Além disso, o custo da aquisição e manutenção de elementos genéticos extras, como por exemplo genes de resistência, há muito vem sendo discutido na literatura (SAN MILLAN; MACLEAN, 2017).

Em termos evolutivos, a tendência é que genes não-funcionais sejam perdidos, no entanto, esse processo se dá dentro de uma escala de tempo maior, o que não exclui a hipótese de que futuramente esses elementos possam até ser perdidos. Entretanto, em contraponto, já foi observado que a ocorrência de mutações que levam a uma restauração do *fitness* bacteriano tende a ocorrer mais frequentemente do que a perda (ANDERSSON; HUGHES, 2011), o que poderia favorecer a manutenção desse elemento. Além disso, como descrito anteriormente, este gene se apresenta num *cluster* genético, ou seja, em um pacote com outros genes. Esses elementos ligados em *clusters* tendem a interagir uns com os outros das mais variadas formas, o que aumenta sua habilidade coletiva de manutenção (SALYYERS; AMÁBILE-CUEVAS, 1997). Outro fator associado a essa manutenção poderia ser explicado pela cosseleção para algum outro determinante de resistência ou mesmo de metabolismo, que confira algum tipo de vantagem seletiva para o microrganismo (ANDERSSON; HUGHES, 2011).

De fato, existe uma forte correlação entre a presença do *cluster aadE-sat4-aph3* e o gene *erm*(B), associado a resistência aos macrolídeos e amplamente disseminado na amostragem, pois ambos estão presentes no Tn5405-like (REHMAN et al., 2018; WERNER

et al., 2003). Além disso, este transposon já foi detectado em diferentes plasmídeos, onde também estava presente o *Tn1546-like*, responsável pela disseminação do gene *vanA*, e que poderia funcionar como fator de cosseleção (LIM et al., 2006; PINHOLT et al., 2017). A ampla circulação deste elemento defeituoso, como observado no trabalho, pode também contribuir para sua disseminação contínua.

Dentre os cosseleccionadores de resistência mais amplamente descritos e estudados, estão os metais pesados e biocidas (BAKER-AUSTIN et al., 2006; PAL et al., 2015; SEILER; BARENDONK, 2012). A contaminação ambiental por estes agentes é amplamente descrita na literatura e os processos de cosseleção ocorrem pela presença simultânea de genes que levam a fenótipos de resistência aos metais pesados em elementos móveis que carregam genes de resistência aos antimicrobianos (LI et al., 2019; RAI et al., 2019; SANKHLA et al., 2016; YU et al., 2017). No presente estudo, quatro genes associados à resistência a metais pesados foram pesquisados nos genomas bacterianos, sendo possível observar a presença de *cadA* em todas as amostras, enquanto que *cueO*, *merA* e *tcrB* estiveram mais relacionados ao ST78.

Dentro do gênero *Enterococcus*, diversos trabalhos já apontaram a correlação da resistência aos metais pesados com a ocorrência de amostras MDR, bem como a disseminação de determinantes de resistência. Kimiran-Erdem et al. (2007) observaram que a maioria das amostras de *Enterococcus* spp. fenotipicamente resistentes à amicacina, canamicina, ácido nalidíxico e estreptomicina, também apresentou resistência ao cádmio e ao cobre. Outro estudo também observou a correlação da presença de resistência fenotípica a diversos metais pesados com a presença dos genes *tet* (resistência à tetraciclina) (LIU et al., 2019).

Em relação a espécie *E. faecium*, um trabalho brasileiro publicado em 2012 observou que amostras isoladas das fezes de pombos urbanos, apresentaram tolerância ao mercúrio, cádmio e cobre (SILVA et al., 2012). Em um estudo recente (VIGNAROLI et al., 2018), apesar dos autores não terem detectado a presença dos genes *cadA*, *tcrB* e *merA*, as amostras de *E. faecium* isoladas de sedimento marinho, foram resistentes ao cádmio e ao cobre em testes fenotípicos e, também, em maioria se mostraram multirresistentes aos antimicrobianos.

O gene *tcrB* que codifica para a resistência ao cobre, dentre os relacionados à resistência aos metais, é o mais amplamente disseminado em *E. faecium*. Alguns estudos já demonstraram que a ocorrência na espécie está associada à presença concomitante de genes de resistência aos antimicrobianos, como como *vanA* e *erm(B)*, presentes em elementos como o *Tn1546*, que possui ampla disseminação através de plasmídeos (HASMAN; AARESTRUP, 2002; 2005; PASQUAROLI et al., 2014). Em um estudo de Fard et al. (2011), 46 amostras de

E. faecium isoladas do intestino de porcos saudáveis foram positivas para o gene *tcrB*, bem como para os genes *erm(B)* e diversos genes da família *tet*.

Embora no presente estudo tenha ficado restrita à linhagem do ST78, a presença da resistência ao cobre parece estar amplamente dispersa nas subpopulações de *E. faecium*. Silveira et al. (2014), demonstraram que 23% e 25% das amostras de *E. faecium* apresentaram o gene *tcrB* e *cueO*, respectivamente, ambos associados à resistência ao cobre. Os autores também relataram que ambos elementos foram detectados nos mais variados STs de diferentes origens. Um estudo com gados em confinamento alimentados com uma dieta rica em cobre demonstrou que 6,9% das amostras de *E. faecium* carregaram o gene *tcrB*, além de *erm(B)* e *tet(M)*. A reconstrução da população, utilizando os resultados obtidos pela técnica de MLVA e o algoritmo eBURST, indicou que essas amostras representavam diferentes grupos populacionais (AMACHAWADI et al., 2013). A presença dos genes *cueO* e *cadA* também foi relatada em amostras pertencentes ao ST796, clone epidêmico na Austrália (BUULTJENS et al., 2017). Por sua vez, o gene *merA*, que confere resistência ao mercúrio, já foi descrito associado a um plasmídeo de resistência à estreptomicina, pPPM1000 (DAVIS et al., 2005).

A utilização do sequenciamento do genoma completo também permitiu observar uma ampla gama de fatores de virulência nas amostras de *E. faecium*. Os dados obtidos corroboraram diversos outros estudos sobre o arsenal de determinantes de virulência presentes na espécie (EISENBERGER et al., 2020; FREITAS, A.R et al., 2018a; LEE et al., 2020; RIOS et al., 2020; SOHEILI et al., 2014). Estes achados demonstraram o potencial de *E. faecium* em exibir atributos de virulência, característica comumente mais associada a espécie *E. faecalis* (GUZMAN PRIETO et al., 2016; LANDETE et al., 2018; SHOKOOHIZADEH et al., 2018). Embora muitos dos determinantes encontrados ainda não tenham um papel definido na patogênese das infecções por *E. faecium* (GAO; HOWDEN; STINEAR, 2018), a presença desses elementos alerta para o potencial patogênico desse microrganismo, reconhecidamente multirresistente aos antimicrobianos.

A circulação de alguns determinantes de virulência, como *hyl* e *scm*, foi associada com algumas linhagens específicas. Os resultados obtidos através da técnica de PCR para o gene *hyl* serviram como indicador de que mudanças populacionais estariam ocorrendo, sendo posteriormente corroborado com a reconstrução da filogenia e a observação que esse determinante não foi encontrado nas amostras do ST78. Contudo, em termos gerais, não é possível inferir que a presença desses determinantes seja um parâmetro indicativo dessas alterações. Deve-se a isso o fato de que tanto *hyl* quanto *scm* já foram observados nas mais

diversas linhagens de *E. faecium*, estando, em geral, associados às amostras de origem humana (EISENBERGER et al., 2020; FREITAS, A.R et al., 2018a; GHATTARGI et al., 2018; LEE et al., 2019b; SOHEILI et al., 2014). Adicionalmente, no recente trabalho de Rios et al. (2020), os autores observaram que o clado onde todas as amostras do ST412 foram agrupadas não foi associado à presença do gene *hyl*.

O mobiloma é definido como o conjunto de todos os elementos genéticos móveis dentro do genoma e contribuem para a sua plasticidade (SANTAGATI; CAMPANILE; STEFANI, 2012). No presente estudo, o mobiloma das amostras de *E. faecium* foi avaliado através da detecção da presença de plasmídeos, profagos intactos e ISs. A média de plasmídeos foi de 3,6 por amostra. No estudo de Arredondo-Alonso et al. (2020), a média de plasmídeos em amostras de *E. faecium* de origem hospitalar foi de 5,7, portanto comparativamente um pouco mais elevada.

Todavia, cabe ressaltar que neste estudo foi utilizada a plataforma *PlasmidFinder*, que permite apenas determinar quantos tipos de famílias Rep estão presentes no genoma, o que não reflete o exato número de plasmídeos presentes em cada amostra. Uma vez que diversos trabalhos já demonstraram a existência de plasmídeos multireplicons (ARREDONDO-ALONSO et al., 2020; FREITAS et al., 2013; MIKALSEN et al., 2015), ou seja, que codificam diferentes proteínas replicadoras (Rep), a determinação do número exato de plasmídeos por amostra demandaria abordagens computacionais específicas, envolvendo a identificação e montagem de *reads* plasmidiais, em separado da montagem do cromossomo.

As proteínas Rep mais abundantes no estudo foram *repUS15* e *rep17*, ambas pertencentes a grande família de plasmídeos conjugativos denominada RepA_N. Esta família é amplamente disseminada em bactérias Gram-positivas, com presença frequente em amostras de *E. faecium* (ARREDONDO-ALONSO et al., 2020; MIKALSEN et al., 2015; WEAVER et al., 2009). Além disso, esta família, e em especial *rep17*, tem sido descrita associada a amostras pertencentes a linhagens altamente adaptadas ao ambiente hospitalar e relacionadas a disseminação do gene *vanA* (ARREDONDO-ALONSO et al., 2020; HEGSTAD et al., 2010; MELEGH et al., 2018; ORAVCOVÁ et al., 2018).

Algumas famílias Rep tiveram maior prevalência em grupos de amostras pertencentes a uma determinada linhagem de ST, como por exemplo *rep11a* no ST963 e ST896, *rep14b* no ST412 e ST963 e *repUS43* no ST896. Um estudo de Mikalsen et al. (2015), demonstrou uma maior prevalência de *rep11* em amostras do ST78, quando comparada ao ST17. No presente estudo, *rep11a* não foi encontrada nas amostras pertencentes ao ST78 analisadas.

Quanto às demais, *rep14b* e *repUS43* pertencem ao grupo Rep_trans, que também já foi apontado como altamente disseminado em amostras de enterococos (LANZA et al., 2015). Já *rep11a*, que pertence à família Rep_3, foi identificada em *Enterococcus* spp. das mais diversas origens, incluindo amostras VRE de complexos clonais associados ao ambiente hospitalar; além de amostras de outros gêneros de cocos Gram-positivos como *Staphylococcus*, *Streptococcus*, *Lactococcus* (CLEWELL et al., 2014; QU et al., 2012). A família Inc18, também observada no presente estudo, pela presença de *rep1*, *rep2* e *repUS7A*, é reconhecida por carrear múltiplos determinantes de resistência aos antimicrobianos e metais pesados, e mais recentemente tem estado associada à disseminação do Tn1546 (CLEWELL et al., 2014; VALDEZATE et al., 2012; WARDAL et al., 2017).

Em geral, a presença de elementos acessórios, tais como os plasmídeos, é interpretada apenas a partir da perspectiva relativa aos benefícios que promovem para a célula bacteriana e dos valores adaptativos e/ou custosos que esses elementos podem trazer ao seu hospedeiro. Todavia, o entendimento de que esses elementos possuem uma dinâmica populacional e evolutiva própria é fundamental para a compreensão dos processos evolutivos bacterianos, dado que são nesses elementos onde grande parte da variação necessária para a evolução bacteriana tem início e é transmitida.

Os plasmídeos são entidades moleculares independentes, concentrados, de forma não literal, apenas em sua própria sobrevivência (SLATER et al., 2008). Geralmente carregam atributos que permitem o acesso e sobrevivência da bactéria hospedeira a novos nichos ambientais, sendo a resistência aos antimicrobianos o mais bem estudado desses processos. Em um estudo recente, Arredondo-Alonso et al. (2020) postularam que a emergência de *E. faecium* no ambiente hospitalar estaria relacionada ao elevado número de plasmídeos nessas amostras e que a configuração diversa desses elementos poderia contribuir com a especificidade necessária para a adaptação da espécie no ambiente hospitalar.

Contudo, uma vez que a condição estressante que exerce pressão seletiva é removida, a presença desses elementos deixa de ser benéfica. Diversos estudos vêm demonstrando que na ausência de pressão seletiva de antimicrobianos, a presença de plasmídeos eleva os custos para a célula hospedeira (SAN MILLAN et al., 2017; VOGWILL; MACLEAN, 2015). Um trabalho de revisão abordando os efeitos da presença de plasmídeos para a célula hospedeira observou uma redução no *fitness* bacteriano variando entre 1% e 28%, sendo os enterococos os que mais sofreram com esses custos (CARROLL; WONG, 2018).

Considerando que em determinado momento os elementos genéticos carregados por plasmídeos sejam benéficos para a célula hospedeira e dado os custos e a instabilidade que

possuem, é razoável admitir que sua integração ao cromossomo bacteriano seria uma possibilidade natural. Em *E. faecium*, diversos estudos demonstraram a integração dos transposons relacionados a resistência à vancomicina no cromossomo bacteriano (HANDWERGER; SKOBLE, 1995; PAPAGIANNITSIS et al., 2017; WARDAL et al., 2017). Em estudo publicado por Bergstrom et al. (2000), os autores apresentaram a hipótese de que caso a mobilidade desses determinantes adaptativos, através de elementos móveis, venha a acarretar custo ao hospedeiro, que não é observado quando os mesmos estão integrados e transportados através do cromossomo, então, eles podem eventualmente perder sua mobilidade. Do mesmo modo, o oposto ocorre; na medida em que a estabilidade do elemento e as oportunidades de serem transmitidos dentro da população decaem, por exemplo, a seleção favorecerá a mobilidade desse elemento.

Entretanto, o que se observa, mesmo diante das situações anteriormente apresentadas, é a persistência desses plasmídeos em populações bacterianas. Essa longa permanência tem sido tratada como “o paradoxo dos plasmídeos” (HARRISON; BROCKHURST, 2012). A resolução desse paradoxo reside em um processo de coevolução, favorecendo mecanismos que reduzem os custos de *fitness* para o hospedeiro (MACLEAN; SAN MILLAN, 2015).

Dentre os EGMs, também observados no estudo, uma diversidade de profagos foi encontrada, com algumas associações em relação à linhagem das amostras. Sanderson et al. (2020) também encontraram diferentes tipos de profagos em amostras de *E. faecium*, porém, nenhum deles semelhantes aos encontrados neste estudo. O profago phage_Bacill_phBC6A52, encontrado associado as amostras do ST412, já foi reportado em amostras clínicas de *E. cecorum* isoladas de aves (SHARMA et al., 2020). As amostras do ST17 já foram relacionadas com a presença dos fagos PhiEnfa001, Enfa002 e Enfa0003 (LAM et al., 2012). Todos os profagos encontrados no presente estudo pertenceram a família *Siphoviridae* que, junto com as famílias *Podoviridae* e *Myoviridae*, são as mais frequentemente encontradas em enterococos (DUERKOP; PALMER; HORSBURGH, 2014).

Um estudo de Fard et al. (2011) demonstrou experimentalmente a transferência de genes de resistência à tetraciclina mediada por fagos em enterococos. Contudo, ainda não está claro se, na natureza, os fagos carregam genes que contribuam para a adaptação de *E. faecium* (DUERKOP; PALMER; HORSBURGH, 2014; MILKALSEN et al., 2015). Um artigo de Van Schaik et al. (2010) demonstrou que entre 2,3% e 5,1% dos genes presentes nos genomas analisados eram oriundos de fagos. Em geral, as amostras de *E. faecium* adaptadas ao ambiente hospitalar parecem estar mais associadas a presença desses elementos (KIM; MARCO, 2014).

No artigo de Lebreton et al. (2013), que demonstrou a divisão de *E. faecium* em clados distintos, as amostras pertencentes ao clado A1, associadas às infecções em humanos, tinham uma média de 1,6 profagos por genoma. O clado B, que agrupou amostras comensais humanas, apresentou em média 0,9 profagos por genoma, enquanto que no clado A2, que reúne amostras bacterianas oriundas de animais, essa média foi de 0,7. No presente estudo, a média geral de profagos no genoma foi de 0,7, com uma predominância maior em amostras do ST78 (média de 1 profago por genoma).

Em termos evolutivos, a presença de fagos pode ser benéfica em um ambiente altamente competitivo. Um estudo de Brown et al. (2006) demonstrou experimentalmente que os fagos servem como uma excelente estratégia para invasão de nichos. Embora possa trazer riscos para a célula hospedeira, pois eventualmente pode resultar em lise celular, eles também podem eliminar os concorrentes. Os autores compararam tais achados a verdadeiras epidemias, demonstrando que os fagos, além de eliminarem concorrentes, também os utilizam para o seu processo de replicação. Assim, pode-se pensar na hipótese de que profagos diferentes relacionados a linhagens distintas, como identificado neste estudo, atuariam como uma estratégia evolutiva em situações competitivas para um mesmo nicho ecológico.

As seqüências de inserção têm importante impacto na estrutura e função do genoma. Um trabalho de Leavis et al. (2007) propôs que a *IS16*, descrita por ser um marcador de amostras de *E. faecium* no ambiente hospitalar, conferiu certa flexibilidade genômica que permitiu a adaptação dessas subpopulações de *E. faecium*. No presente trabalho, essa IS foi encontrada em 70 cópias. As ISs também apresentam impactos individuais em genes, alterando sua expressão. Monteiro da Silva et al. (2020) demonstraram que a *IS19*, também conhecida como *ISEfm1*, foi responsável por manter a expressão constitutiva de uma amostra com fenótipo VanA, por estar inserida no gene *vanS*.

No presente estudo, uma variedade de ISs das mais diversas famílias foram identificadas. Dentre as que estiveram em maior número de cópias, *ISEfa11*, pertencente à família *ISL3*, foi considerada predominante dentre amostras do ST78. Mikalsen et al. (2015) também observaram que esta IS foi mais frequentemente encontrada nas amostras pertencentes aos ST17 e ST78.

A *ISEfa5*, que foi a segunda mais identificada no estudo, já foi associada ao *Tn1546-like* em amostras isoladas nos diversos continentes (CAMARGO et al., 2005; KHAN et al., 2010; LAM et al., 2013; WARDAL et al., 2017). Em um estudo recente, que avaliou o dinamismo a longo prazo do genoma de amostras de colonização em pacientes hospitalizados, os autores identificaram de 17 a 27 cópias de *ISEfa5* nos genomas de *E. faecium* e sugeriram

que o alto número de eventos de excisão e inserção desse elemento contribuiu para a flexibilidade do genoma das amostras (BAYJANOV et al., 2019).

Em 2018, Ghattargi et al., comparando o genoma de uma amostra de *E. faecium*, utilizada como probiótico (170M39), com o de amostras patogênicas e não-patogênicas da espécie, identificaram que *ISEfa8* e *ISEfa11* estavam presentes no genoma apenas de amostras patogênicas; enquanto que, *ISEfa12* foi associada a amostras comensais. Neste estudo, *ISEfa8* e *ISEfa11* foram identificadas em frequência elevada. Entretanto, dentre os 74 genomas aqui avaliados, *ISEfa12* não foi identificada. Assim, comparando os dados obtidos no presente estudo, pode-se sugerir que as amostras representativas de colonização poderiam estar albergando padrão de ISs supostamente mais similar com o já encontrado em amostras patogênicas.

A presença de ISs associadas especificamente a algumas linhagens já foi reportada. No trabalho de Mikalsen et al. (2015), foi descrito que a *IS1380* foi prevalente em amostras do ST78. Esses dados são diferentes daqueles apresentados neste estudo, em que as amostras do ST78 aqui avaliadas estiveram associadas a *IS1216V*. Alguns trabalhos já demonstraram que *IS1216V* pode comumente ser identificada inserida no *Tn1546*, acarretando disrupções no gene *vanS* (DARINI et al., 1999), que pode inclusive levar à expressão do fenótipo VanB em amostras que carregam o genótipo *vanA* (PARK et al., 2008).

A *ISLgar5* de *Lactococcus garvieae*, que neste estudo esteve associada às linhagens que emergiram mais recentemente, apresentou similaridades com *ISEfm1* segundo o estudo de Eraclio et al. (2015). Também foi reportada por estar inserida no *Tn5801*, responsável pela disseminação de *tet(M)*, em uma amostra de *Enterococcus villorum* (LEÓN-SAMPEDRO et al., 2016), além de já ter sido identificada no genoma de amostras de *E. faecium* tanto utilizadas como probióticas, quanto patogênicas (GHATTARGI et al., 2018).

A presença de ISs comuns a outras espécies bacterianas demonstra a elevada capacidade de *E. faecium* em realizar trocas de material genético, bem como confirmam a sua plasticidade genômica. Todavia, esses dados alertam para a capacidade desses microrganismos de realizar habitualmente eventos de transferência de determinantes de resistência e/ou virulência, contribuindo com frequência para a ampliação do vasto arsenal que a espécie já possui. Um estudo de Wagner et al. (2008), demonstrou que, embora raros, os eventos de trocas de IS entre espécies filogeneticamente distintas podem ocorrer. No estudo, os autores sugeriram uma transferência recente da *IS256* entre *E. faecalis* e *Staphylococcus epidermidis*. Outros eventos de transferências de ISs entre espécies distintas também já foram reportados na literatura (TOUCHON; ROCHA, 2007; WAGNER; LEWIS, BICHSEL, 2007).

Além da *ISLgar5*, mencionada anteriormente, outras sequências de inserção oriundas de diferentes espécies bacterianas foram observadas. A *IS1182* foi associada ao transposon de *S. aureus* Tn5405, responsável por carrear o *cluster* genético de resistência aos aminoglicosídeos *aadE-sat4-aph3* (DERBISE; DE; EL, 1997). Também descrita em *S. aureus*, a *IS256*, bastante disseminada no genoma das amostras estudadas, já foi identificada flanqueando o gene de resistência aos aminoglicosídeos *aphD-aac* no transposon Tn4001-like (BYRNE; ROUCH; SKURRAY, 1989). Rice et al. (1994) já indicaram que 85% das amostras clínicas de enterococos apresentaram essa IS. Em *E. faecium*, tem sido responsável pela diversidade do Tn4001 e do Tn5281-like, ambos associados à disseminação de determinantes do fenótipo HLGR (KLIBI et al., 2006; WATANABE et al., 2009).

Nos estudos já mencionados anteriormente de Klibi et al. (2006) e Watanabe et al. (2009), foi observado que a ausência da *IS256* em uma e/ou ambas extremidades do gene *aac(6')-aph(2'')* geraram transposons truncados. Como a excisão de ISs pode levar a perda das sequências nucleotídicas subjacentes, esses achados levantam a hipótese que pode explicar a presença de amostras contendo o gene *aac(6')-aph(2'')* não funcional, como observado no presente trabalho. Entretanto, mais avaliações na sequência deste gene devem ser realizadas para que seja possível obter um resultado confirmatório.

As *ISSIE* e *ISSIN*, que pertencem à família *IS6* e são descritas como comuns em *Lactococcus lactis*, foram encontradas neste estudo em quatro amostras. A *ISSIN* foi descrita por Haandrikman et al. (1990), presente na região *downstream* do gene *prtM* (*proteinase maturation gene*), necessário para a atividade de proteinase no gênero *Lactococcus*. Já *ISSIE* não possui relatos na literatura e, em sua descrição no banco de dados do software *ISFinder*, é apresentada como contendo uma única ORF, responsável por codificar uma transposase. Diferente das anteriores, a *IS1297*, também descrita primeiramente em *L. lactis*, esteve amplamente presente nas amostras analisadas. Esta IS já foi encontrada nos três plasmídeos identificados na amostra de referência *E. faecium* DO (QIN et al., 2012).

Assim como *ISSIE*, não foram encontrados relatos na literatura para *ISCco2*, de *Campylobacter coli*. Na sequência depositada no banco de dados do *ISFinder*, esta IS pertencente à família *IS1595* e possui 7.852 pb divididos em oito ORFs, sendo três deles relacionadas a genes de resistência aos aminoglicosídeos. No presente estudo, a *ISCco2* esteve amplamente disseminada por toda amostragem, porém não se mostrou intacta nos genomas sequenciados.

A *ISWci2* e *IS153*, também observadas entre as amostras de *E. faecium* que compuseram o estudo, pertencem a família *IS3*. A *ISWci2* foi encontrada em uma amostra do

ST896 e a IS153, que foi identificada em apenas uma amostra do ST1626 (descrito pela primeira vez neste estudo), já havia sido descrita por Ehrmann et al. (2001) como possível marcador espécie-específico de *Lactococcus sanfranciscensis*.

Todos esses elementos genéticos móveis aqui mencionados, em sua maioria, são entidades semiautônomas que possuem um processo evolutivo próprio, que visa maximizar sua transmissão e, por isso, são altamente representados no genoma bacteriano. Sendo assim, as pressões seletivas exercidas pelo ambiente tendem a operar, de forma conflitante, nos diferentes níveis de complexidade, por exemplo, no nível do gene, do elemento genético móvel, do genoma, da população, da espécie e assim por diante. Isto posto, é importante entender que os padrões de distribuição diferenciada desses EGMs nas linhagens analisadas são consequência da emergência e expansão desses grupos, com possível contribuição desses elementos no processo, mas também de sua própria evolução desses elementos (BOYD; ALMAGRO-MORENO; PARENT, 2009; HARRISON et al., 2015; HE et al., 2016; LEVIN; BERGSTROM, 2000; RODRÍGUEZ-BELTRÁN et al., 2020).

A THG é parte importante da evolução, acarretando extensas variações no genoma bacteriano. No caso dos enterococos, mesmo em amostras que possuam uma grande identidade genômica, diversos genes únicos à essas amostras podem estar presentes (GHATTARGI et al., 2018; LAM et al., 2013; SANDERSON et al., 2020; ZHONG et al., 2019). A transferência de genes ocorre não apenas entre indivíduos relacionados, como em bactérias de uma mesma espécie ou de mesmo gênero, já tendo sido identificada entre organismos de domínios distintos, por exemplo, entre bactérias e arqueas (REST; MINDELL, 2003). Os mais recentes achados em relação aos impactos da transferência horizontal de genes na evolução bacteriana levam a discussões sobre o conceito de espécie e as reconstruções filogenéticas, e sobre a própria teoria evolutiva.

Em um estudo de 2010, o autor discute que a transferência horizontal em bactérias questiona o conceito clássico do Neo-Darwinismo, em que a evolução é vista como um processo gradual de surgimento de novas características e funções. A aquisição de genes por THG é primariamente não-genealógica e faz com que os indivíduos adquiram novidades evolutivas de forma rápida e não gradual (BOTO, 2010). O mesmo autor lembra que o impacto da transferência horizontal de genes no processo de evolução bacteriana depende de fatores como: o número de genes transferidos, o sucesso em mantê-los no genoma, a extensão do fenômeno e a distância filogenética entre os organismos.

Todavia, mesmo com os diversos fatores que estão envolvidos nos processos de THG, diferentes autores já questionaram a acurácia dos dados obtidos através de árvores

filogenéticas construídas para bactérias (GALTIER, 2007; HAO; PALMER, 2011; LAPIERRE; LASEK-NESSERQUIST; GOGARTEN, 2014; MARTIN, 1999). Em um estudo de 2007, o autor sugere que o impacto da THG na reconstrução da filogenia é que, em vez de se observar árvores, os eventos de transferência horizontal levariam a formação de teias (SAAP, 2007). Para outros, a reconstrução da filogenia somente se mostra possível e verdadeira quando feita através dos chamados *core* genes (genoma central), uma vez que estes, em teoria, nunca seriam transferidos (BROWN, 2003; RILEY; LIZOTTE-WANIEWSKI, 2009; WOLF et al. 2002).

Os impactos dos eventos de THG levam a questionamentos até mesmo mais amplos, como, por exemplo, a definição de espécies bacterianas (FRASER et al. 2009; BAPTESTE; BOUCHER, 2008). O conceito que prevalece, descreve que grupos de indivíduos que sejam genomicamente coesos e compartilhem um elevado grau de similaridade em muitas características independentes, testadas comparativamente sob condições altamente padronizadas, são considerados membros de uma mesma espécie. Em termos práticos, isso significa que uma espécie deve apresentar consistência fenotípica, com valores de hibridização DNA-DNA igual ou superior a 70% e mais de 97% de similaridade nucleotídica para rRNA 16S (GEVERS et al., 2005). Nas análises mais recentes, envolvendo o sequenciamento de genoma completo, a média de identidade nucleotídica deve ser superior a 94% (KONSTANTINIDIS; TIEDJE, 2005). De qualquer forma, com os avanços do conhecimento do genoma bacteriano, o conceito de espécie permanece em discussão. Na última década, diferentes propostas para uma melhor definição do conceito de espécie já foram descritas. Por exemplo, uma delas propõe uma abordagem de sequenciamento, onde a tipificação de múltiplos loci definiria as espécies de um gênero (KONSTANTINIDIS; 2005). Rosselló-Móra et al. (2015) já sugerem que as espécies sejam definidas como populações monofiléticas, que apresentam coerência genômica e fenotípica, em que os indivíduos podem ser claramente discriminados por meio de parâmetros padronizados, como sequências genômicas de alta qualidade ao menos para a amostra-tipo da espécie e a sequência completa do gene do rRNA 16S.

Neste estudo, também foi investigada a presença de fatores descritos por delimitar as trocas de elementos genéticos, como o sistema CRISPR-Cas e os sistemas de restrição e modificação (R-M). Os resultados obtidos não indicaram a presença de regiões associadas ao sistema CRISPR, o que corrobora as hipóteses de que em *E. faecium* esses elementos estariam sob seleção negativa e não atuariam como mecanismo de defesa (PALMER; GILMORE, 2010). Embora tenha sido muito bem caracterizado como barreira protetiva em *E. faecalis*,

para *E. faecium* a presença desse sistema é questionável (BURLEY; SEDGLEY, 2012; SANTOS et al., 2020; PRICE et al., 2016). Deve-se a isso o fato de que diversos outros estudos também não detectaram a sua presença (LINDENSTRAUSS et al., 2011; MIKSLESEN et al., 2015; RIOS et al., 2020; TREMBLAY et al., 2013), ou então identificaram sistemas sem funcionalidade (BUULTJENS et al., 2017; VAN SCHAIK et al., 2010), indicando que possivelmente os processos evolutivos levaram à seleção de amostras que não carregam esse sistema protetivo.

Por outro lado, das 74 amostras analisadas no presente estudo, apenas quatro não apresentaram o sistema R-M. Entretanto, uma extensa variação na sequência da subunidade S foi observada entre amostras de subpopulações distintas. As amostras do ST78 e ST412 compartilharam da mesma sequência da subunidade S, enquanto que amostras do ST896 e ST963 apresentaram sequências distintas. A subunidade S é responsável por especificar a sequência de DNA a ser reconhecida pelo sistema, que posteriormente, baseado em seu status de metilação, fará com que o sistema funcione tanto como uma enzima de restrição (REase), ou seja, clivando sequências não metiladas, quanto como uma metiltransferase (MTase), catalisando a metilação (TOCK; DRYDEN, 2005). Sequências de DNA reconhecidas pela subunidade S que apresentam metilação estão imunes aos processos de clivagem.

A presença de sequências distintas para a subunidade S entre essas linhagens resulta em diferentes padrões de reconhecimento, formando uma barreira frente ao fluxo ilimitado de material genético entre esses grupos e contribuindo para os processos de especiação e de isolamento genético. Esse fato pode ser observado por Huo et al. (2019), que foi o suporte de onde as sequências pesquisadas neste estudo foram extraídas. Os autores demonstraram que as amostras do clado A1 avaliadas (associadas a amostras MDR circulantes no ambiente hospitalar) apresentaram uma subunidade S distinta das observadas no clado B (amostras comensais humanas). Diante disso, sugeriram que essas diferenças no sistema R-M atuaram como barreira para a troca de material genético entre esses dois grupos e podem ter contribuído para a especiação das subpopulações de *E. faecium* que circulam no ambiente hospitalar.

As amostras do ST963 e ST896 apresentaram a subunidade S semelhante às encontradas no clado A1. Já as sequências dos grupos ST78 e ST412 apresentaram baixa homologia com todas as subunidades pesquisadas, indicando possivelmente se tratar de uma nova sequência desta subunidade. Em um estudo recente, Arredondo-Alonso et al. (2020) identificaram oito novas sequências da subunidade S que também apresentavam uma distribuição de acordo com padrões populacionais, reforçando a teoria de que esse sistema

forma uma barreira para THG e auxilia no processo de especiação. Entretanto, cabe ressaltar que o presente estudo é o primeiro a reportar variações na subunidade S dentro de linhagens associadas ao ambiente hospitalar. As diferenças de distribuição da subunidade S reportadas nos estudos anteriores foram observadas comparando amostras hospitalares e não-hospitalares.

Os genes adquiridos por transferência horizontal contribuem para a estrutura do pangenoma. Por conta da grande variabilidade que o genoma bacteriano apresenta, essa coleção de genes pode ser dividida em genes centrais (*core genes*), que compreendem aqueles que são compartilhados por todos os membros da espécie ou grupo; e os genes acessórios, que são aqueles presentes em apenas algumas amostras e refletem essa variabilidade (TETTELIN et al., 2005).

O pangenoma é descrito por possuir uma correlação negativa em relação ao seu tamanho com a proporção de *core genes* presentes nele. No presente estudo, o pangenoma das 74 amostras foi composto por 5.598 genes, sendo que destes, 1.925 pertenciam ao genoma central (*core genes*). Os chamados genomas abertos são grandes em tamanho, entretanto apresentam uma proporção pequena de *core genes* e possuem uma elevada taxa aquisição genética por meio de THG. Já os genomas classificados como fechados apresentam um tamanho menor como consequência da baixa taxa de ganho de genes, mas uma grande proporção de *core genes* (BROCKHURST et al., 2019).

Genomas abertos são comuns em bactérias com tamanho populacional grande, que são capazes de sobreviver nos mais diversos nichos ecológicos e possuem as mais diversas interações com os membros da comunidade ali presente, resultando em mais oportunidades para a transferência horizontal de genes. Diante de todas essas características, é de se observar que *E. faecium* apresenta genomas do tipo aberto (GHATTARGI et al., 2018; VAN SCHAIK et al., 2010; ZHONG et al., 2017).

Em um estudo que avaliou dez genomas de *E. faecium*, representados por amostras patogênicas e não-patogênicas, o genoma acessório apresentou 5.718 genes, enquanto que, o *core* genoma era composto de 1.935 genes (GHATTARGI et al., 2018). Essa grande diferença no conteúdo do genoma acessório, em relação ao observado nas amostras deste presente estudo, pode ser explicada pela origem de isolamento. Neste estudo, todas as amostras foram isoladas do ambiente hospitalar, o que sugere uma maior relação clonal entre elas. Em outro estudo envolvendo amostras da espécie isoladas de estações de tratamento de águas residuais, o pangenoma predito apresentou apenas 3.950 genes, sendo 49,6% desses genes pertencentes

ao *core* genoma. O genoma acessório continha apenas 1.991 genes (SANDERSON et al., 2020).

De fato, o que vem sendo descrito é que as amostras de *E. faecium* circulantes no ambiente hospitalar, em geral, apresentam um pangenoma maior, enriquecido de genes com as mais variadas funções. Um estudo de Kim et al. (2014), reunindo genomas de *E. faecium* depositados no banco de dados do NCBI, demonstrou que as amostras clínicas humanas apresentaram, em média, 182 genes a mais do que amostras não-clínicas. Os genes que contribuíram para esse aumento pertenciam aos mais diversos grupos funcionais (metabolismo de carboidratos, reparo de DNA, fatores de virulência, genes de resistência aos antimicrobianos e elementos móveis).

Todavia, ainda que a presença de elementos de resistência aos antimicrobianos e/ou fatores de virulência seja amplamente relatada em amostras de *E. faecium* circulantes no ambiente hospitalar, são os genes associados ao metabolismo de carboidratos que parecem ter um papel central na adaptação da espécie. No estudo que Lebreton et al. (2013) determinou a divisão da espécie em dois clados distintos, a aquisição de novos sistemas de fosfotransferases (PTS) para a utilização de carboidratos complexos como aminoácidos foi fundamental para a divergência do clado de amostras hospitalares (A1) a partir do clado de amostras comensais (B). Outros estudos também demonstraram que genes de metabolismo de carboidratos estavam mais relacionadas às amostras hospitalares (BAYJANOV et al., 2019; KIM; MARCO, 2014).

No presente estudo, também caracterizamos os genes centrais e acessórios para os diferentes grupos de STs circulantes. Considerando a observação apenas de grupos coesos (excluindo o grupo “Outros STs”), foi observado que o ST78 apresentou um elevado número de genes acessórios (715), sugerindo um processo evolutivo relacionado à aquisição de elementos acessórios. É admitido que a aquisição de elementos acessórios através de THG permite uma adaptação rápida ao ambiente (MEDINI et al., 2005).

Por outro lado, outra estratégia evolutiva que também é atribuída a diversos patógenos, está relacionada a redução do tamanho do genoma, principalmente pela remoção de elementos acessórios (LEE; MARX, 2012; MORAN, 2002; WOLF; KOONIN, 2013). Como mencionado anteriormente, em certas situações, um elevado número de elementos acessórios pode representar custos à célula hospedeira (CARROLL; WONG, 2018; SAN MILLAN et al., 2017; VOGWILL; MACLEAN, 2015). Considerando os dados observados neste estudo, estratégia semelhante pode ter ocorrido nas amostras pertencentes aos ST963 e ST896, que demonstraram a presença de poucos genes acessórios. No entanto, o elevado número de genes

centrais no ST896 pode também ser um indício da incorporação de genes acessórios no genoma central. A expansão do genoma central através da aquisição de elementos acessórios, que em geral trazem alguma vantagem adaptativa, pode ser uma forma de reduzir os custos associados a presença de muitos desses elementos.

Neste estudo, em termos de exclusividade genética, o ST78 foi o que obteve o maior número de genes exclusivos (59 genes), enquanto o ST896 foi o que obteve menos (12 genes). Estes resultados, em parte, são comparáveis aos obtidos por Ghattargi et al. (2018), que registraram um número de genes amostra-específicos variando entre 10 e 87 por genoma de amostras de *E. faecium*.

Também, no presente estudo, as mais diversas categorias funcionais estiveram envolvidas nesses processos de aquisição, embora muitos dos genes exclusivos em cada grupo apresentaram funções ainda desconhecidas. Dentre os que possuíam função conhecida, as categorias adquiridas variaram intra e inter-grupos de STs, como no caso das duas emergências do ST412. Estas observações provavelmente refletem alguns dos processos adaptativos que foram necessários para a expansão dessas subpopulações.

Um dos genes que se mostrou exclusivo de ST78 foi associado à produção de um sistema toxina-antitoxina (TA) chamado epsilon/zeta. Nesse sistema, ambos os componentes formam um complexo no qual a síntese e/ou a atividade da toxina é inibida pela antitoxina (FERNÁNDEZ-GARCÍA et al., 2016). Os sistemas TA estão relacionados ao processo de morte celular programada, que se ativa principalmente em condições de estresse nutricional. Nesses casos, as antitoxinas são degradadas, favorecendo a ação da toxina (MUTSCHLER; MEINHART, 2011). Em alguns casos, a toxina ativada apenas inibe processos essenciais na célula (YAMAGUCHI; PARK; INOUE, 2011).

Nos enterococos, o sistema TA epsilon/zeta está amplamente disseminado em plasmídeos que carregam determinantes de resistência à vancomicina (MORITZ et al., 2007; ROSVOLL et al., 2009), onde desempenha papel importante para a estabilidade na herança plasmidial, induzindo a morte dos descendentes privados de plasmídeo (ZIELENKIEWICZ; CEGLOWSKI, 2005). A presença exclusiva desse sistema nas amostras do ST78 pode explicar o elevado número de genes acessórios, que em geral são carregados pelos plasmídeos, identificados neste grupo.

Os dois genes encontrados exclusivamente em amostras do ST412 estão relacionados com a produção de uma beta-lactamase e com a biossíntese do ácido siálico. É reconhecido que esses dois genes são carregados conjuntamente num locus genético associado ao metabolismo de carboidratos, adquirido por amostras de *E. faecium* circulantes no ambiente

hospitalar (GILMORE; LEBRETON; VAN SCHAİK, 2013; NOVAIS et al., 2016). O locus denominado *epa-like* tem sido descrito por supostamente estar envolvido na biossíntese de um polissacarídeo antigênico de parede celular e se apresenta em pelo menos quatro versões distintas. Dentre as variantes descritas, uma delas apresenta os genes *neuABCD* e a beta-lactamase inseridos na região *downstream* do locus (PALMER et al., 2012; QIN et al., 2012). Múltiplos eventos de recombinação na região *epa-like* têm sido observados e sugeriram variações capsulares em *E. faecium* (DE BEEN et al., 2013). A variação capsular é uma estratégia bem-sucedida que bactérias patogênicas utilizam para evadir do sistema imune (BEDDEK et al., 2009; MERINO; TOMÁS, 2015; MILLARD et al., 2012; TEMIME et al., 2008), que também pode ter contribuído no sucesso dessa linhagem.

Outro processo substancial para a evolução bacteriana é a ocorrência de mutações pontuais. No presente estudo, o número de SNPs de cada amostra foi obtido através de sua comparação com a amostra Aus0004, posteriormente foi obtida uma média para cada grupo ST. O grupo denominado “Outros STs” obteve o maior número de SNPs e isso certamente é decorrente deste grupo representar a reunião de amostras que apresentaram um grande distanciamento filogenético. A observação de um elevado número de SNPs no ST412, segundo grupo com mais variações, despertaram hipóteses relativas ao processo evolutivo desse grupo.

O baixo grau de diversidade observado entre as principais linhagens que circularam nas instituições hospitalares do estado do Rio de Janeiro, durante o período do estudo, com médias de SNPs que variaram de 4.895 a 6.069, corroborou o que já havia sido reportado sobre a expansão clonal de amostras hospitalares. Um estudo de Raven et al. (2016), com amostras de *E. faecium* isolados em hospitais do Reino Unido e da Irlanda, durante o período de 2001 a 2011, indicou que a expansão do clado A (associado a amostras hospitalares) apresentou uma baixa diversidade genética (média de 2.174 SNPs) quando comparado ao clado B (amostras comunitárias), que obteve uma média de 11.673 SNPs.

Entre as variações genômicas observadas, foi identificado um elevado número de processos de inserção e deleção de nucleotídeos, eventos importantes que causam deslocamento do quadro de leitura, afetando a síntese proteica. Apesar de observado apenas para o grupo do ST896, já foi sugerido que o mais comum é a ocorrência de um maior número de deleções do que de inserções (MIRA; OCHMAN; MORAN, 2001).

Esses eventos ocorrem quando a pressão seletiva não é forte o suficiente para manter determinados genes, que podem ser mais facilmente eliminados se estiverem inativados (BOBAY; OCHMAN, 2017; BRYANT; CHEWAPREECHA; BENTLEY, 2012) De fato,

dentro das linhagens que se expandiram no ambiente hospitalar, o ST896 foi o que obteve o maior número de deleções e o menor número de genes acessórios, o que pode sugerir que sua evolução incluiu a redução do genoma pela eliminação de genes após eventos de deleção.

A importância dos efeitos mutacionais para o processo evolutivo foi corroborada diante dos resultados observados na avaliação de novidades genéticas (GNU). Dentro das variações presentes no genoma, uma média de 86 GNUs foi identificada, não sendo encontradas em nenhum outro genoma presente no banco de dados que foi construído. Os resultados reforçaram os arranjos obtidos para a determinação das relações filogenéticas observadas, demonstrando que um menor número de GNUs foi compartilhado entre STs que apresentaram um maior distanciamento genético. Um dado interessante observado dentre os STs de interesse se refere a linhagem ST78 que apresentou o menor número médio de SNPs. Por outro lado, esta linhagem foi a que obteve um dos maiores números de unidades de novidade genética exclusivas.

As mutações podem levar a efeitos de alto impacto pois o ganho ou a perda de um códon de parada, resulta em sequências truncadas, também chamadas de pseudogenes (TUTAR, 2012). Diversos estudos já demonstraram que, embora os pseudogenes estejam frequentemente presentes no genoma bacteriano, os processos evolutivos tendem a levar a eliminação dessas sequências não funcionais (KUO; OCHMAN, 2010; LERAT; OCHMAN, 2005; MIRA; OCHMAN; MORAN, 2001). A presença de pseudogenes amplamente distribuídos na população reflete os diferentes impactos do processo evolutivo na arquitetura genômica bacteriana. O genoma conservado possui uma origem ancestral e reflete a evolução vertical da espécie; sua divergência está mais associada ao acúmulo lento de polimorfismos. Por outro lado, o genoma acessório reflete uma evolução mais recente, ocorrida lateralmente e de forma mais rápida, por isso tende a ser menos impactado por polimorfismos. Sua variação é dada pelas diferenças de conteúdo (AZIZ; NIZET, 2010; CORDERO; POLZ, 2014).

A presença de sequências com alterações disruptivas tem impacto não apenas na evolução da arquitetura do genoma, como também tem efeitos na funcionalidade da célula bacteriana (DINA; MALBRUNY; LECLERCQ, 2003; TEIXEIRA et al., 2012). Em uma amostra de *E. faecium*, portadora do gene *vanA*, Kerbauy et al. (2011) demonstraram que uma inserção no gene *ddl*, responsável pela produção da proteína D-Alanina-D-Alanina ligase, envolvida no processo de síntese da parede celular, acarretou a geração prematura de um códon de parada. Como efeito fenotípico, a amostra apresentou um crescimento dependente da presença de vancomicina.

Uma dessas proteínas afetadas pelo códon de parada foi a proteína de reparo de DNA *mismatch*, que esteve associada às amostras do ST412, e pode ser a explicação para o elevado acúmulo de mutações nesse grupo. Os sistemas MMR possuem importância central na integridade do material genético ao reparar incompatibilidades (*mismatch*) de pareamento, ocorridas após o processo de replicação do DNA (FUKUI, 2010). A perturbação desses processos de reparo leva à instabilidade do genoma bacteriano, que, no entanto, é essencial para o processo evolutivo (WILLIAMS, 2016).

A instabilidade é capaz de gerar a diversidade genética necessária para os processos de adaptação, seja pela permissividade em gerar mutantes como também por facilitar a aquisição de elementos por THG (DARMON; LEACH, 2014). Em *E. faecium*, a formação de hipermutantes por erros no funcionamento do sistema de reparo está associada a emergência de amostras resistentes aos antimicrobianos como a linezolida e daptomicina, que são de última escolha para o tratamento das infecções por este microrganismo (MATONO et al., 2016; WILLEMS et al., 2003).

A resistência aos β-lactâmicos e às quinolonas também está relacionada a eventos de mutação. No presente estudo, a avaliação das mutações nos genes *gyrA* e *parC* revelaram padrões de substituição associados às mudanças populacionais. As substituições encontradas estão de acordo com o reportado na literatura. Em um estudo de 2019, Ostrer et al., analisando um extenso banco de dados de sequências dos genes *gyrA* e *parC*, demonstraram que em *E. faecium* a posição 84 do gene *gyrA* é a que apresenta a maior frequência de mutação (0,64). Já em *parC*, a posição 82 é mais frequentemente alterada (0,66). Assim, os autores estabeleceram que o caminho hierárquico para a aquisição de resistência às quinolonas seria: *parC82* → *gyrA84* → *gyrA88* → *parC86*. No presente estudo, não foram identificadas alterações nas posições 88 de *gyrA* e 86 de *parC*.

Em um estudo com amostras isoladas no Japão, Urushibara et al. (2018) observaram que no gene *gyrA* a substituição Ser84Ileu era a mais frequente, seguida por Ser84Arg e Ser84Tyr. Esses resultados diferiram do observado para nossas amostras, em que a presença de isoleucina nessa posição foi menos encontrada. Já para *parC*, nosso estudo obteve o mesmo padrão observado por Urushibara et al., 2018, com a predominância da substituição Ser82Arg. Em relação aos genes *gyrB* e *parE*, poucos estudos já relataram mutações nesses dois genes. Em 2020, Yang et al. encontraram uma amostra apresentando a substituição de serina por glicina na posição 378 do gene *parE*. No gene *gyrB*, Tremblay et al. (2013) encontraram alterações nas posições 371 (leucina→triptofano), 436 (ácido aspático→asparagina) e 455 (prolina→serina). Cabe ressaltar que no presente estudo foi

possível identificar três amostras que não apresentaram, concomitantemente, mutações nos genes *gyrA* e *parC*, mas possuíam alterações nos genes *gyrB* e/ou *parE*. Entretanto, para confirmar que a presença dessas alterações teria participação no fenótipo de resistência, análises adicionais seriam necessárias.

O gene *pbp5*, associado a resistência aos β -lactâmicos, apresentou uma série de polimorfismos em sua estrutura. A maior parte destas substituições já foi reportada na literatura (GALLOWAY-PEÑA; RICE; MURRAY, 2011; HSIEH et al., 2006; KLIBI et al., 2008; LÓPEZ et al., 2009; POETA et al., 2007; RUIZ-RIPA et al., 2020; SADOWY; LUCZKIEWICZ, 2014; SILVA et al., 2018; TREMBLAY et al., 2013; VALDEZATE et al., 2012). Entretanto, até onde foi possível observar, as alterações nas posições 56 (leucina→metionina), 227 (alanina→serina); 378 (alanina→ácido glutâmico); 401 (alanina→serina); 408 (glutamina→histidina/tirosina); 476 (alanina→serina) e 497 (fenilalanina→leucina), observadas pontualmente em algumas amostras, neste estudo, ainda não haviam sido descritas.

Embora uma grande variação em relação a esse gene seja observada, a substituição da metionina por alanina na posição 485, localizada dentro do sítio ativo da molécula, é que vem sendo descrita como responsável pelos fenótipos de resistência, além de uma inserção na posição 466 (MOHN et al., 2004; RICE et al., 1995; SIFAOUY et al., 2001). O elevado acúmulo de mutações nesse gene pode ser explicado por sua própria importância no processo de síntese da parede celular, uma vez que já foi postulado que os genes que são altamente expressos tendem a apresentar uma taxa de mutação maior que o restante do genoma (BRYANT; CHEWAPREECHA; BENTLEY, 2012).

Dentre os diversos perfis de variação no gene *pbp5*, dois corresponderam a quase 90% das amostras (P1 e P2). E assim como observado para os genes *gyrA* e *parC*, pode-se observar uma certa distribuição relacionada aos diferentes STs. No estudo de Novais et al. (2016), os autores observaram mutações no gene da PBP5 estavam relacionadas com a diversidade filogenética. Além disso, comprovaram que esse gene pode ser transferido entre as amostras. Posteriormente, Morroni et al. (2019) corroboraram esses achados, ao observar a transferência de *pbp5* através de plasmídeos.

Diante da observação de que algumas das substituições nos genes anteriormente mencionados não apenas emergiram junto com as novas linhagens, como também foram disseminadas nas gerações seguintes, neste estudo foram investigados os processos evolutivos que poderiam estar envolvidos nesse fenômeno. O teste de neutralidade D de Tajima

distingue se uma sequência de DNA evolui de forma aleatória, através da neutralidade, ou por um processo não-aleatório (selecionista) (TAJIMA, 1989).

Proposta por Motoo Kimura (1968), após a observação, maior do que a esperada, de polimorfismos no genoma, a teoria neutralista da evolução sugere que a maior parte das mutações no genoma não apresentaria vantagens adaptativas, sendo consideradas neutras, e que a evolução no nível molecular é dominada por processos aleatórios como a deriva gênica. O mesmo autor ressaltou que esta teoria não se apresenta como uma oposição à teoria da seleção natural proposta por Darwin (KIMURA, 1968).

Dessa forma, pode-se observar que genes *gyrA* e *pbp5* não estiveram sobre processos de evolução neutra. Suportados estatisticamente, os dois genes tiveram valores de D estimados em 2,54 (*gyrA*) e -2,43 (*pbp5*). Em geral, valores de D acima de zero indicam a existência de poucas variantes raras, uma situação que pode ser encontrada em processos de seleção balanceadora ou estrutural. Por outro lado, quando o valor de D se encontra abaixo de zero, as variantes raras são abundantes, situação que pode ser encontrada em processos de expansão populacional após “efeitos de gargalo”, ou mesmo em processos de seleção, seja ela positiva ou purificadora.

A recombinação genética é tida como uma das maiores forças que governam a evolução bacteriana. Foi primeiramente identificada em decorrência da observação da formação de mosaicismos no locus que codifica moléculas imunogênicas e de resistência aos antimicrobianos (FEAVERS et al., 1992; SPRATT et al., 1992). O processo de recombinação em bactérias é unidirecional e definido como a incorporação de DNA de uma célula doadora para uma receptora. Os eventos de recombinação podem ser classificados de acordo com o mecanismo pelo qual o DNA é introduzido na célula hospedeira em: transdução mediada por bacteriófagos, conjugação ou transformação (VILLA et al., 2019). A recombinação é baseada no tipo de material genético transferido, podendo ser uma sequência homóloga, que limitará a variação a no máximo o aparecimento de um novo alelo para aquela região; ou não-homóloga, em que um novo gene pode ser transferido (VOS; DIDELOT, 2009).

Uma importante função biológica primária gerada por recombinação homóloga é o reparo de DNA (REDFIELD, 2001). A troca de segmentos de DNA idênticos entre grupos bacterianos filogeneticamente relacionados ou mesmo entre membros da mesma espécie serve como molde para o reparo de danos no DNA, que não tenham sido naturalmente reconhecidos pela maquinaria que a célula possui como, por exemplo, aqueles causados por quebras da fita ou alterações estruturais (MICHOD; BERNSTEIN; NEDELCO, 2008).

No presente trabalho, a taxa estimada de recombinação no genoma da população total (R/θ) foi de 0,0680, com um impacto relativo de recombinação para a mutação (r/m) de 1,87. Um estudo de Homan et al. (2002) com amostras de *E. faecium* isoladas de diferentes origens estimou um valor de r/m de 1,1. Em um estudo que avaliou um conjunto de 146 amostras isoladas numa instituição hospitalar australiana, durante diferentes períodos, os autores estimaram um valor de r/m de 6,1 com R/θ de 0,22 (VAN HAL et al., 2016). Outros estudos também já relataram taxas elevadas de recombinação em amostras de *E. faecium* (DE BEEN et al., 2013; TURNER et al., 2007; WILLEMS; VAN SCHAİK, 2009). Todavia, cabe ressaltar que variação nas taxas de recombinação são comumente relatadas, seja devido a diferentes metodologias utilizadas nas análises, seja pelo tipo de região analisada (VOS, 2009).

Anteriormente, taxas de recombinação eram comumente estimadas a partir dos genes de MLST (DEN BAKKER et al., 2008; RAVEN et al., 2016; SHEPPHARD et al; 2008; WILLEMS et al., 2012; WIRTH et al., 2007). Porém, com o advento do sequenciamento de genoma, os eventos de recombinação passaram a ser avaliados mais amplamente. Além disso, os métodos de análise são influenciados em relação a amostragem a ser estudada e taxas de recombinação podem diferir ao serem analisados representantes de uma população inteira ou apenas de uma linhagem específica daquela população (DIDELOT; MAIDEN, 2010).

Ainda que diferentes taxas de recombinação tenham sido observadas em estudos relativos a *E. faecium*, é posto que os processos de recombinação homóloga foram fundamentais para evolução da espécie. Esses eventos levaram à diversificação das populações e foram importantes para a emergência de linhagens mais adaptadas no ambiente hospitalar, como o anteriormente denominado, complexo clonal 17 (WILLEMS et al., 2005). Os eventos de recombinação que atingiram o *core* genoma também tiveram papel fundamental no processo adaptativo da espécie no ambiente hospitalar (DE BEEN et al., 2013). Adicionalmente, Garcia-Solache et al. (2016), descreveram que a ocorrência de recombinação homóloga em grandes trechos do cromossomo facilitou a aquisição de resistência aos β -lactâmicos pela espécie.

Como observado no presente estudo, os eventos de recombinação tiveram um impacto significativo na reconstrução da filogenia das amostras estudadas. Ainda que a organização das relações filogenéticas tenha se mantido, a filtragem dos eventos de recombinação levou ao encurtamento dos braços e a uma maior coesão entre os diferentes clados. As mesmas observações foram feitas no estudo de De Been et al. (2013), que reconstruiu a filogenia dos dois principais clados de *E. faecium* em humanos (Clado A e B), utilizando o alinhamento de

SNPs obtidos do concatenamento dos genes que compõem o *core* genoma. Os achados de nosso estudo corroboram a hipótese de que os eventos de recombinação tendem a influenciar as reconstruções filogenéticas, como postulado por Posada et al. (2002).

O presente estudo também identificou que a maioria dos eventos de recombinação ocorreu em nodos ancestrais, com poucos eventos sendo detectados individualmente nas amostras analisadas. A observação da ocorrência de poucos eventos de recombinação nas amostras nos leva a considerar a hipótese de que esse resultado seja ratificado pelos achados de Willems et al. (2012), que propôs que ao adentrar no ambiente hospitalar essas linhagens especializadas começaram um processo de isolamento genético, levando ao declínio na frequência de eventos recombinatórios.

Diversos outros patógenos experimentaram um decréscimo dos eventos de recombinação, como consequência por se tornarem altamente especializados, como por exemplo *Yersinia pestis* (ACHTMAN et al., 2004), *Bacillus anthracis* (VAN ERT et al., 2006) e *Mycobacterium tuberculosis* (LIU et al., 2006). Com a adaptação e a especiação, a oportunidade de recombinação com outras linhagens diminui e tende a ficar restrita àquelas que compartilham o mesmo nicho. Uma ocorrência de recombinação homóloga entre grupos próximos é mais provável primeiro em razão da proximidade física em que se encontram e segundo por conta da dependência de homologia para que o evento venha a ter sucesso (DIDELOT; MAIDEN, 2010).

Os dados obtidos neste estudo, relativos ao processo recombinação, quando avaliados em uma representação gráfica (como em Resultados – Figura 19), revelaram que esses eventos dentro dessas linhagens apresentaram algum grau de homoplasia. Os processos homoplásicos não decorrem a partir de uma ancestralidade comum, mas sim por múltiplas causas; dentre elas a convergência evolutiva que, entre outros fatores, pode ser produzida através do aumento na pressão seletiva pelo uso massivo de antimicrobianos, o que favorece o surgimento de múltiplas populações independentes (FARHAT et al., 2013; ROKAS; CARROLL, 2008). Entretanto, os processos de recombinação homóloga também podem ser responsáveis pela introdução de sítios homoplásicos e são um indicativo de sua ocorrência (CRISPELL; BALAZ; GORDON, 2019).

Para observar se os eventos de recombinação que ocorreram tanto nos nodos ancestrais quanto nas amostras decorreram de trocas de material genético apenas entre as linhagens circulantes, ou se havia o envolvimento de outras populações, seriam necessárias análises em relação ao nível de mistura dessa população. De Been et al. (2013) sugeriram que os maiores

reservatórios de material genético para o clado de amostras hospitalares (Clado A) eram amostras comensais humanas (Clado B).

Embora tenha sido observado que quase todos os nodos da árvore de filogenia foram impactados por eventos de recombinação, alguns merecem destaque e podem indicar mais sobre a emergência e expansão de algumas linhagens. Ao que parece, as duas emergências do ST412, detectadas neste estudo, podem ter ocorrido por processos de recombinação homóloga, observados pela presença de eventos no nodo ancestral. Entretanto, a expansão dessas duas subpopulações se deu por outros mecanismos, uma vez que não foram observados eventos de recombinação nos nodos do ancestral comum mais recente (MRCA).

Por outro lado, o surgimento e expansão do ST963 e ST896 pode estar relacionado com processos de recombinação, uma vez que todos os nodos ancestrais, incluindo o do MRCA, foram afetados por eventos dessa natureza. O MRCA do ST896 foi o segundo mais afetado por eventos de recombinação, oito eventos no total com uma extensão de 47.934 pb. Os nodos relacionados ao surgimento e expansão do ST963 somam juntos 14 eventos de recombinação com 218.061 pb de bases afetados, sendo que o nodo NODO 91, ancestral que deu origem a este clado, foi o que apresentou sozinho a maior extensão de eventos de recombinação, 100.427 pb. Estes dados nos remetem ao estudo de Buultjens et al. (2017), que determinou a origem evolutiva do ST796, um clone epidêmico na Austrália, através da utilização de genômica comparativa. Os autores descreveram que a mais provável explicação para o surgimento dessa linhagem foi a ocorrência de eventos de recombinação que foram detectados em seu ancestral recente mais comum. Os mesmos autores relataram que uma região genômica de 170 kb, constituída de genes codificadores de proteínas de transporte associadas à parede celular, incluindo os translocadores de cádmio e cobre, foi envolvida nos eventos recombinatórios propulsores do surgimento do referido clone (BUULTJENS et al., 2017).

Neste estudo, de forma interessante, foi observada a presença de eventos de recombinação no MRCA que divide o ST78 do ST412. Entretanto, o mesmo não se deu para o nodo que divide o ST412 do ST963/ST896. Estes achados podem ser resultado do que foi identificado quanto às diferenças na subunidade S do sistema R-M. Pode-se supor que, por compartilharem a mesma subunidade de especificação, as trocas de material genético entre o ST78 e o ST412 poderiam ocorrer de forma mais constante. Por outro lado, como não possuem a mesma sequência da subunidade S presente nos ST963 e ST896, as trocas de material entre esses grupos podem ter sido impedidas por esse sistema.

O impacto dos eventos de recombinação também foi avaliado nos genes que fazem parte do esquema de caracterização por MLST, bem como naqueles relacionados a resistência às quinolonas e aos β -lactâmicos, particularmente, por terem um padrão de distribuição característico nas diferentes linhagens identificadas. Os resultados revelaram que apenas os genes *atpA* e *ddl* estavam presentes em áreas sob efeito de recombinação. Os alelos para o gene *atpA* são semelhantes nas quatro linhagens de maior frequência. Entretanto, a mudança do alelo 1 (ST78-ST412) para o alelo 2 (ST963-ST896) no gene *ddl* pode residir na presença de eventos de recombinação nesse gene dentro do respectivo nodo ancestral. A mudança do alelo de *purK* não pode ser explicada por processos de recombinação. Além disso, como já demonstrado, este gene também não se apresentou sob processos de seleção adaptativa.

A presença de eventos de recombinação em genes de manutenção celular que compõem o esquema de MLST já foi amplamente reportado na literatura (HOWDEN et al., 2013; RAVEN et al., 2016; VAN HAL et al., 2016; WILLEMS et al., 2010) e, por essa razão, a reconstrução da filogenia baseada em algoritmos como o eBURST não se mostram verdadeiras. A análise da filogenia utilizando inferência bayesiana tem se mostrado mais eficaz para demonstrar as relações evolutivas de *E. faecium* e permitiu determinar que a formação do anteriormente chamado CC17 não se deu a partir de um único fundador (ST17), mas sim por múltiplos fundadores (WILLEMS et al., 2012).

Neste estudo, a análise bayesiana corroborou os achados de estrutura populacional e as relações filogenéticas. A observação de que o grupo do ST412 apresentava uma origem polifilética com dois processos de emergência e expansão foram confirmadas. O mesmo ocorreu para o ST963, em que foi possível observar de uma forma mais clara a divisão desse grupo em dois subclados. A taxa de mutação observada na população foi de $2,34 \times 10^{-7}$ SNPs por sítio por ano.

Num estudo similar, Raven et al. (2017), analisando 293 amostras de *E. faecium* isoladas de uma instituição hospitalar no Reino Unido no período de 2006 a 2012, obtiveram, através do emprego de análise bayesiana, uma taxa de mutação de $2,3 \times 10^{-6}$ SNPs por sítio por ano, valor dez vezes menor que o observado no presente estudo. A taxa de mutação encontrada para amostras pertencentes ao ST17 e ST252 que foram isoladas em instituições hospitalares australianas num período de 12 anos também foi menor do que a relatada no presente estudo, $1,5 \times 10^{-6}$ SNPs por sítio por ano (HOWDEN et al., 2013). Assim, os achados sugerem fortemente que as amostras de *E. faecium* que circularam nas instituições hospitalares do Rio de Janeiro apresentaram uma evolução mais rápida do que tem sido observado em instituições hospitalares dos países mencionados.

No estudo que propôs a divisão de *E. faecium* em diferentes clados, os autores determinaram que a taxa de mutação do clado A1, onde se encontram as amostras que circulam no ambiente hospitalar, seria de $1,49 \times 10^{-5}$ SNPs por sítio por ano. As amostras do clado A2, que estão associadas a isolados animais, tiveram uma taxa de mutação bastante superior $3,6 \times 10^{-6}$ SNPs por sítio por ano. Nas amostras comensais, clado B, a taxa de mutação foi de $1,3 \times 10^{-5}$ SNPs por sítio por ano. Entretanto, como já mencionado, dados de outros autores não suportam essa subdivisão do clado A (LEBRETON et al., 2013).

Em um estudo recente, Rios et al. (2020) estimaram a taxa de mutação de todo o clado A em $1,21 \times 10^{-6}$ SNPs por sítio por ano. Quando avaliado o subclado onde estavam o ST78 e ST412, a taxa de mutação foi de $4,67 \times 10^{-7}$ SNPs por sítio por ano. Já as amostras do ST17 e ST18, que pertenciam a outro subclado, obtiveram uma taxa de mutação estimada em $6,43 \times 10^{-7}$. Esses valores, embora maiores do que o observado no presente estudo, se apresentaram na mesma ordem de grandeza. Entretanto, cabe ressaltar que variações na estimativa da taxa de mutação podem ocorrer por diversas razões, desde o conjunto de amostras avaliado até o tipo de análise empregada para obtenção dos dados.

As análises das taxas de mutação também permitiram estimar o tempo de divergência clonal e datar o MRCA das diferentes linhagens observadas. Os dados indicaram que a emergência do ST78 se iniciou em 1999 e que cerca de seis anos depois, em 2005, ocorreu o início da primeira emergência do ST412. Esta foi prontamente seguida pela sua segunda emergência, com MRCA datado de 2006. Neste mesmo ano, ocorreu o surgimento dos dois MRCAs que deram origem ao ST963. A emergência do ST896 ocorreu mais recentemente, em 2012.

Além disso, os resultados obtidos neste estudo apontaram que a divisão dessas grandes linhagens ocorreu em um período anterior: (i) a divisão do ST78 das linhagens ST412/ST896/ST963 ocorreu em 1958, ano em que foi estimado o surgimento do MRCA desses grupos; (ii) a subsequente divisão do ST412 com o ST963/ST896 ocorreu seis anos depois, em 1964, quando foi posicionado o MRCA deste grupo; (iii) a posterior divisão do ST963 e ST896 ocorreu já no século XXI, no ano de 2004.

Entretanto, a emergência de *E. faecium* no ambiente hospitalar parece ter ocorrido em um período mais distante, embora as datas específicas ainda sejam controversas. No primeiro estudo que propôs a divisão de *E. faecium* em um clado hospitalar e um clado de amostras comunitárias, os autores sugeriram que a separação desses grupos ocorreu há 300 mil anos (GALLOWAY-PEÑA et al., 2012). Posteriormente, outro estudo que sugeriu que as amostras humanas que circulavam no ambiente hospitalar eram relacionadas com amostras animais, e

propôs a divisão em clado A1 (amostras humanas hospitalares) e clado A2 (amostras animais), os autores determinaram que a separação desses cladogramas e a emergência de *E. faecium* no ambiente hospitalar teria ocorrido 75 anos antes (1938), concomitante com a introdução dos antimicrobianos. A separação do clado A2 do clado B, que seriam as amostras comensais humanas, teria ocorrido há 3 mil anos, mesmo período que registra o início da domesticação animal (LEBRETON et al., 2013).

Em um estudo recente, os autores sugeriram que a separação do clado A do clado B teria ocorrido há 2.765 anos. A emergência de amostras humanas que circulam no ambiente hospitalar também teria datado em torno de 502 anos atrás. As posteriores divisões de linhagens que circulam dentro do ambiente hospitalar, no caso o subclado ST78-ST412 e o subclado ST17-ST18, teriam ocorrido em torno do ano de 1718, ou seja, há 302 anos (RIOS et al., 2020).

Através das datas estimadas para emergência das diferentes linhagens, foi possível neste estudo reconstruir uma linha do tempo, que demonstrou a ocorrência de diversas ondas de circulação de STs. As diferentes linhagens apresentaram um respectivo período de emergência, de expansão, e sequencialmente um decréscimo da sua circulação (declínio da curva), que se deu no mesmo momento em que outros STs começaram sua expansão.

Manter a diversidade genotípica representa uma estratégia evolutiva pois, primariamente, aumenta as chances de que um genótipo adaptativo emergir e se fixar. A longo prazo, a circulação concomitante de diferentes linhagens tende a influenciar o sucesso ecológico dessa população naquele ambiente, uma vez que, se qualquer tipo de mudança extrema ocorrer, haverá uma probabilidade maior de que ao menos um deles seja capaz de se adaptar, garantindo a sobrevivência da população. Além disso, a diversidade genotípica aumenta a variedade de ambientes e recursos aos quais a população está adaptada, diminuindo a competição entre essas linhagens (CORDERO; POLZ, 2014; GALHARDO; HASTINGS; ROSENBERG, 2007; HUGHES et al., 2008).

Com o emprego de testes estatísticos, foi possível prever que algumas características estavam associadas aos diferentes grupos de STs estudados. Tais análises demonstraram que o ST78 foi significativamente o que apresentou o menor número de SNPs e o maior número de CDSs preditos, bem como em seu genoma havia uma maior presença de ISs e profagos. Assim, pode-se sugerir que a emergência e expansão do ST78 se deu principalmente por alargamento do seu genoma acessório. Este fato torna-se mais relevante considerando que esse grupo de ST apresentou comparativamente o maior número desses genes e demonstrou portar um sistema toxina-antitoxina associado a estabilização de elementos móveis. Ainda que

estatisticamente tenha apresentado menos mutações que o restante dos grupos, um grande número dessas variações foi de novidades genéticas exclusivas desse grupo, sugerindo que nesse caso tenha acarretado algum benefício evolutivo para este grupo.

Considerando que um dos principais objetivos deste estudo consistiu na busca dos eventos que determinaram a emergência do ST412, os resultados obtidos nos levam a sugerir as seguinte propostas para responder esta questão: (i) as duas emergências do ST412 podem ter ocorrido após os eventos de recombinação observados em seu nodo ancestral; (ii) a expansão dessas duas linhagens, possivelmente, ocorreu por processo de formação de hipermutantes, corroborada pela observação de um alto número de SNPs nesse grupo, possivelmente explicado pelo ganho de um códon de parada num sistema de reparo de DNA; (iii) seu sucesso no ambiente hospitalar pode ter sido decorrente do ganho exclusivo de genes de parede celular, promovendo variações antigênicas em componentes associados ao ácido siálico (*neuABC*).

Além disso, pode-se sugerir que os processos de recombinação no MRCA dos ST963 e ST896 levaram a emergência desses grupos. Posteriormente, um balanço entre a redução do genoma acessório, através da expurgação de genes afetados por deleções e a expansão do genoma central, com uma maior aquisição de genes de resistência e de genes de metabolismo estatisticamente significativa resultaram no sucesso do ST896 no ambiente hospitalar. O grande número de unidades de novidade genética e o ganho de genes exclusivos associados à parede celular, metabolismo de aminoácidos e biossíntese de metabólitos secundários podem ter sido responsáveis pelo sucesso do ST963.

Ante o exposto, fica clara a importância de se elucidar a estrutura populacional e os processos evolutivos de um microrganismo que possui notável relevância no cenário hospitalar e que parece que cada vez mais tem se tornado apto a adquirir um arsenal de mecanismos que aumentam sua adaptação e circulação nesse ambiente. A determinação da estrutura populacional e a observação de diversos mecanismos evolutivos podem apontar as pistas que auxiliem as ações de controle dos futuros processos que possam resultar na emergência de novas linhagens ainda mais adaptados. As mudanças importantes no cenário epidemiológico de *E. faecium* sustentam ainda mais a necessidade do rápido rastreamento dessas linhagens, para que possa basear ações efetivas na melhoria de protocolos de controle, detecção e intervenção.

Ademais, uma das importantes observações feitas neste estudo diz respeito ao rápido processo evolutivo no qual a espécie *E. faecium* está envolvida e o quanto esse tempo é superior a capacidade humana de criar novas ferramentas de combate e controle. A evolução é

um fato e o surgimento de linhagens ainda mais adaptadas é, possivelmente, apenas uma questão de tempo. É de fato premente que sejam continuamente avaliados os tipos de conduta que são aplicadas nos ambientes hospitalares, o funcionamento dos sistemas de saúde, o uso consciente de substâncias antimicrobianas e, de maneira ainda mais geral, substâncias poluentes que impactam o ambiente. Desse modo, pode-se estender às próximas gerações a chance de manter o uso de antimicrobianos para o tratamento eficaz das doenças infecciosas bacterianas. Caso contrário, até mesmo infecções, ainda que consideradas como de menor gravidade, podem vir a configurar sérios problemas futuros de saúde pública.

CONCLUSÃO

Neste estudo foi determinada a estrutura populacional e a análise dos processos evolutivos que ocorreram em amostras de *Enterococcus faecium* que circularam em instituições hospitalares do estado do Rio de Janeiro no período de 2009 a 2016.

A caracterização inicial de 205 amostras bacterianas revelou taxas elevadas de não suscetibilidade a pelo menos dez antimicrobianos distintos, caracterizando a maior parte da amostragem como multirresistente. Determinantes de resistência comuns na espécie, detectados através de testes de PCR, apresentaram ampla distribuição na amostragem avaliada.

Foram sequenciados os genomas de 74 dessas amostras, cuja análise indicou que pelo menos 13 STs distintos circularam no período estudado, e confirmou a substituição da linhagem anteriormente circulante ST78, pelo ST412. Além disso, foi identificada a emergência e circulação de outras duas linhagens não pertencentes ao complexo clonal circulante no ambiente hospitalar CC17, ST963 e ST896. A reconstrução filogenética permitiu determinar que as principais linhagens circulantes pertenceram a diferentes braços filogenéticos e indicaram que múltiplas linhagens distintas circularam em nosso meio, no período avaliado.

Essas linhagens emergentes resultaram em mudanças na estrutura populacional da espécie, sugerindo a presença do alelo *purk44* como marcador desses novos STs circulantes. Além disso, a avaliação do conteúdo genético permitiu identificar padrões diferenciados de circulação de genes associados à resistência e à virulência, como também de elementos móveis. Os diferentes perfis de substituição para genes de resistência às quinolonas e aos β -lactâmicos também acompanharam o padrão de circulação dessas linhagens. Porém os resultados apontaram que alguns desses genes não acompanharam um padrão evolutivo de neutralidade.

Com a utilização dos recursos de análise do sequenciamento do genoma completo, foi possível identificar ao menos 21 genes distintos associados à resistência aos antimicrobianos e 13 genes associados à determinantes de virulência, bem como quatro genes envolvidos na resistência aos metais pesados. Alguns genes apresentaram uma ampla distribuição na população. Ressalta-se que foram identificados genes com pouco relato na literatura, como *tet(S)* e *tet(U)*, relacionados a resistência à tetraciclina, e *eatA*, associado ao fenótipo LS_AP (lincosamidas-streptogramina A-pleuromutilinas).

Uma extensa diversidade de famílias de plasmídeos, ISs e profagos foi identificada na amostragem, incluindo alguns que possuem como origem outras espécies bacterianas. A detecção de um mobiloma diverso reforçou o conceito da ampla plasticidade genética de *E. faecium*.

A análise conjunta de técnicas fenotípicas de caracterização da susceptibilidade a níveis elevados de aminoglicosídeos e sequenciamento do genoma permitiu identificar a ocorrência de incongruências fenótipo-genótipo. Estes achados foram fundamentais para caracterização de uma variante do gene *ant(6')-Ia* que apresentou modificações estruturais que levaram a perda de sua funcionalidade.

Resultados relativos ao encontro de variações no sistema de defesa R-M permitiram supor que estas podem ter ajudado no processo de expansão de algumas linhagens, por atuarem como barreira para troca de material genético.

A composição do genoma central e do genoma acessório foi bastante diversa entre os grupos de STs e demonstrou diferentes estratégias evolutivas para a manutenção da arquitetura genômica. A expansão do genoma acessório através da aquisição de genes acessórios foi importante para o processo adaptativo do ST78, corroborado pelo elevado número de IS e profagos identificados nesse grupo.

Os processos mutacionais levaram a uma baixa diversidade das amostras, porém se mostraram importantes para as linhagens ST412 e ST78, principalmente quando associados à efeitos de alto impacto nas proteínas e geração de unidades de novidade genética.

Os eventos de recombinação foram de extrema importância para o surgimento das linhagens ST896 e ST96, sendo identificados primordialmente em nodos ancestrais. Além disso, foram responsáveis por possíveis mudanças de alelo no gene *ddl* dentro dessas subpopulações.

Os eventos que determinaram a emergência e expansão da linhagem ST412 estiveram relacionados a processos de recombinação em seu nodo ancestral e, posteriormente, a formação de hipermutantes através de modificações em sistema de reparo de DNA. O ganho de genes de parede celular, promovendo a formação de variantes antigênicas pode também ter contribuído para o seu processo adaptativo.

A inferência bayesiana permitiu estimar o tempo de divergência clonal e indicou a presença de sucessivas ondas de emergência e expansão de STs iniciada a partir do ano de 1999. Essas expansões estiveram relacionadas com diversos processos evolutivos no genoma, como a aquisição de material genético, eventos mutacionais e de recombinação.

REFERÊNCIAS

- ABABNEH, M. et al. Trends in aminoglycoside use and gentamicin-resistant Gram-negative clinical isolates in US academic medical centers: Implications for antimicrobial stewardship. **Infect. Control Hosp. Epidemiol.**, v. 33, n. 6, p. 594-601, 2012.
- ABDELBARY, M.H.H. et al. Whole-genome sequencing revealed independent emergence of vancomycin-resistant *Enterococcus faecium* causing sequential outbreaks over 3 years in a tertiary care hospital. **Euro J. Clin. Microbiol.**, v. 38, p. 1163-1170, 2019.
- ABELE-HORN, M. et al. Molecular epidemiology of hospital-acquired vancomycin-resistant enterococci. **J. Clin. Microbiol.**, v. 44, p. 4009-4013, 2006.
- ACHTMAN, M. et al. Microevolution and history of the plague bacillus, *Yersinia pestis*. **Proc. Natl. Acad. Sci. USA**, v. 101, p. 17837-17842, 2004.
- AGERSØ, Y.; PEDERSEN, A.G.; AARESTRUP, F.M. Identification of Tn5397-like and Tn916-like transposons and diversity of the tetracycline resistance gene *tet(M)* in enterococci from humans, pigs and poultry. **J. Antimicrob. Chemother.**, v. 57, n. 5, p. 832-839, 2006.
- AKPAKA, P.E. et al. Genetic characteristics and molecular epidemiology of vancomycin-resistant enterococci isolates from Caribbean countries. **PLoS ONE**, v. 12, n. 10, p. e0185920, 2017.
- ALCOCK, B.P. et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. **Nucleic Acids Res.**, v. 48, n. D1, p. D517-D525, 2020.
- ALEVIZAKOS, M. et al., Colonization with vancomycin-resistant enterococci and risk for bloodstream infection among patients with malignancy: A systematic review and meta-analysis. **Open Forum Infect. Dis.**, v. 4, n. 1, p. ofw246, 2017.
- ALMEIDA, B.M.M.; BRENDA, G.L.; SILVA, M.G. Proporção de bactérias multirresistentes de um hospital público sul-brasileiro. **Rev. Med. UFPR.**, v. 1, p. 5-9, 2014.
- ALTSCHUL, S.F. et al. Basic local alignment search tool. **J. Mol. Biol.**, v. 215, p. 403-410, 1990.
- ALVES, G.S.A. et al. Clonal dissemination of vancomycin-resistant *Enterococcus faecium* ST412 in a Brazilian region. **Braz. J. Infect. Dis.**, v. 21, p. 656 – 659, 2017.
- AMACHAWADI, R.G. et al. Occurrence of the transferable copper resistance gene *tcrB* among fecal enterococci of U.S. feedlot cattle fed copper-supplemented diets. **Appl. Environ. Microbiol.**, v. 79, n. 14, p. 4396-4375, 2013.
- AMIN, N.A.; JALAL, S.; WRETLIND, B. Alterations in GyrA and ParC associated with fluoroquinolone resistance in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 43, p. 947-949, 1999.

- ANDERSSON, D.I.; HUGHES, D. Persistence of antibiotic resistance in bacterial populations. **FEMS Microb. Rev.**, v. 35, n. 5, p. 901-911, 2011.
- ANDREW, J.P. et al. Roary: rapid large-scale prokaryote pan genome analysis. **Bioinformatics**, v. 31, n. 22, p. 3691-3693, 2015.
- ANDREW, S. **FastQC**. Versão 0.11.8. [Cambridge]: Brabraham Institute, 2018. Disponível em: <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>. Acesso em: 20 maio 2020.
- ARIAS, C.A. et al. Clinical and microbiological aspects of linezolid resistance mediated by the *cfr* gene encoding a 23S rRNA methyltransferase. **J. Clin. Microbiol.**, v. 46, p. 892-896, 2008.
- ARIAS, C. A.; MURRAY, B.E. Emergence and management of drug-resistant enterococcal infections. **Expert Rev. Anti-Infect. Ther.**, v. 6, p. 637-655, 2008.
- ARIAS, C. A.; MURRAY, B.E. The rise of the *Enterococcus*: Beyond vancomycin resistance. **Nat. Rev. Microbiol.**, v. 10, p. 266-278, 2012.
- ARNDT, D. et al. PHASTER: a better, faster version of the PHAST phage search tool. **Nucleic Acids Res.**, v. 44, n. W1, p. W16-W21, 2016.
- ARREDONDO-ALONSO, S. et al. Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. **mBIO**, v. 11, n. 1, p. e03284-19, 2020.
- ARTHUR, M. et al. Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptideoglycan precursors in *Enterococcus faecium* BM4147. **J. Bacteriol.**, v. 175, p. 117-127, 1993.
- ARTHUR, M.; QUINTILIANI, R., Jr. Regulation of VanA- and VanB-type glycopeptide resistance in enterococci. **Antimicrob. Agents Chemother.**, v. 45, p. 375-338, 2001.
- ASLAM, M.; DIARRA, M.S.; MASSON, L. Characterization of antimicrobial resistance and virulence genotypes of *Enterococcus faecalis* recovered from a pork processing plant. **J. Food Prot.**, v. 75, n. 8, p. 1486-1491, 2012.
- AZIZ, R.K.; NIZET, V. Pathogen microevolution in high resolution. **Sci. Transl. Med.**, v. 2, n. 16, p. 16ps4, 2010.
- BABADY, N.E. Hospital-associated infections. **Microbiol. Spectr.**, v. 4, n. 3, p. 10.1128/microbiolspec.DMIH2-0003-2015, 2016.
- BAKER-AUSTIN, C. et al. Co-selection of antibiotic and metal resistance. **Trends Microbiol.**, v. 14, p. 176-182, 2006.
- BALDASSARRI, L. et al. Variant *esp* gene in vancomycin-sensitive *Enterococcus faecium*. **Lancet**, v. 357, p. 1802, 2001.
- BAPTESTE, F.; BOUCHER, Y. Lateral gene transfer challenges principles of microbial systematic. **Trends Microbiol.**, v. 16, p. 200-207, 2008.

BAQUERO, F.; COQUE, T.M.; CANTON, R. Antibiotics, complexity, and evolution. **ASM News**, v. 69, p. 547-552, 2003.

BARROS, M.; MARTINELLI, R.; ROCHA, H. Enterococcal urinary tract infections in university hospital: clinical studies. **Braz. J. Infect. Dis.**, v. 13, n. 4, p. 294-296, 2009.

BAYJANOV, J.R. et al. *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. **Microb. Genom.**, v. 5, n. 7, p. e000277, 2019.

BEDDEK, A. et al. Evidence of capsule switching between carried and disease-causing *Neisseria meningitidis* strains. **Infect. Immun.**, v. 77, n. 7, p. 2989-2994, 2009.

BEHNOOD, A. et al. Prevalence of *aac(6')-Ie-aph(2'')-Ia* resistance gene and its linkage to Tn5281 in *Enterococcus faecalis* and *Enterococcus faecium* isolates from Tabriz hospitals. **Iran J. Microbiol.**, v. 5, n. 3, p. 203-208, 2013.

BENDER, J.K. et al. Detection of a *cfr(B)* variant in German *Enterococcus faecium* clinical isolates and the impact of linezolid resistance in *Enterococcus* spp. **PLoS ONE**, v. 11, n. 11, p. e0167042, 2016a.

BENDER, J.K. et al. Population structure and acquisition of the *vanB* resistance determinant in German clinical isolates of *Enterococcus faecium* ST192. **Sci. Rep.**, v. 6, n. 21847, Não paginado, 2016b.

BEN SAID, L. et al. Prevalence, antimicrobial resistance and genetic lineages of *Enterococcus* spp. from vegetable food, soil and irrigation water in farm environments in Tunisia. **J. Sci. Food Agric.**, v. 96, n. 5, p. 1627-1633, 2016.

BERGSTROM, C.T.; LISITCH, M.; LEVIN, B.R. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. **Genetics**, v. 155, n. 4, p. 1505-1519, 2000.

BERTELLONI, F. et al. Antimicrobial resistance in *Enterococcus* spp. isolated from laying hens of backyard poultry flocks. **Ann. Agric. Environ. Med.**, v. 22, n. 4, p. 665-669, 2015.

BETHEA, J.A.; WALKO, C.M.; TARGOS, P.A. Treatment of vancomycin-resistant *Enterococcus* with quinupristin/dalfopristin and high-dose ampicillin. **Ann. Pharmacother.**, v. 38, n. 6, p. 989-991, 2004.

BILLSTRÖM, H. et al. Frequent occurrence of multidrug-resistant CC17 *Enterococcus faecium* among clinical isolates in Sweden. **J. Appl. Microbiol.**, v. 108, p. 1810-1816, 2010.

BINNS, D. et al. QuickGO: a web-based tool for Gene Ontology searching. **Bioinformatics**, v. 25, n. 22, p. 3045-3046, 2009.

BIRMINGHAM, M.C. et al. Linezolid for the treatment of multidrug-resistant, gram-positive infections: Experience from a compassionate-use program. **Clin. Infect. Dis.**, v. 36, p. 159-168, 2003.

BOBAY, L.M.; OCHMAN, H. The evolution of bacterial genome architecture. **Front. Genet.**, v. 8, n. 72, 2017. DOI: <https://doi.org/10.3389/fgene.2017.00072>

- BONILLA, H.F.; ZERVOS, M.J.; KAUFFMAN, C.A. Long-term survival of vancomycin-resistant *Enterococcus faecium* on contaminated surface. **Infect. Control. Hosp. Epidemiol.**, v. 12, n.12, p. 770-772, 1996.
- BONORA, M.G. et al. Phylogenetic analysis of vancomycin-resistant *Enterococcus faecium* genotypes associated with outbreaks or sporadic infections in Italy. **Microb. Drug Resist.**, v. 13, n. 3, p. 171-177, 2007.
- BONORA, M.G. et al. Vancomycin-resistant *Enterococcus faecium* isolates causing hospital outbreaks in Northern Italy belong to the multilocus sequence typing C1 lineage. **Microb. Drug Resist.**, v. 10, n. 2, p. 114-123, 2004.
- BOTO, L. Horizontal gene transfer in evolution: facts and challenges. **Proc. R. Soc.** v. 277, p. 819-827, 2010.
- BOUCKAERT, R. et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. **PLoS Comput. Biol.**, v. 15, n. 4, p. e1006650, 2019.
- BOYD, E.F.; ALMAGRO-MORENO, S.; PARENT, M.A. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. **Trends Microbiol.**, v. 17, n. 2, p. 47-53, 2009.
- BRISSE, S. et al. Association of alterations in ParC and GyrA proteins with resistance of clinical isolates of *Enterococcus faecium* to nine different fluoroquinolones. **Antimicrob. Agents Chemother.**, v. 43, p. 2513-2516, 1999.
- BROCKHURST, M.A. et al. The ecology and evolution of pangenomes. **Curr. Biol.**, v. 29, n. 20, p. PR1094-PR1103, 2019.
- BRODRICK, H.J. et al. Whole-genome sequencing reveals transmission of vancomycin-resistant *Enterococcus faecium* in a healthcare network. **Genome Med.**, v. 8, n. 4, p. 4, 2016.
- BROUWER, M.S.; MULLANY, P.P; ROBERTS, A.P. Characterization of the conjugative transposon Tn6000 from *Enterococcus casseliflavus* 664.1H1. (formerly *Enterococcus faecium* 664.1H1). **FEMS Microbiol. Lett.**, v. 309, p. 71-76, 2010.
- BROWN, E. et al. Use of whole-genome sequencing for food safety and public health in the United States. **Foodborne Pathog. Dis.**, v. 16, n. 7, p. 441-450, 2019.
- BROWN, J. Ancient horizontal gene transfer. **Nat. Rev. Genet.**, n. 4, p. 121-132, 2003.
- BROWN, S.P. et al. Ecology of microbial invasions: Amplification allows virus carriers to invade more rapidly when rare. **Curr. Biol.**, v. 16, p. 2048-2052, 2006.
- BRYANT, J.; CHEWAPREECHA, C.; BENTLEY, S.D. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. **Future Microbiol.**, v. 7, n. 11, p. 1283-1296, 2012.

BRYNILDSRUD, O. et al. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. **Genome Biol.**, v. 17, p. 238, 2016.

BURLEY, K.; SEDGLEY, C.M. CRISPR-Cas, a prokaryotic adaptive immune system, in endodontic, oral, and multidrug-resistant hospital-acquired *Enterococcus faecalis*. **J. Endod.**, v. 38, n. 11, p. 1511-1515, 2012.

BUULTJENS, A.H. et al. Evolutionary origins of the emergent ST796 clone of vancomycin resistant *Enterococcus faecium*. **PeerJ**, v. 5, n. e2916, 2017. DOI: <https://doi.org/10.7717/peerj.2916>.eCollection 2017.

BYAPPANAHALLI, M.N. et al. Enterococci in the environment. **Microbiol. Mol. Biol. Rev.**, v. 76, n. 4, p. 685-706, 2012.

BYRNE, M.E.; ROUCH, D.A.; SKURRAY, R.A. Nucleotide sequence analysis of IS256 from the *Staphylococcus aureus* gentamicin-tobramycin resistance transposon Tn4001. **Gene**, v. 81, p. 361-367, 1989.

CAI, J. et al. Enterococcal isolates carrying the novel oxazolidinone resistance *optrA* from hospitals in Zhejiang, Guangdong, and Henan, China, 2010-2014. **Clin. Microb. Infect.**, v. 21, n. 12, p. 1095.e1-1095.e4, 2015.

CAMARGO, C.H. et al. Prevalence and phenotypic characterization of *Enterococcus* spp. isolated from food in Brazil. **Braz. J. Microbiol.**, v. 45, n. 1, p. 111-115, 2014.

CAMARGO, I.L.B.C. et al. Identification of an unusual VanA element in glycopeptide-resistant *Enterococcus faecium* in Brazil following international transfer of a bone marrow transplant patient. **Can. J. Microbiol.**, v. 50, n. 9, p. 767-770, 2004.

CAMARGO, I.L.B.C. et al. Occurrence of insertion sequences within the genomes and Tn1546-like elements of glycopeptide-resistant enterococci isolated in Brazil, and identification of a novel element, ISEfa5. **Int. J. Med. Microbiol.**, v. 294, p. 513-519, 2005.

CAMPOS, P.A. et al. A sustained endemic outbreak of vancomycin-resistant *Enterococcus faecium*: A 30-month surveillance study. **Scand. J. Infect. Dis.**, v. 46, p. 547-554, 2014.

CANTÓN, R. et al. A potential role for daptomycin in enterococcal infections: what is the evidence? **J. Antimicrob. Chemother.**, v. 65, p. 1126-1136, 2010.

CARATTOLI, A. et al. PlasmidFinder and pMLST: in silico detection and typing of plasmids. **Antimicrob. Agents Chemother.**, v. 58, n. 7, p. 3895-3903, 2014.

CARROL, A.C.; WONG, A. Plasmid persistence: costs, benefits, and the plasmid paradox. **Can. J. Microbiol.**, v. 64, n. 5, p. 293-304, 2018.

CARTER, G.P. et al. Emergence of endemic MLST non-typeable vancomycin-resistant *Enterococcus faecium*. **J. Antimicrob. Chemother.**, v. 71, n. 12, p. 3367-3371, 2016.

CATTOIR, V. et al. Trimethoprim resistance genes in vancomycin-resistant *Enterococcus faecium* clinical isolates from France. **Int. J. Antimicrob. Agents**, v. 34, n. 4, p. 390-392, 2009.

CENTERS FOR DISEASE CONTROL AND PREVENTION. Antibiotic Resistance Threats in the United States, 2019. Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2019. Disponível em: <<https://www.cdc.gov/DrugResistance/Biggest-Threats.html>>. Acesso em: 20 maio 2020

CHAGUZA, C. et al. Recombination in *Streptococcus pneumoniae* lineages increase with carriage duration and size of the polysaccharide capsule. **mBio**, v. 7, n. 5, p. e01053-16, 2016.

CHAJĘCKA-WIERZCHOWSKA, W.; ZADERNOWSKA, A.; ŁANIEWSKA-TROKENHEIM, L. Virulence factors, antimicrobial resistance and biofilm formation in *Enterococcus* spp. isolated from retail shrimps. **LWT – Food Sci. Technol.**, v. 69, p. 117-122, 2016.

CHANG, C.M. et al. Characterization of vancomycin-resistant enterococci from hospitalized patients at a tertiary center over a seven-year period. **J. Hosp. Infect.**, v. 74, p. 377- 384, 2010.

CHIVERS, L.S. et al. Vancomycin-resistant enterococci: 15 years and counting. **J. Hosp. Infect.**, v. 53, n. 3, p. 159-171, 2003.

CHEN, C.H. et al. Clinical and microbiological characteristics of vancomycin-resistant *Enterococcus faecium* bloodstream infection in Central Taiwan. **Medicine (Baltimore)**, v. 96, n. 49, p. e900, 2017.

CHEN, H. et al. Molecular characteristics of oxazolidinone resistance in enterococci from a multicenter study in China. **BMC Microbiol.**, v. 19, n. 1, p. 162, 2019.

CHOW, J.W. Aminoglycoside resistance in enterococci. **Clin. Infect. Dis.**, v. 31, p. 586 - 589, 2000.

CINGOLANI, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. **Fly (Austin)**, v. 6, n. 2, p. 80-92, 2012a.

CINGOLANI, P. et al. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. **Front. Genet.**, v. 3, p. 35, 2012b.

CLEWELL, D.B. et al. Extrachromosomal and mobile elements in enterococci: transmission, maintenance, and epidemiology. In: GILMORE, M.S. et al. (ed). **Enterococci: From commensals to leading causes of drug resistant infection**. Boston, MA: Massachusetts Eye and Ear Infirmary, 2014.

CLINICAL LABORATORY STANDARDS INSTITUTE. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically: approved standard. 10th ed. Wayne, PA: CLSI, 2015. (Clinical Laboratory Standards Institute Document M07-A10, v. 35, n. 2).

CLINICAL LABORATORY STANDARDS INSTITUTE. Performance Standards for Antimicrobial Susceptibility Testing. 29th ed. Wayne, PA:CLSI, 2019. (Clinical Laboratory Standards Institute Document M100)

CONCEICAO, N. et al. Trends in antimicrobial resistance among clinical isolates of enterococci in a Brazilian tertiary hospital: a 4-year study. **Rev. Soc. Bras. Med. Trop.**, v. 44, n. 2, p. 177-181, 2011.

COOMBS, G.W. et al. Australian group on antimicrobial resistance (AGAR) Australian Enterococcal sepsis outcome programme (AESOP) Annual report 2018. **Commun. Dis. Intell.**, v. 44, 2020. DOI: <https://doi.org/10.33321/cdi.2020.44.19>.

COOMBS, G.W. et al. Molecular epidemiology of enterococcal bacteremia in Australia. **J. Clin. Microbiol.**, v. 52, n. 3, p. 897-905, 2014.

COQUE, T.M. et al. Characterization of dihydrofolate reductase genes from trimethoprim-susceptible and trimethoprim-resistant strains of *Enterococcus faecalis*. **Antimicrob. Agents Chemother.**, v. 43, n. 1, p. 141-147, 1999.

COQUE, T.M. et al. High occurrence of *esp* among ampicillin-resistant and vancomycin-susceptible *Enterococcus faecium* clones from hospitalized patients. **J. Antimicrob. Chemother.**, v. 50, p. 1035-1038, 2002.

COQUE, T.M. et al. Population structure of *Enterococcus faecium* causing bacteremia in a Spanish University Hospital: Setting the scene for a future increase in vancomycin resistance? **Antimicrob. Agents Chemother.**, v. 49, n. 7, p. 2693-2700, 2005.

CORANDER, J. et al. Population structure in the *Neisseria*, and the biological significance of fuzzy species. **J. R. Soc. Interface**, v. 9, p. 1208-1215, 2012.

CORDERO, O.X.; POLZ, M.F. Explaining microbial genomic diversity in light of evolutionary ecology. **Nat. Rev. Microbiol.**, v. 12, p. 263-273, 2014.

CORREIA, S. et al. Mechanisms of quinolone action and resistance, where do we stand? **J. Med. Microbiol.**, v. 66, p. 551-559, 2017.

COURVALIN, P. Vancomycin resistance in gram-positive cocci. **Clin. Infect. Dis.**, v. 42, p. S25-S34, 2006.

COURVIN, D. et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. **Nucleic Acids Res.**, v. 46, n. W1, p. W246-W251, 2018.

COX, C.R.; GILMORE, M.S. Native microbial colonization of *Drosophila melanogaster* and its use as a model of *Enterococcus faecalis* pathogenesis. **Infect. Immun.**, v. 75, n. 4, p. 1565-1576, 2007.

CRETI, R. et al. Survey for virulence determinants among *Enterococcus faecalis* isolated from different sources. **J. Med. Microbiol.**, v. 53, n. 1, p. 13-20, 2004.

CRISPELL, J.; BALAZ, D.; GORDON, S.V. HomoplasmyFinder: a simple tool to identify homoplasies on phylogeny. **Microb. Genom.**, v. 5, n. 1, p. e000245, 2019.

DAMANI, A. et al. A newly described vancomycin-resistant ST412 *Enterococcus faecium* predominant in Greek hospitals. **Eur. J. Clin. Microbiol.**, v. 29, n. 3, p. 329-331, 2010.

DARINI, A.L.C. et al. Disruption of *vanS* by *IS1216V* in a clinical isolate of *Enterococcus faecium* with VanA glycopeptide resistance. **Antimicrob. Agents Chemother.**, v. 43, n. 4, p. 995-996, 1999.

DARMON, E.; LEACH, D.R.F. Bacterial genome instability. **Microbiol. Mol. Biol. Rev.**, v. 78, n. 1, p. 1-39, 2014.

DAVIS, I.J. et al. Linkage of a novel mercury resistance operon with streptomycin resistance on a conjugative plasmid in *Enterococcus faecium*. **Plasmid**, v. 54, p. 26-38, 2005.

D' COSTA, V.M. et al. Antibiotic resistance is ancient. *Nature*, v. 477, p. 457-461.

DE BEEN, M. et al. Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. **Genome Biol. Evol.**, v. 5, n. 8, p. 1524-1535, 2013.

DE GRAEF, E.M. et al. Prevalence and mechanism of resistance against macrolides, lincosamides, and streptogramins among *Enterococcus faecium* isolates from food-producing animals and hospital patients in Belgium. **Microb. Drug Resist.**, v. 13, n. 2, p. 135-141, 2007.

DEN BAKKER, H.C. et al. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. *BMC Evol. Biol.*, v. 8, n. 277, 2008. DOI: <https://doi.org/10.1186/1471-2148-8-277>.

DERBISE, A.; DE, C.G.; EL, SN. Nucleotide sequence of the *Staphylococcus aureus* transposon, Tn5405, carrying aminoglycosides resistance genes. **J. Basic Microbiol.**, v. 37, n. 5, p. 379-384, 1997.

DESHPANDE, L.M. et al. Antimicrobial resistance and molecular epidemiology of vancomycin-resistant enterococci from North America and Europe: a report from the SENTRY antimicrobial surveillance program. **Diagn. Micr. Infect. Dis.**, v. 58, p. 163-170, 2007.

DESHPANDE, L.M. et al. Detection of a new *cfr*-like gene, *cfr(B)*, in *Enterococcus faecium* isolates recovered from human specimens in the United States as part of the SENTRY antimicrobial surveillance program. **Antimicrob. Agents Chemother.**, v. 59, p. 10, n. 6256-6261, 2015.

DESHPANDE, L.M. et al. Evolving oxazolidinone resistance mechanisms in a worldwide collection of enterococcal clinical isolates: results from the SENTRY antimicrobial surveillance program. **J. Antimicrob. Chemother.**, v. 73, n. 9, p. 2314-2322, 2018.

DICUONZO, G. et al. Antibiotic resistance and genotypic characterization by PFGE of clinical and environmental isolates of enterococci. **FEMS Microbiol. Lett.**, v. 201, n. 2, p. 205-211, 2001.

DIDELLOT, X. et al. Transforming clinical microbiology with bacterial genome sequencing. **Nat. Rev. Genet.**, v. 13, n. 9, p. 601-612, 2012.

DIDELLOT, X.; MAIDEN, M.C.J. Impact of recombination on bacterial evolution. **Trends Microbiol.**, v. 18, n. 7, p. 315-322, 2010.

DIDELLOT, X.; WILSON, D.J. ClonalFrameML: Efficient inference of recombination in whole bacterial genomes. **PLoS Comput. Biol.**, v. 11, n. 2, p. e1004041, 2015.

DINA, J.; MALBRUNY, B.; LECLERCQ R. Nonsense mutations in the *lsa-Like* gene in *Enterococcus faecalis* isolates susceptible to lincosamides and streptogramins A. **Antimicrob. Agents Chemother.**, v. 47, n. 7, p. 2307-2309, 2003.

DONABEDIAN, S.M. et al. Quinupristin-Dalfopristin resistance in *Enterococcus faecium* isolates from humans, farm animals, and grocery store meat in the United States. **J. Clin. Microb.**, v. 44, p. 3361-3365, 2006.

DONSKEY, C.J. et al. Effect of antibiotic therapy on the density of vancomycin-resistant enterococci in the stool of colonized patients. **New Eng. J. Med.**, v. 343, n. 26, p. 1925-1932, 2000.

DUERKOP, B; PALMER, K.L.; HORSBURGH, M. Enterococcal bacteriophages and genome defense. In: GILMORE, M.S. et al. (ed). **Enterococci: From commensals to leading causes of drug resistant infection**. Boston: Massachusetts Eye and Ear Infirmary, 2014.

DUTKA-MALEN, S.; EVERS, S.; COURVALIN, P. Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant enterococci by PCR. **J. Clin. Microbiol.**, v. 33, p. 24 – 27, 1995.

EDELSBERG, J. et al. Prevalence of antibiotic resistance in US hospitals. **Diagn. Microbiol. Infect. Dis.**, v. 78, p. 255-262, 2014.

EHRMANN, M.; VOGEL, R.F. Characterization of IS153, an IS3-family insertion sequence isolated from *Lactobacillus sanfranciscensis* and its use for strain differentiation. **System. Appl. Microbiol.**, v. 24, p. 443-450, 2001.

EICHEL, V. et al. Challenges in interpretation of WGS and epidemiological data to investigate nosocomial transmission of vancomycin-resistant *Enterococcus faecium* in an endemic region: incorporation of patient movement network and admission screening. **J. Antimicrob. Chemother.**, v. 75, n. 7, p. 1716-1721, 2020.

EISENBERGER, D. et al. Whole-genome analysis of vancomycin-resistant *Enterococcus faecium* causing nosocomial outbreaks suggests the occurrence of few endemic clonal lineages in Bavaria, Germany. **J. Antimicrob. Chemother.**, v. 75, n. 6, p. 1398-1404, 2020.

- EL HADDAD, L. et al. Single nucleotide polymorphism (SNP) analysis reveal potential vancomycin-resistant enterococci (VRE) transmission networks between rooms and patients on stem cell transplant (SCT) and leukemia units. **Biol. Blood Marrow Transplant**, v. 24, p. S25-S118, 2018.
- ELIOPOULOS, G.M. Vancomycin-resistant enterococci: Mechanism and clinical relevance. **Infect. Dis. Clin. North Am.**, v. 11, n. 4, p. 851-865, 1997.
- ELLINGTON, M.J. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST subcommittee. **Clin. Microbiol. Infect.**, v 23, p. 2-22, 2017.
- EL-MAHDY, R.; MOSTAFA, A.; EL-KANNISHY, G. High level aminoglycoside resistant enterococci in hospital-acquired urinary tract infections in Mansoura, Egypt. **Germs**, v. 8, n. 4, p. 186-190, 2018.
- ERACLIO, G.; RICCI, G.; FORTINA, M.G. Insertion sequence elements in *Lactococcus garvieae*. **Gene**, v. 55, n. 2, p. 291-296, 2015.
- EUROPEAN COMMITTEE ON ANTIMICROBIAL SUSCEPTIBILITY TESTING. Breaking point tables for interpretation of MICs and zone diameters. Copenhagen: European Society of Clinical Microbiology and Infectious Disease, 2019. v. 9.0.
- FABRETTI, F. et al. Alanine esters of enterococcal lipoteichoic acid play a role in biofilm formation and resistance to antimicrobial peptides. **Infect. Immun.**, v. 74, p. 4164-4171, 2006.
- FACKLAM, R.R.; CARVALHO, M.G.S.; TEIXERA, L.M. History, taxonomy, biochemical characteristics, and antibiotic susceptibility testing of enterococci. In: GILMORE, M.S. et al (ed). **The Enterococci: pathogenesis, molecular biology, and antibiotic resistance**. Washington, DC: ASM Press, 2002. p. 1-54.
- FALGENHAUER, L. et al. Near-ubiquitous presence of a vancomycin-resistant *Enterococcus faecium* ST117/CT71/*vanB* – clone in the Rhine-Main metropolitan area of Germany. **Antimicrob. Resist. Infect. Control.**, v. 8, n. 128, 2019. DOI: <https://doi.org/10.1186/s13756-019-0573-8>. eCollection 2019.
- FARD, R.M.; BARTON, M.D.; HEUZENROEDER, MW. Bacteriophage-mediated transduction of antibiotic resistance in enterococci. **Lett. Appl. Microbiol.**, v. 52, n. 6, p. 559-564, 2011.
- FARD, R.M.; HEUZENROEDER, MW.; BARTON, M.D. Antimicrobial and heavy metal resistance in commensal enterococci isolated from pigs. **Vet. Microbiol.**, v. 148, p. 276-282, 2011.
- FARHAT, M.R. et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. **Nat. Genet.**, v. 45, p. 1183-1189, 2013.
- FARIA, A.R. **Diversidade e estrutura populacional de *Enterococcus faecalis* resistentes a níveis elevados de aminoglicosídeos**. 2017. 180f. Tese (Doutorado em Microbiologia) -

Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, 2017.

FARIA, A.R. *Enterococcus faecium*: caracterização molecular de amostras isoladas no período anterior e posterior à emergência da resistência à vancomicina no Estado do Rio de Janeiro. 2012. 110f. Dissertação (Mestrado em Microbiologia) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, 2012.

FEAVERS, I.M. et al. Role of horizontal genetic exchange in the antigenic variation of the class 1 outer membrane protein of *Neisseria meningitidis*. **Mol. Microbiol.**, v. 6, 489-495, 1992.

FERNÁNDEZ-GARCIA, L. Toxin-Antitoxin systems in clinical pathogens. **Toxins**, v. 8, n. 7, p. 227-250, 2016.

FERNÁNDEZ-HIDALGO, N.; ESCOLÀ-VERGÉ, L.; PERICÀS, J.M. *Enterococcus faecalis* endocarditis: what's next? **Future Microbiology**, v. 15, n. 5, p. 349-364, 2020.

FRANYÓ, D. et al. Characterization of clinical vancomycin-resistant *Enterococcus faecium* isolated in Eastern Hungary. **Microb. Drug Resist.**, v. 24, n. 10, p. 1559-1567, 2018.

FRASER, C. et al. The bacterial species challenge: making sense of genetic and ecological diversity. **Science**, v. 323, p. 741-746, 2009.

FRAZZON, A.P.G. et al. Prevalence of antimicrobial resistance and molecular characterization of tetracycline resistance mediated by *tet(M)* and *tet(L)* genes in *Enterococcus* spp. isolated from food in Southern Brazil. **World J. Microbiol. Biotechnol.**, v. 26, p. 365-370, 2010.

FREITAS, A.A. et al. Distribution of species and antimicrobial resistance among enterococci isolated from the fecal microbiota of captive blue-fronted parrot (*Amazona aestiva*) in Rio de Janeiro, Brazil. **Sci. Total Environ.**, v. 15, p. 1428-1437, 2018.

FREITAS, A.R. et al. Distribution of putative virulence markers in *Enterococcus faecium*: towards a safety profile review. **J. Antimicrob. Chemother.**, v. 73, n. 2, p. 306-319, 2018a.

FREITAS, A.R. et al. Global spread of the *hyl_{Efm}* colonization-virulence gene in megaplasmids of the *Enterococcus faecium* CC17 polyclonal subcluster. **Antimicrob. Agents Chemother.**, v. 54, n. 6, p. 2660-2665, 2010.

FREITAS, A.R. et al. High rates of colonisation by ampicillin-resistant enterococci in residents of long-term care facilities in Porto, Portugal. **Int. J. Antimicrob. Agents.**, v. 51, p. 503-507, 2018b.

FREITAS, A.R. et al. Microevolutionary events involving narrow host plasmids influence local fixation of vancomycin-resistance in *Enterococcus* populations. **PLoS ONE**, v. 8, n. 3, p. e60589, 2013.

FREITAS, A.R. et al. Multilevel population genetic analysis of *vanA* and *vanB* *Enterococcus faecium* causing nosocomial outbreaks in 27 countries (1986-2012). **J. Antimicrob. Chemother.**, v. 71, n. 12, p. 3351-3366, 2016.

- FRIDKIN, S.K. et al. Epidemiology of a dominant clonal strain of vancomycin-resistant *Enterococcus faecium* at separate hospitals in Boston Massachusetts. **J. Clin. Microbiol.**, v. 36, p. 965-970, 1998.
- FUKUI, K. DNA mismatch repair in eukaryotes and bacteria. **J. Nucleic. Acids**, v. 2010, n. 2605512, 2010. DOI: 10.4061/2010/260512.
- GALHARDO, R.S.; HASTINGS, P.J.; ROSENBERG, S.M. Mutations as a stress response and the regulation of evolvability. **Crit. Rev. Biochem. Mol. Biol.**, v. 42, p. 399-435, 2007.
- GALLOWAY-PEÑA, J. et al. Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of *Enterococcus faecium*. **PLoS ONE**, v. 7, n. 1, p. e30187, 2012.
- GALLOWAY-PEÑA, J.; RICE, L.B.; MURRAY, B.E. Analysis of PBP5 of early U.S. isolates of *Enterococcus faecium*: Sequence variation alone does not explain increasing ampicillin resistance over time. **Antimicrob. Agents Chemother.**, v. 55, n. 7, p. 3272-3277, 2011.
- GALTIER, N. A model of horizontal gene transfer and the bacterial phylogeny problem. **Syst. Biol.**, v. 56, n. 4, p. 633-642, 2007.
- GAO, W.; HOWDEN, B.P.; STINEAR, T.P. Evolution of virulence in *Enterococcus faecium*, a hospital-adapted opportunistic pathogen. **Curr. Opin. Microb.**, v. 41, p. 76-82, 2018.
- GARCIA-SOLACHE, M. et al. Homologous recombination within large chromosomal regions facilitates acquisition of β -Lactam and vancomycin resistance in *Enterococcus faecium*. **Antimicrob. Agents. Chemother.**, v. 60, n. 10, p. 5777-5786, 2016.
- GEISS-LIEBISCH, S. et al. Secondary cell wall polymers of *Enterococcus faecalis* are critical for resistance to complement activation via mannose-binding lectin. **J. Biol. Chem.**, v. 287, p. 37769-37777, 2012.
- GETACHEW, Y. et al. Genetic variability of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* isolates from humans, chickens, and pigs in Malaysia. **Appl. Environ. Microbiol.**, v. 79, n. 15, p. 4528-4533, 2013.
- GEVERS, D. Opinion: Re-evaluating prokaryotic species. **Nat. Rev. Microbiol.**, v. 3, p. 733-739, 2005.
- GHATTARGI, V.C. et al. Comparative genome analysis reveals key genetic factors associated with probiotic property in *Enterococcus faecium* strains. **BMC genomics**, v. 19, n. 652, 2018. DOI: <https://doi.org/10.1186/s12864-018-5043-9>.
- GILMORE, M.S.; LEBRETON, F.; VAN SCHAIK, W. Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. **Curr. Opin. Microbiol.**, v. 16, n. 1, p. 10-16, 2013.

- GOULIOURIS, T. et al. Detection of vancomycin-resistant *Enterococcus faecium* hospital-adapted lineages in municipal wastewater treatment plants indicates widespread distribution and release into the environment. **Genome Res.**, v. 29, p. 626-634, 2019.
- GRAYSON, M.L. et al. Failure of trimethoprim-sulfamethoxazole therapy in experimental enterococcal endocarditis. **Antimicrob. Agents Chemother.**, v. 34, n. 9, p. 1792-1794, 1990.
- GRIFFIN, P.M. et al. Use of matrix-assisted laser desorption ionization-time of flight mass spectrometry to identify vancomycin-resistant enterococci and investigate the epidemiology of an outbreak. **J. Clin. Microbiol.**, v. 50, p. 2918-2931, 2012.
- GUDIOL, C. et al. Increase in bloodstream infection due to vancomycin-susceptible *Enterococcus faecium* in cancer patients: Risk factors, molecular epidemiology and outcomes. **PLoS ONE**, v. 8, n. 9, p. e74734, 2013.
- GUREVICH, A. et al. QUASt: quality assessment tool for genome assemblies. **Bioinformatics**, v. 28, n. 8, p. 1072-1075, 2013.
- GUZMÁN PRIETO, A.M. et al. Global emergence and dissemination of enterococci as nosocomial pathogens: Attach of the clones? **Front. Microbiol.**, v. 7, n. 788, 2016. DOI: 10.3389/fmicb.2016.00788.
- GUZMÁN PRIETO, A.M. et al. The two-component system *ChtRS* contributes to chlorhexidine tolerance in *Enterococcus faecium*. **Antimicrob. Agents. Chemother.**, v. 61, n.5, p. e02122-16, 2017.
- HAANDRIKMAN, A.J. et al. Insertion elements on lactococcal proteinase plasmids. **Appl. Environ. Microb.**, v. 56, n. 6, p. 1890-1896, 1990.
- HAGHI, F.; LOHRASBI, V.; ZEIGHAMI, H. High incidence of virulence determinants, aminoglycoside and vancomycin resistance in enterococci isolated from hospitalized patients in Northwest Iran. **BMC Infect. Dis.**, v. 19, n. 744, 2019. DOI: 10.1186/s12879-019-4395-3.
- HANCOCK, L.E.; GILMORE, M.S. The capsular polysaccharide of *Enterococcus faecalis* and its relationship to other polysaccharides in the cell wall. **Proc. Natl. Acad. Sci. USA**, v. 99, p. 1574-1579, 2002.
- HANDWERGER, S.; SKOBLE, J. Identification of chromosomal mobile element conferring high-level vancomycin resistance in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 39, n. 11, p. 2446-2453, 1995.
- HANRAHAN, J.; HOYEN, C.; RICE, L.B. Geographic distribution of a large mobile element that transfer ampicillin and vancomycin resistance between *Enterococcus faecium* strains. **Antimicrob. Agents Chemother.**, v. 44, p. 1349-1351, 2000.
- HAO, W.; PALMER, J. HGT turbulence. **Mob. Genet. Elements**, v. 1, n. 4, p. 256-304, 2011.
- HARRISON, E.; BROCKHURST, M.A. Plamid-mediated horizontal gene transfer is a coevolutionary process. **Trends Microbiol.**, v. 20, n. 6, p. 262-267, 2012.

HARRISON, E. et al. Parallel compensatory evolution stabilizes plasmids across the parasitism-mutualism continuum. **Curr. Biol.**, v. 25, n. 15, p. 2034, 2039, 2015.

HASMAN, H.; AARESTRUP, F.M. Relationship between copper, glycopeptide, and macrolide resistance among *Enterococcus faecium* strains isolated from pigs in Denmark between 1997 and 2003. **Antimicrob. Agents Chemother.**, v. 49, p. 454-456, 2005.

HASMAN, H.; AARESTRUP, F.M. *tcxB*, a gene conferring transferable copper resistance in *Enterococcus faecium*: occurrence, transferability, and linkage to macrolide and glycopeptide resistance. **Antimicrob. Agents Chemother.**, v. 46, p. 1410-1416, 2002.

HE, S. et al. Mechanisms of evolution in high-consequence drug resistance plasmids. **mBio**, v. 7, n. 6, p. e01987-16, 2016.

HEGSTAD, K. et al. Mobile genetic elements and their contribution to the emergence of antimicrobial resistant *Enterococcus faecalis* and *Enterococcus faecium*. **Clin Microbiol. Infect.**, v. 16, n. 6, p. 541-554, 2010.

HEIKENS, E.; BONTEN, M.J.; WILLEMS, R.J. Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. **J. Bacteriol.**, v. 189, p. 8233- 8240, 2007.

HEIKENS, E. et al. Contribution of the Enterococcal surface protein Esp to pathogenesis of *Enterococcus faecium* endocarditis. **Microbes Infect.**, v. 13, p. 1185-1190, 2011.

HENDRICKX, A.P. et al. Five genes encoding surface-exposed LPXTG proteins are enriched in hospital-adapted *Enterococcus faecium* clonal complex 17 isolates. **J. Bacteriol.**, v. 189, p. 8321- 8332, 2007.

HENDRICKX, A.P. et al. SgrA, a nidogen-binding LPXTG surface adhesin implicated in biofilm formation, and EcbA, a collagen binding MSCRAMM, are two novel adhesins of hospital-acquired *Enterococcus faecium*. **Infect. Immun.**, v. 77, n. 11, p. 5097-5106, 2009.

HENDRIKSEN, R.S. et al. Using genomics to track global antimicrobial resistance. **Front. Public Health**, v. 7, n. 242, 2019. DOI: 10.3389/fpubh.2019.00242.

HERZSTEIN, J. et al. Optimal therapy for enterococcal endocarditis. **Am. J. Med.**, v. 76, p. 186–191, 1984.

HIDRON, A.I. et al. NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007. **Infect. Control. Hosp. Epidemiol.**, v. 29, n. 11, p. 996 –1011, 2008.

HOFFMANN, M. et al. Complete genome sequence of a vancomycin-resistant sequence type 203 *Enterococcus faecium* strain with *vanA* belonging to complex type 859. **Microbial Resour. Announc.**, v. 7, n. 2, p. e00815-00818, 2018.

- HOLLENBECK, B.L.; RICE, L.B. Intrinsic and acquired resistance mechanisms in *Enterococcus*. **Virulence**, v. 3, p. 421-569, 2012.
- HOMAN, W.L. et al. Multilocus sequence typing scheme for *Enterococcus faecium*. **J. Clin. Microbiol.**, v. 40, p. 1963-1971, 2002.
- HOWDEN, B.P. et al. Genomic insights to control the emergence of vancomycin- resistant enterococci. **MBio**, v.4, n.4, p. e00412-13, 2013.
- HSIEH, S.E. et al. Importance of amino acid alterations and expression of penicillin-binding protein 5 to ampicillin resistance of *Enterococcus faecium* in Taiwan. **Int. J. Antimicrob. Agents**, v. 28, p. 514-519, 2006.
- HSIEH, Y.C. et al. Clonal spread of CC17 vancomycin-resistant *Enterococcus faecium* with multilocus sequence type 78 (ST78) and a novel ST444 in Taiwan. **Eur. J. Microbiol.**, v. 29, p. 25–30, 2010.
- HUERTA-CEPAS, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. **Nucleic Acids Res.**, v. 47, n. D1, p. D309-D314, 2019.
- HUGHES, A.R. et al. Ecological consequences of genetic diversity. **Ecol. Lett.**, v. 11, p. 609-623, 2008.
- HUO, W. et al. A type I restriction-modification system associated with *Enterococcus faecium* subspecies separation. **Appl. Environ. Microb.**, v.85, n. 2, p. e02174-18, 2019.
- ISNARD, C. et al. Genetic basis for in vitro and in vivo resistance to lincosamides, streptogramins A, and pleuromutilins (LSAP phenotype) in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 57, n. 9, p. 4463-4469, 2013.
- JABBARI SHIADEH, S.M. et al. Global prevalence of antibiotic resistance in blood-isolated *Enterococcus faecalis* and *Enterococcus faecium*: a systematic review and meta-analysis. **Infect. Drug Resist.**, v. 12, p. 2713-2725, 2019.
- JACKSON, C.R. et al. Genetic relatedness of high-level aminoglycoside-resistant enterococci isolated from poultry carcasses. **Avian Dis.**, v. 48, p. 100-107, 2004.
- JACKSON, C.R.; FEDORKA-CRAY, P.J.; BARRETT, J. B. Use of a genus- and species-specific multiplex PCR for identification of enterococci. **J. Clin. Microbiol.**, v. 42, p. 3558-3565, 2004.
- JACOB, J. Characterization of the *sat4* gene encoding a streptothricin acetyltransferase in *Campylobacter coli* BE/G4. **FEMS Microbiol. Lett.**, v. 120, p. 13-18, 1994.
- JASPAN, H.B. et al. Multidrug-resistant *Enterococcus faecium* meningitis in a toddler: characterization of the organism and successful treatment with intraventricular daptomycin and intravenous tigecycline. **Pediatr. Infect. Dis. J.**, v. 29, n. 4, p. 379-381, 2010.

- JOENSEN, K.G. et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. **J. Clin. Microbiol.**, v. 52, n. 5, p. 1501-1510, 2014.
- JOHNSEN, P.J. et al. Persistence of animal and human glycopeptide-resistant enterococci on two Norwegian poultry farms formerly exposed to avoparcin is associated with a widespread plasmid-mediated *vanA* element within a polyclonal *Enterococcus faecium* population. **Appl. Environ. Microbiol.**, v. 71, p. 159-168, 2005.
- JOHNSON, P.D.R. et al. A sustained hospital outbreak of vancomycin-resistant *Enterococcus faecium* bacteremia due to emergence of *vanB E. faecium* sequence type 203. **J. Infect. Dis.**, v. 202, p. 1278-1286, 2010.
- JOHNSTON, L.M.; JAYKUS, L. Antimicrobial resistance in *Enterococcus* species isolated from produce. **Appl. Environ. Microbiol.**, v. 70, n. 5, p. 3133-3137, 2004.
- JOLLEY, K.A.; BRAY, J.E.; MAIDEN, M.C. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. **Wellcome Open Res.**, v. 3, n. 124, doi:10.12688/wellcomeopenres.14826.1, 2018.
- JOLLEY, K.A.; MAIDEN, M.C. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*. v. 11, n. 595, 2010. DOI: <https://doi.org/10.1186/1471-2105-11-595>
- JUNG, A. et al. A review of *Enterococcus cecorum* infection in poultry. **Avian Dis.**, v. 62, n. 3, p. 261-271, 2018.
- JUREEN, R. et al. Comparative analysis of amplified fragment length polymorphism and pulsed-field gel electrophoresis in a hospital outbreak and subsequent endemicity of ampicillin-resistant *Enterococcus faecium*. **FEMS Immunol. Med. Microbiol.**, v. 40, p. 33-39, 2004.
- KANG, Z.Z. et al. Detection of transferable oxazolidinone resistance determinants in *Enterococcus faecalis* and *Enterococcus faecium* of swine origin in Sichuan province, China. **J. Glob. Antimicrob. Resist.**, v. 19, p. 333-337, 2019.
- KELLEY, L.A. The Phyre2 web portal for protein modeling, prediction and analysis. **Nat. Protoc.**, v. 10, p. 845-858, 2015.
- KERBAUY, G. et al. Vancomycin-dependent *Enterococcus faecium vanA*: characterization of the first case isolated in a university hospital in Brazil. **Braz. J. Med. Bio. Res.**, v. 44, n. 3, p. 253-257, 2011.
- KERSCHNER, H. et al. Hospital outbreak caused by linezolid resistant *Enterococcus faecium* in Upper Austria. **Antimicrob. Resist. Infect. Control**, v. 8, n. 150, 2019. DOI: <https://doi.org/10.1186/s13756-019-0598-z>.
- KHAN, M.A. et al. High prevalence of ST-78 infection associated vancomycin-resistant *Enterococcus faecium* from hospitals in Asunción, Paraguay. **Clin. Microbiol. Infect.**, v. 16, p. 624-627, 2010.

- KIM, E.B.; MARCO, M.L. Nonclinical and clinical *Enterococcus faecium*, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. **Appl. Environ. Microbiol.**, v. 80, n. 1, p. 154-165, 2014.
- KIM, M.C.; WOO, G.J. Characterization of antimicrobial resistance and quinolone resistance factors in high-level ciprofloxacin-resistant *Enterococcus faecalis* and *Enterococcus faecium* isolates obtained from fresh produce and fecal samples of patients. **J. Sci. Food Agric.**, v. 97, p. 2858-2864, 2017.
- KIM, Y.B. et al., Characteristics of high-level ciprofloxacin-resistant *Enterococcus faecalis* and *Enterococcus faecium* from retail chicken meat in Korea. **J. Food Prot.**, v. 81, p. 1357-1363, 2018.
- KIMIRAN-ERDEN, A. et al. Isolation and identification of enterococci from seawater samples: Assessment of their resistance to antibiotics and heavy metals. **Environ. Monit. Assess.**, v. 125, p. 219-228, 2007.
- KIMURA, M. Evolutionary rate at the molecular level. **Nature**, v. 217, p. 624-626, 1968.
- KLARE, I. et al. Occurrence and spread of antibiotic resistances in *Enterococcus faecium*. **Int. J. Food Microbiol.**, v. 88, p. 269– 290, 2003.
- KLARE, I. et al. Spread of ampicillin/vancomycin-resistant *Enterococcus faecium* of the epidemic-virulent clonal complex-17 carrying the genes *esp* and *hyl* in German hospitals. **Eur. J. Clin. Microbiol.**, v. 24, p. 815-825, 2005.
- KLEIN, E.Y. et al. Assessment of WHO antibiotic consumption and access targets in 76 countries, 2000-2015: an analysis of pharmaceutical sales data. **Lancet Infect. Dis.**, 2020. In Press.
- KLIBI, N. et al. Diversity of structure carrying the *aac(6')-aph(2'')* gene in clinical *Enterococcus faecalis* and *Enterococcus faecium* strains isolated in Tunisia. **J. Chemother.**, v. 18, n. 4, p. 353-359, 2006.
- KLIBI, N. et al. Polymorphism in *pbp5* gene detected in clinical *Enterococcus faecium* strains with different ampicillin MICs from a Tunisian hospital. **J. Chemother.**, v. 20, n. 4, p. 436-440, 2008.
- KONSTANTINIDIS, K.T.; TIEDJE, J.M. Genomic insights that advance the species definition for prokaryotes. **Proc. Natl. Acad. Sci. USA**, v. 102, p. 2567-2572, 2005.
- KÖSER, C.U. et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. **PLoS Pathog.**, v. 8, n. 8, p. e1002824, 2012.
- KÖSER, C.U.; ELLINGTON, M.J.; PEACOCK, S.J. Whole-genome sequencing to control antimicrobial resistance. **Trends Genet.**, v. 30, n. 9, p. 401-407, 2014.
- KOTEVA, K. et al. A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. **Nat. Chem. Biol.**, v. 6, p. 327-329, 2010.

KROPEC, A. et al. Identification of SagA as a novel vaccine target for the prevention of *Enterococcus faecium* infections. **Microbiology**, v. 157, p. 3429-3434, 2011.

KRUEGER, F. **Trim Galore!**. Versão 0.6.2. [Cambridge]: Brabraham Institute, 2015. Disponível em: <http://www.bioinformatics.brabraham.ac.uk/projects/trim_galore/>. Acesso em: 20 maio 2020

KUHN, L.G. et al. Biochemical fingerprint compared with ribotyping and pulsed-field gel electrophoresis of DNA for epidemiology typing of enterococci. **J. Clin. Microb.**, v. 33, n. 11, p. 2812-2817, 1995.

KUO, A.J. et al. National surveillance on vancomycin-resistant *Enterococcus faecium* in Taiwan: Emergence and widespread of ST414 and a Tn1546-Like element with simultaneous insertion of IS1251-Like and IS1678. **PLoS ONE**, v. 9, n. 2, p. e115555, 2014.

KUO, C.H.; OCHMAN, H. The extinction dynamics of bacterial pseudogenes. **PLoS Genet.**, v. 6, n. 8, p. e1001050, 2010.

KWONG, J. cfml-maskrc. [S. I.; s. n.], 2018. Disponível em: <<https://github.com/kwongj/cfml-maskrc>>. Acesso em: 10 de julho de 2020.

KWONG, J.C. et al. Whole genome sequencing in clinical and public health microbiology. **Pathology**, v. 47, n. 3, p. 199-210, 2015.

LAM, M.M.C. et al. Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. **BMC Genomics**, v. 14, p. 595, 2013.

LAM, M.M.C. et al. Comparative analysis of the first complete *Enterococcus faecium* genome. **J. Bacteriol.**, v. 194, n. 9, p. 2334-2341, 2012.

LANCEFIELD, R.C. A serological differentiation of human and other groups of hemolytic streptococci. **J. Exp. Med.**, v. 57, p. 571-595, 1933.

LANDETE, J.M. et al. Virulence and antibiotic resistance of enterococci isolated from healthy breastfed infants. **Microb. Drug Resist.**, v. 24, n. 1, p. 63-69, 2018.

LANZA, V.F. et al. The plasmidome of firmicutes: Impact on the emergence and spread of resistance to antimicrobials. **Microbiol Spectr.**, v. 3, n. 2, 2015. DOI: 10.1128/microbiolspec.PLAS-0039-2014.

LAPIERRE, P.; LASEK-NESELQUIST, E.; GOGARTEN, J.P. The impact of HGT on phylogenomic reconstruction methods. **Brief. Bioinform.**, v. 15, n. 1, p. 79-90, 2014.

LARSEN, M.V. et al. Multilocus sequence typing of total genome sequenced bacteria. **J. Clin. Microbiol.**, v. 50, n. 4, p. 1355-1361, 2012.

LASCH, P. et al. Insufficient discriminatory power of MALDI-TOF mass spectrometry for typing of *Enterococcus faecium* and *Staphylococcus aureus* isolates. **J. Microbiol.Meth.**, v. 100, p. 58-69, 2014.

- LAVERDE GOMEZ, J.A. et al. A multiresistance megaplasmid pLG1, bearing a *hyE_fm* genomic island in hospital *Enterococcus faecium* isolates. **Int. J. Med. Microb. Suppl.**, v. 301, p. 16S-17S, 2011.
- LEAVIS, H.L.; BONTEN, M.J.M.; WILLEMS, R.J.L. Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. **Curr. Opin. Microbiol.**, v. 9, n. 5, p. 454-460, 2006.
- LEAVIS, H.L. et al. Epidemic and nonepidemic multidrug-resistant *Enterococcus faecium*. **Emerg. Infect. Dis.**, v. 9, n. 9, p. 1108-1115, 2003.
- LEAVIS, H.L. et al. High-level ciprofloxacin resistance from point mutations in *gyrA* and *parC* confined to global hospital-adapted clonal lineage CC17 of *Enterococcus faecium*. **J. Clin. Microb.**, v. 44, p. 1059-1064, 2006.
- LEAVIS, H.L. et al. Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. **PLoS Pathog.**, v. 3, n.1, p. e7, 2007.
- LEBRETON, F. et al. D-Ala-D-Ser VanN-type transferable vancomycin resistance in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 55, p. 4606-4612, 2011.
- LEBRETON, F. et al. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. **mBio**, v. 4, n. 4, p. e00534-13, 2013.
- LEBRETON, F. et al. Tracing the enterococci from Paleozoic origins to the hospital. **Cell**, v. 169, p. 849-861, 2017.
- LEBRETON, F.; WILLEMS, R.J.L.; GILMORE, M.S. *Enterococcus* diversity, origins in Nature, and gut colonization. In: GILMORE, M.S. et al. (ed). **Enterococci: From commensals to leading causes of drug resistant infection**. Boston: Massachusetts Eye and Ear Infirmary, 2014.
- LECLERCQ, R. et al. Changes in enterococcal populations and related antibiotic resistance along medical center-wastewater treatment plant-river continuum. **Appl. Environ. Microbiol.**, v. 79, n. 7, p. 2428-2434, 2013.
- LEE, I. et al. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. **Int. J. Syst. Evol. Microbiol.**, v. 66, p. 1100-1103, 2015.
- LEE, M.C.; MARX, C.J. Repeated, selection-driven genome reduction of accessory genes in experimental population. **PLoS Genet.**, v. 8, n. 5, p. e1002651, 2012.
- LEE, T. et al. Antimicrobial-resistant CC17 *Enterococcus faecium*: The past, the present and the future. **J. Glob. Antimicrob. Resist.**, v. 16, P. 36-47, 2019a.
- LEE, T. et al. Molecular characterization and evolution of the first outbreak of vancomycin resistant *Enterococcus faecium* in Western Australia. **Int. J. Antimicrob. Agents**, v. 53, n. 6, p. 814-819, 2019b.

- LEE, T. et al. A three-year whole genome sequencing perspective of *Enterococcus faecium* sepsis in Australia. **PLoS ONE**, v. 15, n. 2, p. e0229781, 2020.
- LEE, W.G.; LEE, S.M.; KIM, Y.S. Molecular characterization of *Enterococcus faecium* isolated from hospitalized patients in Korea. **Lett. Appl. Microbiol.**, v. 43, p. 274-279, 2006.
- LEELAPORN, A. et al. A novel structure of Tn4001-truncated element, type V, in clinical enterococcal isolates and multiplex PCR for detecting aminoglycoside resistance genes. **Int. J. Antimicrob. Agents**, v. 31, p. 250-254, 2008.
- LEMONIDIS, K. et al. Emergence of an Australian-like *pstS*-null vancomycin resistant *Enterococcus faecium* clone in Scotland. **PLoS ONE**, v. 14, n. 6, p. e0218185, 2019.
- LEÓN-SAMPEDRO, R. et al. Diversity and evolution of the Tn5801-*tet(M)*-Like integrative and conjugative elements among *Enterococcus*, *Streptococcus*, and *Staphylococcus*. **Antimicrob. Agents Chemother.**, v. 60, n. 3, p. 1736-1746, 2016.
- LEONG, K.W.C. et al. Emergence of vancomycin-resistant *Enterococcus faecium* at an Australian hospital: A whole genome sequencing analysis. **Sci. Rep.**, v. 8, n. 6274, 2018. DOI: <https://doi.org/10.1038/s41598-018-24614-6>.
- LERAT, E.; OCHMAN, H. Recognizing the pseudogenes in bacterial genomes. **Nucleic Acids Res.**, v. 33, n. 10, p. 3125-3132, 2005.
- LETUNIC, I.; BORK, P. Interactive tree of life (iTOL) v4: recent updates and new developments. **Nucleic Acids Res.**, v. 47, n. W1, p. W256-W259, 2019.
- LEVIN, B.R.; BERGSTROM, C.T. Bacteria are different: Observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. **Proc. R. Soc.**, v. 97, n. 13, p. 6981-6985, 2000.
- LI, C. et al. A review on heavy metals contamination in soil: Effects, sources and remediation techniques. **Soil Sediment Contam.**, v. 28, n. 4, p. 380-394, 2019.
- LI, H.; DURBIN, R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**, v. 25, n. 14, p. 1754-1760, 2009.
- LI, H. et al. The sequence alignment/map format and SAMtools. **Bioinformatics.**, v. 25, n. 16, p. 2078-2079, 2009.
- LIM, S.K. Pheromone-responsive conjugative vancomycin resistance plasmids in *Enterococcus faecalis* isolates from humans and chicken feces. **Appl. Environ. Microbiol.**, v. 72, n. 10, p. 6544-6553, 2006.
- LINDENSTRAUSS, A.G. et al. Comparison of genotypic and phenotypic cluster analyses of virulence determinants and possible role of CRISPR elements towards their incidence in *Enterococcus faecalis* and *Enterococcus faecium*. **Syst. Appl. Microbiol.**, v. 34, n. 8, p. 553-560, 2011.

LIU, B. et al. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. **Nucleic Acids Res.**, v. 47, n. D1, p. D687-D692, 2019.

LIU, K. et al. Coexistence and association between heavy metals, tetracycline and corresponding resistance genes in vermicomposts originating from different substrates. **Environ. Pollut.**, v. 244, p. 28-37, 2019.

LIU, X. et al. Evidence for recombination in *Mycobacterium tuberculosis*. **J. Bacteriol.**, v. 188, p. 8169-8177, 2006.

LONG, K.S. et al. The *cfr* rRNA methyltransferase confers resistance to phenicols, lincosamides, oxazolidinones, pleuromutilins, and streptogramin A antibiotics. **Antimicrob. Agents Chemother.**, v. 50, p. 2500-2505, 2006.

LÓPEZ, F. et al. Antimicrobial susceptibility and macrolide resistance genes in *Enterococcus faecium* with reduced susceptibility to quinupristin-dalfopristin: level of quinupristin-dalfopristin resistance is not dependent on *erm(B)* attenuator region sequence. **Diagn. Microbiol. Infect. Dis.**, v. 66, p. 73-77, 2010.

LÓPEZ, M. et al. Characterization of the mechanisms of fluoroquinolone resistance in vancomycin-resistant enterococci of different origins. **J. Chemotherapy**, v. 23, p. 87-91, 2011.

LÓPEZ, M. et al. Clonal dissemination of *Enterococcus faecalis* ST201 and *Enterococcus faecium* CC17-ST64 containing Tn5382 – *vanB2* among 16 hospitals in Chile. **Clin. Microbiol. Infect.**, v. 15, p. 586-588, 2009.

MACLEAN, R.C.; SAN MILLAN, A. Microbial evolution: Towards resolving the plasmid paradox. **Curr. Biol.**, v. 25, n. 17, p. R764-R767, 2015.

MADRAZO, M. et al. Predictive factors for *Enterococcus faecalis* in complicated community-acquired urinary tract infections in older patients. **Geriatr. Gerontol. Int.**, v. 20, n. 3, p. 183-186, 2020.

MADSEN, K.T. et al. Virulence factors associated with *Enterococcus faecalis* infective endocarditis: A mini review. **Open Microbiol. J.**, v. 11, p. 1-11, 2017.

MAGIORAKOS, A.P. et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. **Clin. Microbiol. Infect.**, v. 18, n. 3, p. 268-281, 2012.

MAHBUB ALAM, M. et al. Detection of a novel *aph(2'')* allele (*aph[2'']-Ie*) conferring high-level gentamicin resistance and a spectinomycin resistance gene *ant(9)-Ia* (*aad 9*) in clinical isolates of enterococci. **Microb. Drug. Resist.**, v. 11, p. 239-247, 2005.

MAHONY, A.A. et al. Vancomycin-resistant *Enterococcus faecium* sequence type 796 – rapid international dissemination of a new epidemic clone. **Antimicrob. Resist. Infect. Control**, v. 7, p. 44, 2018.

MAIDEN, M.C.J. et al. MLST revisited: the gene-by-gene approach to bacterial genomics. **Nat. Rev. Microbiol.**, v. 11, n.10, p. 728-736, 2013.

MAIDEN, M.C. J. et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. **Proc. Natl. Acad. Sci. USA**, v. 95, p. 3140-3145, 1998.

MALATHUM, K. et al. Repetitive sequence-based PCR versus pulsed-field gel electrophoresis for typing of *Enterococcus faecalis* at the subspecies level. **J. Clin. Microbiol.**, v. 6, p. 211-215, 1998.

MANCINI, E.M.; BOTEN, M.J. Vancomycin-resistant enterococci: consequences for therapy and infection control. **Clin. Microbiol. Infect.**, v. 11, p. 43-56, 2005.

MARACCINI, P.A.; FERGUSON, D.M.; BOEHM, A.B. Diurnal variation in *Enterococcus* species composition in polluted ocean water and a potential role for the enterococcal carotenoid in protection against photoinactivation. **Appl. Environ. Microbiol.**, v. 78, p. 305-310, 2012.

MARCHI, A.P. et al. Vancomycin-resistant enterococci isolates colonizing and infecting haematology patients: clonality and virulence and resistance profile. **J. Hosp. Infect.**, v. 99, p. 346-355, 2018.

MARKWART, R. et al. The rise in vancomycin-resistant *Enterococcus faecium* in Germany: data from the German antimicrobial resistance surveillance (ARS). **Antimicrob. Resist. Infect. Control.**, v. 8, 2019. DOI: 10.1186/s13756-019-0594-3.

MARSHALL, C.G. et al. Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. **Antimicrob. Agents Chemother.**, v. 42, p. 2215-2220, 1998.

MARTIN, W. Mosaic bacterial chromosomes: a challenge in route to a tree of genomics. **Bioessays**, v. 21, p. 99-104, 1999.

MASON, A. et al. Accuracy of different bioinformatics methods in detecting antibiotic resistance and virulence factors from *Staphylococcus aureus* whole-genome sequences. **J. Clin. Microb.**, v. 56, n. 9, p. e01815-e01817, 2018.

MATO, R. et al. Multiplicity of genetic background among vancomycin-resistant *Enterococcus faecium* isolates recovered from an outbreak in a New York city hospital. **Microb. Drug. Res.**, v. 2, p. 309-317, 1996.

MATONO, T. et al. Emergence of a daptomycin-non-susceptible *Enterococcus faecium* strain that encodes mutations in DNA repair genes after high-dose daptomycin therapy. **BMC Res. Notes**, v. 9, n. 197, 2016. DOI: 10.1186/s13104-016-2003-9.

MAYORAL-TERÁN, C. et al. High phenotypic and genotypic diversity of *Enterococcus faecium* from clinical and commensal isolates in third level hospital. **Microb. Drug. Resist.**, v.26, n. 3, p. 227-237, 2020.

MCCRACKEN, M. et al. Molecular epidemiology of vancomycin-resistant enterococcal bacteremia: results from the Canadian Nosocomial Infection Surveillance Program, 1999-2009. **J. Antimicrob. Chemother.**, v. 68, p. 1505-1509, 2013.

MCDERMOTT, P.F. et al. Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal *Salmonella*. **Antimicrob. Agents Chemother.**, v. 60, p. 15-20, 2016.

MCGANN, P. et al. Real time application of whole-genome sequencing for outbreak investigation – What is an achievable turnaround time? **Diagn. Microbiol. Infect. Dis.**, v. 85, n. 3, p. 277-282, 2016.

MCGOWAN, J.E. Debate-Guidelines for control of glycopeptide-resistant enterococci (GRE) have not yet worked. **J. Hosp. Infect.**, v. 57, p. 281-284, 2004.

MEDEIROS, A.W. et al. *Enterococcus* species diversity in fecal samples of wild marine species as determined by real-time PCR. **Can. J. Microbiol.**, v. 63, n. 2, p. 129-136, 2017.

MEDINI, D. et al. The microbial pan-genome. **Curr. Opin. Genet. Dev.**, v. 15, p. 589-594, 2005

MEDINI, D. et al. Microbiology in the post-genomic era. **Nat. Rev. Microbiol.**, v. 6, n. 6, p. 419-430, 2008.

MELEGH, S. et al. Dissemination of *vanA*-type *Enterococcus faecium* isolates in Hungary. **Microb. Drug Resist.**, v. 24, n. 9, p. 1376-1390, 2018.

MENDES, R.E. et al. Longitudinal (2001-14) analysis of enterococci and VRE causing invasive infections in European and US hospitals, including a contemporary (2010-13) analysis of oritavancin in vitro potency. **J. Antimicrob. Chemother.**, v. 71, p. 3453-3458, 2016.

MERINO, S.; TOMÁS, J.M. Bacterial capsule and evasion immune responses. In: eLS. Chichester: John Wiley & Sons, 2015. DOI: <https://doi.org/10.1002/9780470015902.a0000957.pub4>

MERLO, T.P.; DABUL, A.N.G.; CAMARGO, I.L.B.C. Different VanA elements in *E. faecalis* and in *E. faecium* suggest at least two origins of Tn1546 among VRE in a Brazilian hospital. **Microb. Drug Resist.**, v. 21, n. 3, p. 320-328, 2015.

MERQUIOR, V.L.C. et al. Emergence and characterization of *vanB* vancomycin resistant *Enterococcus faecalis* in Rio de Janeiro, Brazil. **Mem. Inst. Oswaldo Cruz**, v. 107, p. 557-560, 2012.

MICHAEL, K.E.; NO, D.; ROBERTS, M.C. *vanA* positive multidrug-resistant *Enterococcus* spp. isolated from surfaces of a US hospital laundry facility. **J. Hosp. Infect.**, v. 95, n. 2, p. 218-223, 2017.

MICHOD, R.E.; BERNSTEIN, H.; NEDELCO, A.M. Adaptive value of sex in microbial pathogens. **Infect. Genet. Evol.**, v. 8, p. 267-285, 2008.

- MIKALSEN, T. et al. Investigating the mobilome in clinically important lineages of *Enterococcus faecium* and *Enterococcus faecalis*. **BMC Genomics**, v. 16, p. 282, 2015.
- MILLARD, C.M. et al. Evolution of the capsular operon of *Streptococcus iniae* in response to vaccination. **Appl. Environ. Microbiol.**, v. 78, n. 23, p. 8219-8226, 2012.
- MILLER, W.R. et al. Vancomycin-resistant enterococci: Therapeutic challenges in the 21st Century. **Infect. Dis. Clin. North. Am.**, v. 30, p. 415–439, 2016.
- MINGEOT-LECLERCQ, M.P.; GLUPCZYNSKI, Y.; TULKENS, P.M. Aminoglycosides: activity and resistance. **Antimicrob. Agents Chemother.**, v. 43, p. 727-737, 1999.
- MIRA, A.; OCHMAN, H.; MORAN, N.A. Deletional bias and the evolution of bacterial genomes. **Trends Genet.**, v. 17, n. 10, p. 589-596, 2001.
- MOHN, S.C. Duplex real-time PCR assay for rapid detection of ampicillin-resistant *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 48, p. 556-560, 2004.
- MONTEIRO DA SILVA, B.N. et al. Expression of vanA-type vancomycin resistance in a clinical isolate *Enterococcus faecium* showing insertion of IS19 in the *vanS* gene. **Int. J. Antimicrob. Agents**, v. 55, n. 4, p. 105897, 2020.
- MORAN, N.A. Microbial minimalism: Genome reduction in bacterial pathogens. **Cell**, v. 108, n. 5, p. 583-586, 2002.
- MORAN, R.A. et al. Prediction of antibiotic resistance from antibiotic resistance genes detected in antibiotic-resistant commensal *Escherichia coli* using PCR or WGS. **J. Antimicrob. Chemother.**, v. 72, p. 700-704, 2017.
- MORITZ, E.M.; HERGENROTHER, P.J. Toxin-antitoxin systems are ubiquitous and plasmid-encoded in vancomycin-resistant enterococci. **Proc. Natl. Acad. Sci. USA**, v. 104, p. 311-316, 2007.
- MORRONI, G. et al. Characterization of a new transferable MDR plasmid carrying the *pbp5* gene from a clade B commensal *Enterococcus faecium*. **J. Antimicrob. Chemother.**, v. 74, n. 4, p. 843-850, 2019.
- MOUSTAFA, A.M.; LAL, A.; PLANET, P.J. Comparative genomics in infectious disease. **Curr. Opin. Microbiol.**, v. 53, p. 61-70, 2020.
- MOUSTAFA, A.M.; PLANET, P.J. WhatsGNU: a tool for identifying proteomic novelty. **Genome Biol.**, v. 58, 2020. DOI: <https://doi.org/10.1186/s13059-020-01965-w>.
- MUKAKA, M.M. A guide to appropriate use of correlation coefficient in medical research. **Malawi Med. J.**, v 24, n. 3, p. 69-71, 2012.
- MUNDY, L.M; SAHM, D.F.; GILMORE, M. Relationship between enterococcal virulence and antimicrobial resistance. **Clin. Microbiol. Rev.**, v. 13, n. 4, p. 513-522, 2000.

- MUNITA, J.M. et al. Correlation between mutations in *liaFSR* of *Enterococcus faecium* and MIC of daptomycin: Revisiting daptomycin breakpoints. **Antimicrob. Agents Chemother.**, v. 56, p. 4354-4359, 2012.
- MUTSCHLER, H.; MEUNHART, A. ε/ξ: their role in resistance, virulence, and their potential for antibiotic development. **J. Mol. Med.**, v. 89, p. 1183-1194, 2011
- NALLAPAREDDY, S.R. et al. A functional collagen adhesin gene, *acm*, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen. **Infect. Immun.**, v. 76, p. 4110-4119, 2008.
- NASER, S.M. et al., *Enterococcus canintestini* sp. nov., from faecal samples of healthy dogs. **Int. J. Syst. Evol. Microbiol.**, v. 55, n. 5, p. 2177-2182, 2005.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **BLAST®**: Basic Local Alignment Search Tool. Bethesda: NCBI, 2020
- NEVES CASTRO, P.B. et al. Antibiotic consumption in developing countries defies global commitments: an overview on Brazilian growth in consumption. **Environ. Sci. Pollut. Res. Int.**, v. 27, p. 21013-21020, 2020.
- NIGO, M. What's new in treatment of enterococcal endocarditis? **Curr. Infect. Dis. Rep.**, v. 10, p. 431, 2014.
- NISHIDA, H. Comparative analyses of base compositions, DNA sizes, and dinucleotide frequency profiles in archaeal and bacterial chromosomes and plasmids. **Int. J. Evol. Biol.**, v. 2012, 2012. DOI: <https://doi.org/10.1155/2012/342482>.
- NISHIDA, H. Genome DNA sequence variation, evolution and function in bacteria and archaea. **Curr. Issues Mol. Biol.**, v. 15, p. 19-24, 2013.
- NISHIMOTO, Y. et al. Analysis of the prevalence of tetracycline resistance genes in clinical isolates of *Enterococcus faecalis* and *Enterococcus faecium* in a Japanese Hospital. **Microb. Drug Resist.**, v. 11, n. 2, p. 146-153, 2005.
- NIU, H. et al. The prevalence of aminoglycoside-modifying enzyme and virulence genes among enterococci with high-level aminoglycoside resistance in Inner Mongolia, China. **Braz. J. Microbiol.**, v. 47, n. 3, p. 691-696, 2016.
- NOSKIN, G.A. et al. High-level gentamicin resistance in *Enterococcus faecalis* bacteremia. **J. Infect. Dis.**, v. 164, p. 1212-1215, 1991.
- NOSKIN, G.A. et al. Persistent contamination of fabric-covered furniture by vancomycin-resistant enterococci: Implications for upholstery selection in hospitals. **Am. J. Infect. Control**, v. 28, n. 4, p. 311-313, 2000.
- NOVAIS, C. et al. Co-diversification of *Enterococcus faecium* core genomes and PBP5: Evidences of *pbp5* horizontal transfer. **Front. Microbiol.**, v. 7, n. 1581, 2016. DOI: 10.3389/fmicb.2016.01581.

NOVAIS, C. et al. Different genetic supports for the *tet(S)* gene enterococci. **Antimicrob. Agents Chemother.**, v. 56, n. 11, p. 6014-6018, 2012.

NOVAIS, C. et al. High occurrence and persistence of antibiotic-resistant enterococci in poultry food samples in Portugal. **J. Antimicrob. Chemother.**, v. 56, p. 1139-1143, 2005.

NOVAIS, C. et al. Spread of multidrug-resistant *Enterococcus* to animals and humans: an underestimated role for the pig farm environment. **J. Antimicrob. Chemother.**, v. 68, n. 12, p. 2746-2754, 2013.

NOWAKIEWICZ, A. et al. Determination of resistance and virulence genes in *Enterococcus faecalis* and *E. faecium* strains isolated from poultry and their genotypic characterization by ADSRRS-fingerprint. **Poult. Sci.**, v. 96, n. 4, p. 986-996, 2017.

OCHOA, S.A. et al. Molecular analysis and distribution of multidrug-resistant *Enterococcus faecium* isolates belonging to clonal complex 17 in a tertiary care center in Mexico city. **BMC Microbiol.**, v. 13, p. 291, 2013.

OH, W.S. et al. High rate of resistance to quinupristin-dalfopristin in *Enterococcus faecium* clinical isolates from Korea. **Antimicrob. Agents Chemother.**, v. 19, n. 12, p. 5176-5178, 2005.

ORAVCOVÁ, V. Wild corvid birds colonized with vancomycin-resistant *Enterococcus faecium* of human origin harbor epidemic *vanA* plasmids. **Environmental International**, v. 118, p. 125-133, 2018.

OSTRER, L. et al. Analysis of mutational patterns in quinolone resistance-determining regions of GyrA and ParC of clinical isolates. **Int. J. Antimicrob. Agents**, v. 53, p. 318-324, 2019.

PADMASINI, E.; PADMARAJ, R.; SRIVANI-RAMESH, S. High level aminoglycoside resistance and distribution of aminoglycoside resistant genes among clinical isolates of *Enterococcus* species in Chennai, India. **ScientificWorldJournal**, eCollection 329157, 2014. DOI: <https://doi.org/10.1155/2014/329157>.

PAGANELLI, F.L. et al. Distinct *SagA* from hospital-associated clade A1 *Enterococcus faecium* strains contributes to biofilm formation. **Appl. Environ. Microbiol.**, v. 81, p. 6873-6882, 2015.

PAL, C. et al. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. **BMG Genomics**, v. 16, n. 964, 2015. DOI: <https://doi.org/10.1186/s12864-015-2153-5>

PALAZZO, I.C. et al. Changes in vancomycin-resistant *Enterococcus faecium* causing outbreaks in Brazil. **J. Hosp. Infect.**, v. 79, p. 70-74, 2011.

PALEPOU, M.F.I. et al. Molecular analysis of diverse elements mediating *vanA* glycopeptide resistance in enterococci. **J. Antimicrob. Chemother.**, v. 42, p. 605-612, 1998.

- PALMER, K.L. et al. Comparative genomics of enterococci: Variation in *Enterococcus faecalis*, clade structure in *Enterococcus faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. **mBio**, v. 3, n. 1, p. e00318-11, 2012.
- PALMER, K.L.; GILMORE, M.S. Multidrug-resistant enterococci lack CRISPR-Cas. **mBio**, v.1, n. 4, p. e00227-10, 2010.
- PANESSO, D. et al. Molecular epidemiology of vancomycin-resistant *Enterococcus faecium*: a prospective, multicenter study in South American hospitals. **J. Clin. Microbiol.**, v. 48, p. 1562-1569, 2010.
- PANESSO, D. et al. The *hylefm* gene in pHylEfm of *Enterococcus faecium* is not required in pathogenesis of murine peritonitis. **BMC Microbiol.**, v. 11, p. 20, 2011.
- PAPADIMITRIOU-OLIVGERIS, M. et al., Risk factors for enterococcal infection and colonization by vancomycin-resistant enterococci in critically ill patients. **Infection**, v. 42, p. 1013-1022, 2014.
- PAPAGIANNITSIS, C.C. et al. First description in Europe of the emergence of *Enterococcus faecium* ST117 carrying both *vanA* and *vanB* genes, isolated in Greece. **J. Glob. Antimicrob. Resist.**, v. 11, p. 68-70, 2017.
- PARK, I.J. VanB phenotype-vanA genotype *Enterococcus faecium* with heterogeneous expression of teicoplanin resistance. **J. Clin. Microb.**, v. 46, n. 9, p. 3091-3093, 2008.
- PARK, S.H. et al. Molecular epidemiology of vancomycin-resistant *Enterococcus faecium* bloodstream infections among patients with neutropenia over a 6-year period in South Korea. **Microb. Drug Resist.**, v 17, n. 1, p. 59-65, 2011.
- PARTE, A.C. List of prokaryotic names with standing in nomenclature (bacterio.net), 20 years on. **Int. J. Syst. Evol. Microbiol.**, v. 68, p. 1825-1829, 2018.
- PASQUAROLI, S. et al. Erythromycin- and copper-resistant *Enterococcus hirae* from marine sediment and co-transfer of *erm(B)* and *tcrB* to human *Enterococcus faecalis*. **Diagn. Microbiol. Infect. Dis.**, v. 80, p. 26-28, 2014.
- PEGUES, D.A. et al. Emergence and dissemination of a highly vancomycin-resistant *vanA* strain of *Enterococcus faecium* at a large teaching hospital. **J. Clin. Microbiol.**, v. 35, p. 1565-1570, 1997.
- PETSARIS, O. et al. Combined antimicrobial resistance in *Enterococcus faecium* isolated from chickens. **Appl. Environ. Microbiol.**, v. 71, n. 5, p.2796-2799, 2005.
- PEYVASTI, V. S. et al. High-level aminoglycoside resistance and distribution of aminoglycoside resistance genes among *Enterococcus* spp. clinical isolates in Tehran, Iran. **J. Global Antimicrob. Res.**, v. 20, p. 318-323, 2020.
- PFALLER, M.A. et al. Temporal and geographic variation in antimicrobial susceptibility and resistance patterns of enterococci: Results from the SENTRY antimicrobial surveillance program, 1997-2016. **Open Forum Infect. Dis.**, v. 6, n. supplement_1, p. S54-S62, 2019.

PUBLIC HEALTH ENGLAND. Health Protection Report: latest surveillance report. **HPR**, v. 11, n. 15, 2016. Disponível em: <<https://www.gov.uk/government/publications/enterococcus-spp-bacteraemia-voluntary-surveillance>>. Acesso em: 20 maio 2020.

PUBLIC HEALTH ENGLAND. Health Protection Report: latest surveillance report. **HPR**, v. 13, n. 33, 2019. Disponível em: <<https://www.gov.uk/government/publications/enterococcus-spp-bacteraemia-voluntary-surveillance>>. Acesso em: 20 maio 2020.

PIDOT, S.J. et al. Increasing tolerance of hospital *Enterococcus faecium* to handwash alcohols. **Sci. Transl. Med.**, v. 10, n. 452, p. eaar6115, 2018.

PINHOLT, M. et al. Genomic analysis of 495 vancomycin-resistant *Enterococcus faecium* reveals broad dissemination of *vanA* plasmid in more than 19 clones from Copenhagen, Denmark. **J. Antimicrob. Chemother.**, v. 72, n. 1, p. 40-47, 2017.

PINHOLT, M. et al. Multiple hospital outbreaks of *vanA* *Enterococcus faecium* in Denmark, 2012-13, investigated by WGS, MLST and PFGE. **J. Antimicrob. Chemother.**, v. 70, n. 9, p. 2474-2482, 2015.

PINHOLT, M. et al. WGS of 1058 *Enterococcus faecium* from Copenhagen, Denmark, reveals rapid clonal expansion of vancomycin-resistant clone ST80 combined with widespread dissemination of *vanA*-containing plasmid and acquisition of a heterogeneous accessory genome. **J. Antimicrob. Chemother.**, v. 74, n. 7, p. 1776-1785, 2019.

PINTO, T.C.A. et al. Phenotypic and molecular characterization of optochin-resistant *Streptococcus pneumoniae* isolates from Brazil, with description of five novel mutations in the ATPC gene. **J. Clin. Microb.**, v. 51, p. 3242-3249, 2013.

POETA, P. et al. Polymorphisms of the *pbp5* gene and correlation with ampicillin resistance in *Enterococcus faecium* isolates of animal origin. **J. Med. Microbiol.**, v. 56, p. 236-240, 2007.

POETA, P. et al. Study of faecal colonization by *vanA*-containing *Enterococcus* strains in healthy humans, pets, poultry and wild animals in Portugal. **J. Antimicrob. Chemother.**, v. 55, n. 2, p. 278-280, 2005.

POSADA, D.; CRANDALL, K.A. The effect of recombination on the accuracy of phylogeny estimation. **J. Mol. Evol.**, v. 54, n. 396-402, 2002.

PRADO, G.V.B. et al. Virulence and resistance pattern of a novel sequence type of linezolid-resistant *Enterococcus faecium* identified by whole-genome sequencing. **J. Glob. Antimicrob. Resist.**, v. 6, p. 27-31, 2016.

PRICE, V.J. et al. CRISPR-Cas and restriction-modification act additively against conjugative antibiotic resistance plasmid transfer in *Enterococcus faecalis*. **mSphere**, v. 1, n. 3, p. e00064-16, 2016.

PUTNAM, S.D. et al. Worldwide summary of telavancin spectrum and potency against Gram-positive pathogens: 2007 to 2008 surveillance results. **Diagn. Microbiol. Infect. Dis.**, v.67, p. 359-368, 2010.

- QIN, X. et al. Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. **BMC Microb.**, v. 12, n. 135, 2012. DOI: <https://doi.org/10.1186/1471-2180-12-135>
- QU, T.T. et al. Novel vancomycin-resistance transposon plasmid replicon types, and virulence factors of vancomycin-resistant enterococci in Zhejiang, China. **Microb. Drug Resist.**, v.18, n. 2, p. 183-188, 2012.
- RAI, P.K. et al. Heavy metals in food crops: Health risks, fate, mechanisms, and management. **Environ. Int.**, v. 125, p. 365-385, 2019.
- RAJKUMARI, N.; MATHUR, P.; MISRA, M.C. Soft tissue and wound infections due to *Enterococcus* spp. among hospitalized trauma patients in a developing country. **J. Glob. Infect. Dis.**, v. 6, n. 4, p. 189-193, 2014.
- RAMBALDI, M.L. et al. Combination of co-trimoxazole and ciprofloxacin as therapy of a patient with infective endocarditis caused by an enterococcus highly resistant to gentamicin. **J. Antimicrob. Chemother.**, v. 40, p. 737-738, 1997.
- RAMBAUT, A. FigTree. Versão 1.4.4. [S. I.; s. n.], 2018. Disponível em: <http://tree.bio.ed.ac.uk/software/figtree/>. Acesso em: 05 de junho de 2020.
- RAMBAUT, A. et al. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. **Syst. Biol.**, v. 67, n. 5, p. 901-904, 2018.
- RAMIREZ, M.S.; TOLMASKY, M.E. Aminoglycoside modifying enzymes. **Drug Resist. Updat.**, v. 13, n. 6, p. 151-171, 2010.
- RATHNAYAKE, I.; HARGREAVEZ, M.; HUYGENS, F. Antibiotic resistance and virulence traits in clinical and environmental *Enterococcus faecalis* and *Enterococcus faecium* isolates. **Syst. Appl. Microbiol.**, v. 35, n. 5, p. 326-333, 2012.
- RATHNAYAKE, I.; HARGREAVEZ, M.; HUYGENS, F. SNP diversity of *Enterococcus faecalis* and *Enterococcus faecium* in a South East Queensland waterway, Australia, and associated antibiotic resistance gene profiles. **BMC Microbiol.**, v. 11, n. 201, 2011. DOI: 10.1186/1471-2180-11-201.
- RAVEN, K.E. et al. A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland. **Genome Res.**, v. 26, n. 10, p. 1388-1396, 2016.
- RAVEN, K.E. et al. Complex routes of nosocomial vancomycin-resistant *Enterococcus faecium* transmission revealed by genome sequencing. **Clin. Infect. Dis.**, v. 64, p. 886-893, 2017.
- REDFIELD, R.J. Do bacteria have sex? **Nat. Rev. Genet.**, v. 2, p. 634-639, 2001.
- REHMAN, M.A. et al. Genotypes and phenotypes of enterococci isolated from broiler chickens. **Front. Sustain. Food Syst.**, v. 2, n. 83, 2018. DOI: <https://doi.org/10.3389/fsufs.2018.00083>.

- RESENDE, M. et al. Emergence of *vanA* vancomycin-resistant *Enterococcus faecium* in a hospital in Porto Alegre, South Brazil. **J. Infect. Dev. Ctries.**, v. 8, p. 160-167, 2014.
- REST, J.S.; MINDELL, D.P. Retroids in Archaea: phylogeny and lateral origin. **Mol. Biol. Evol.**, v. 20, p. 1134-1142, 2003.
- REYNOLDS, C.R.; ISLAM, S.A.; STERNBERG, M.J.E. EzMol: A web server wizard for the rapid visualization and image production of protein and nucleic acid structures. **J. Mol. Biol.**, v. 430, n. 15, p. 2244-2248, 2018.
- RICE, L.B. et al. A potential virulence gene, *hyl_{Efm}*, predominates in *Enterococcus faecium* of Clinical Origin. **J. Infect. Dis.**, v. 187, n. 3, p. 508-512, 2003.
- RICE, L.B. et al. Impact of specific *pbp5* mutations on expression of B-lactam resistance in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 48, n. 8, p. 3028-3032, 2004.
- RICE, L.B. et al. Occurrence of high-level aminoglycoside resistance in environmental isolates of enterococci. **Appl. Environ. Microbiol.**, v. 61, p. 374-376, 1995.
- RICE, L.B.; THRISDOTTIR, A.S. The prevalence of sequences homologous to IS256 in clinical enterococcal isolates. **Plasmid**, v. 32, p. 344-349, 1994.
- RIDENHOUR, M.B. et al. A novel tetracycline-resistant determinant, *tet(U)*, is encoded on the plasmid pKQ10 in *Enterococcus faecium*. **Plasmid**, v. 36, p. 71-80, 1996.
- RILEY, M.A.; LIZOTTE-WANIEWSKI, M. Population genomics and the bacterial species concept. In: GOGARTEN, M.B.; GOGARTEN, J.P.; OLENDZENSKI, L.C. (ed). **Horizontal gene transfer: Genomes in Flux**. New York: Humana Press, 2009. p. 367-377.
- RIOS, R. et al. Genomic epidemiology of vancomycin-resistant *Enterococcus faecium* (VRE_{fm}) in Latin America: Revisiting the global VRE population structure. **Sci. Rep.**, v. 10, n. 5636, 2020. DOI: <https://doi.org/10.1038/s41598-020-62371-7>.
- RIZZOTT, L. et al. Molecular diversity and transferability of the tetracycline resistance gene *tet(M)*, carried on Tn916-1545 family transposons, in enterococci from a total food chain. **Antonie van Leeuwenhoek**, v. 96, p. 42-52, 2009.
- ROBERTS, A.P. et al. Characterization of the ends and target site of a novel tetracycline resistance-encoding conjugative transposon from *Enterococcus faecium* 664.1H1. **J. Bacteriol.**, v. 188, p. 4356-4361 2006.
- ROBERTS, M.C. et al. Vancomycin resistant *Enterococcus* spp. from crows and their environment in metropolitan Washington State, USA: Is there a correlation between VRE positive crows and the environment? **Vet. Microbiol.**, v. 194, p. 48-54, 2016.
- ROCHA, E.P.C.; DANCHIN, A. Base composition bias might result from competition for metabolic resources. **Trends Genetic.**, v. 18, n. 6, p. 291-294, 2002.

RODRÍGUEZ-BELTRÁN, J. et al. Genetic dominance governs the evolution and spread of mobile genetic elements in bacteria. **Proc. R. Soc.**, v. 117, n. 12, p. 15755-15762, 2020.

ROKAS, A.; CARROLL, S.B. Frequent and widespread parallel evolution of protein sequences. **Mol. Biol. Evol.**, v. 25, p. 1943-1953, 2008.

ROSSELLÓ-MÓRA, R.; AMANN, R. Past and future species definition for bacteria and archaea. **System. Appl. Microbiol.**, v. 38, p. 209-216, 2015.

ROSVOLL, T.C.S. et al. PCR-based plasmid typing in *Enterococcus faecium* strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHT β -related replicons associated with glycopeptide resistance and stabilizing toxin-antitoxin systems. **FEMS Immunol. Med. Microbiol.**, v. 58, n. 2, p. 254-268, 2010.

ROZAS, J. et al. DnaSP 6: DNA sequence polymorphism analysis of large datasets. **Mol. Biol. Evol.**, v. 34, p. 3299-3302, 2017.

RUIZ-RIPA, L. et al. Detection of *poxxA*- and *optrA*- carrying *E. faecium* isolates in air samples of a Spanish swine farm. **J. Glob. Antimicrob. Res.**, v. 22, p. 28-31, 2020.

RUZON, F.I. et al. Virulence determinants in Vancomycin-resistant *Enterococcus faecium* *vanA* isolated from different sources at University Hospital of Londrina, Paraná, Brazil. **J. Microbiol.**, v. 48, n. 6, p. 814-821, 2010.

RYAN, L. et al. Epidemiology and molecular typing of VRE bloodstream isolates in an Irish tertiary care hospital. **J. Antimicrob. Chemother.**, v. 70, n. 10, p. 2718-2724, 2015.

SAAP, J. The structure of microbial evolutionary theory. **Stud. Hist. Philos. Sci.**, v. 38, n. 4, p. 780-795, 2007.

SACRAMENTO, A.G. et al. Changed epidemiology during intra and interhospital spread of high-risk clones of *vanA*-containing *Enterococcus* in Brazilian hospitals. **Diagn. Microbiol. Infect. Dis.**, v. 88, p. 348-351, 2017.

SADER, H.S. et al. Tigecycline antimicrobial activity tested against clinical bacteria from Latin American medical centres: results from SENTRY antimicrobial surveillance program (2011-2014). **Int. J. Antimicrob. Agents**, v. 48, p. 144-150, 2016.

SADOWY, E.; LUCZKIEWICZ, A. Drug-resistant and hospital-associated *Enterococcus faecium* from wastewater, riverine estuary and anthropogenically impacted marine catchment basin. **BMC Microb.**, v. 14, n. 66, 2014. DOI: 10.1186/1471-2180-14-66.

SALYERS, A.A.; AMÁBILE-CUEVAS, C.F. Why are antibiotic resistance genes so resistant to elimination? **Antimicrob. Agents Chemother.**, v. 41, n. 11, p. 2321-2325, 1997.

SANDERSON, H. et al. Comparative genomics of multidrug-resistant *Enterococcus* spp. isolated from wastewater treatment plants. **BMC Microbiol.**, v. 20, n. 1, p. 20, 2020.

SANKHLA, M.S. Heavy metals contamination in water and their hazardous effect on human health- A review. **Int. J. Curr. Microbiol. App. Sci.**, v. 5, n. 10, p. 759-766, 2016.

SAN MILLAN, A.; MACLEAN, R.C. Fitness costs of plasmids: a limit to plasmid transmission. **Microbiol. Spectrum**, v. 5, n. 5, p. 0016, 2017.

SANTAGATI, M.; CAMPANILE, F.; STEFANI, S. Genomic diversification of enterococci in hosts: the role of the mobilome. **Front. Microbiol.**, v. 3, p. 95, 2012.

SANTAJIT, S.; INDRAWATTANA, N. Mechanisms of antimicrobial resistance in ESKAPE Pathogens. **Biomed. Res. Int.** 2475067, v. 2016, p. 2475067, 2016.

SANTOS, B.A. CRISPR elements and their association with antimicrobial resistance and virulence genes among vancomycin-resistant and vancomycin-susceptible enterococci recovered from human and food sources. **Infect. Genet. Evol.**, v. 80, n. 104183, 2020. DOI: 10.1016/j.meegid.2020.104183.

SANTOS, B.A. et al. Major globally disseminated clonal complexes of antimicrobial resistant enterococci associated with infections in cancer patients in Brazil. **Infect. Genet. Evol.**, v. 55, p. 56-62, 2017.

SCHABERG, D.R.; CULVER, D.H.; GAYNES, R.P. Major trends in the microbial etiology of nosocomial infections. **Am. J. Med.**, v. 91, p. 72S-75S, 1991.

SCHLEIFER, K.H.; KLIPPER-BALZ, R. Transfer of *Streptococcal faecalis* and *Streptococcus faecium* to the genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. **Int. J. Syst. Bacteriol.**, v. 34, p. 31-34, 1984.

SCHLOISSNIG, S. et al. Genomic variation landscape of human gut microbiome. **Nature**, v. 493, p. 45-50, 2013.

SCHULTHESS, B. et al. Identification of Gram-positive cocci by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry: comparison of different preparation methods and implementation of a practical algorithm for routine diagnostics. **J. Clin. Microbiol.**, v. 51, p. 1834-1840, 2013.

SEEDAT, J. et al. Rapid emergence of resistance to linezolid during linezolid therapy of an *Enterococcus faecium* infection. **Antimicrob. Agents Chemother.**, v. 50, p. 4127-4219, 2006.

SEEMANN, T. mlst. Versão 2.19.0. [S. l.; s. n.], 2015. Disponível em: <<https://github.com/tseemann/mlst>>. Acesso em: 30 maio 2020.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics.**, v. 30, n. 14, p. 2068-2069, 2014.

SEEMANN, T. snippy. Versão 4.6.0. [S. l.; s. n.], 2015. Disponível em: <<https://github.com/tseemann/snippy>>. Acesso em: 30 maio 2020.

SEILER, C.; BERENDONK, T.U. Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. **Front. Microbiol.**, v. 3, n. 399, 2012. DOI: <https://doi.org/10.3389/fmicb.2012.00399>.

- SHARMA, P. et al. Comparison of antimicrobial resistance and pan-genome of clinical and non-clinical *Enterococcus cecorum* from poultry using whole genome sequencing. **Foods**, v. 9, n. 6, p. 686, 2020.
- SHEPPARD, S.K. et al. Convergence of *Campylobacter* species: implications for bacterial evolution. **Science**, v. 320, p. 237-239, 2008.
- SHEPPARD, S.K.; MAIDEN, M.C.J. The evolution of *Campylobacter jejuni* and *Campylobacter coli*. **Cold Spring Harb. Perspect. Biol.**, v. 7, p. a018119, 2015.
- SHOKOOHIZADEH, L. et al. Antimicrobial resistance patterns and virulence factors of enterococci isolates in hospitalized burn patients. **BMC Res. Notes**, v. 11, n. 1, p.1, 2018.
- SIFAOU, F. et al. Role of penicillin-binding protein 5 in expression of ampicillin resistance and peptidoglycan structure in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 45, p. 2594-2697, 2001.
- SIGUIER, P. et al. ISfinder: the reference centre for bacterial insertion sequences. **Nucleic Acids Res.**, v. 34, p. D32-D36. 2006.
- SILLANPAA, J. et al. Identification and phenotypic characterization of a second collagen adhesin, Scm, and genome-based identification and analysis of 13 other predicted MSCRAMMs, including four distinct pilus loci, in *Enterococcus faecium*. **Microbiology**, v. 154, p. 3199-3211, 2008.
- SILVA, L.P.P. et al. Genetic features and molecular epidemiology of *Enterococcus faecium* isolated in two university hospital in Brazil. **Diagn. Microbiol. Infect. Dis.**, v. 74, p. 267–271, 2012.
- SILVA, V. et al. Genetic characterization of *vanA-Enterococcus faecium* isolates from wild red-legged partridges in Portugal. **Microb. Drug Resist.**, v. 24, n. 1, p. 89-94, 2018.
- SILVA, V.L. et al. Occurrence of multidrug-resistant and toxic-metal tolerant enterococci in fresh feces from urban pigeons in Brazil. **Microbes Environ.**, v. 27, n. 2, p. 179-185, 2012.
- SILVEIRA, E. et al. Co-transfer of resistance to high concentration of copper and first-line antibiotics among *Enterococcus* from different origins (humans, animals, the environment and foods) and clonal lineages. **J. Antimicrob. Chemother.**, v. 69, n. 4, p. 899-906, 2014.
- SIVERTSEN, A. et al. A multicenter hospital outbreak in Sweden caused by introduction of *vanB2* transposon into a stably maintained pRUM-plasmid in an *Enterococcus faecium* ST192 clone. **PLoS ONE**, v. 8, n. 8, p. e103274, 2014.
- SLATER, F.R. et al. Progress towards understanding the fate of plasmids in bacterial communities. **FEMS Microbiol. Ecol.**, v. 66, n. 1, p. 3-13, 2008.
- SMALLA, K. et al. Distribution of streptothricin acetyltransferase encoding determinants among environmental bacteria. **Mol. Ecol.**, v. 2, p. 27-33, 1993.

SOHEILI, S. et al. Wide distribution of virulence genes among *Enterococcus faecium* and *Enterococcus faecalis* clinical isolates. **ScientificWorldJournal**, v. 2014, n. 623174, 2014. DOI: 10.1155/2014/623174.

SONG, J.H. et al. High frequency of vancomycin-resistant *Enterococcus faecium* isolates with VanB phenotype and *vanA* genotype in Korean hospitals. **Diagn. Microbiol. Infec. Dis.**, v. 56, p. 401-406, 2006.

SPRATT, B.G. et al. Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria* species. **J. Mol. Evol.**, v. 34, p. 115-125, 1992.

STAMATAKIS, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. **Bioinformatics**, v. 30, n. 9, p. 1312-1313, 2014.

STAMPONE, L. et al. Clonal spread of a vancomycin-resistant *Enterococcus faecium* among bloodstream-infecting isolates in Italy. **J. Clin. Microbiol.**, v. 43, p. 1574-1580, 2005.

SU, M.; SATOLA, S.W.; READ, T.D. Genome-based prediction of bacterial antibiotic resistance. **J. Clin. Microbiol.**, v. 57, n. 3, p. e01405-e01418, 2019.

SUNDGQIST, M. et al. Within-population distribution of trimethoprim resistance in *Escherichia coli* before and after a community-wide intervention of trimethoprim use. **Antimicrob. Agents Chemother.**, v. 58, p. 7492-7500, 2014.

SUPPOLA, J. et al. *vanA* and *vanB* incorporate into an endemic ampicillin-resistant vancomycin-sensitive *Enterococcus faecium*: effect on interpretation of clonality. **J. Clin. Microbiol.**, v. 37, p. 3934-3939, 1999.

SUZUKI, M. et al. Diversity of vancomycin-resistant enterococci in a low endemicity area. **J. Glob. Antimicrob. Res.**, v. 2, n. 2, p. 115-118, 2014.

SWENSON, J.M. et al. Multilaboratory evaluation of screening methods for detection of high-level aminoglycoside resistance in enterococci. **J. Clin. Microb.**, v. 33, p. 3008-3018, 1995.

TACCONELLI, E.; CATALDO, M.A. Vancomycin-resistant enterococci (VRE): transmission and control. **Int. J. Antimicrob. Agents**, v. 31, n. 2, p. 99-106, 2008.

TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics.**, v. 123, n. 3, p. 585-595, 1989.

TEIXEIRA, L.M. et al. *Enterococcus*. In: CARROLL, K.C. et al. (org.). **Manual of Clinical Microbiology**. 12th ed. Washington, DC : ASM Press., 2019. v. 1, p. 418-435.

TEIXEIRA, N. et al. The incongruent gelatinase genotype and phenotype in *Enterococcus faecalis* due to shutting off the ability to respond to the gelatinase biosynthesis-activating pheromone (GBAP) quorum-sensing signal. **Microbiology**, v. 158, p. 519-528, 2012.

TEMIME, L. et al. Impact of capsular switch on invasive pneumococcal disease incidence in a vaccinated population. **PLoS ONE**, v. 3, n. 9, p. e3244, 2008.

TENG, F. et al. An *Enterococcus faecium* secreted antigen, SagA, exhibits broad-spectrum binding to extracellular matrix proteins and appears essential for *E. faecium* growth. **Infect. Immun.**, v. 71, p. 5033-5041, 2003.

TENOVER, P.M.; MCDONALD, L.C. Vancomycin-resistant staphylococci and enterococci epidemiology and control. **Curr. Opin. Infect. Dis.**, v. 18, p. 300-305, 2005.

TETTELIN, H. et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". **Proc. Natl. Acad. Sci. USA**, v. 102, n. 39, p. 13950-13955, 2005.

TEUBER, M.; SCHWARZ, F.; PERRETEN, V. Molecular structure and evolution of the conjugative multiresistance plasmid pRE25 *Enterococcus faecalis* isolated from a raw-fermented sausage. **Int. J. Food Microb.**, v. 88, p. 325-329, 2003.

THE UNIPROT CONSORTIUM. UniProt: a worldwide hub of protein knowledge. **Nucleic Acids Res.**, v. 47, p. D506-D515, 2019.

THOMAS, M. et al. Whole genome sequencing-based detection of antimicrobial resistance and virulence in non-typhoidal *Salmonella enterica* isolated from wildlife. **Gut. Pathogen.**, v. 9. p. 66-75, 2017.

THOMPSON, R.L.; LAVIN, B.; TALBOT, G.H. Endocarditis due to vancomycin-resistant *Enterococcus faecium* in an immunocompromised patient: cure by administering combination therapy with quinupristin/dalfopristin and high-dose ampicillin. **South Med. J.**, v. 96, n. 8, p. 818-820, 2003.

TITZE-DE-ALMEIDA, R. et al. Multilocus sequence typing of hospital-associated *Enterococcus faecium* from Brazil reveals their unique evolutionary history. **Microb. Drug Resist.**, v. 12, n. 2, p. 121, 2006.

TOCK, M.R.; DRYDEN, D.T.F. The biology of restriction and anti-restriction. **Curr. Opin. Microbiol.**, v. 8, n. 4, p. 466-472, 2005.

TOMITA, H. et al. Highly conjugative pMG1-Like plasmids carrying Tn1546-Like transposons that encode vancomycin resistance in *Enterococcus faecium*. **J. Bacteriol.**, v. 185, n. 23, p. 7024-7028, 2003.

TOMITA, H. et al. Possible connection between a widely disseminated conjugative gentamicin resistance (pMG1-Like) plasmid and the emergence of vancomycin resistance *Enterococcus faecium*. **J. Clin. Microbiol.**, v. 40, n. 9, p. 3326-3333, 2002.

TOP, J. et al. Emergence of clonal complex 17 *Enterococcus faecium* in the Netherlands. **J. Clin. Microbiol.**, v. 46, n. 1, p. 214-219; 2008.

TOP, J.; WILLEMS, R.J.L.; BONTEN, M. Emergence of CC17 *Enterococcus faecium*: from comensal to hospital-adapted pathogen. **FEMS Immunol. Med. Microbiol.**, v. 52, n. 3, p. 297-308, 2008.

- TORELL, E. et al. Clonality among ampicillin-resistant *Enterococcus faecium* isolates in Sweden and relationship with ciprofloxacin resistance. **Clin. Microb. Infect.**, v.9, p. 1011-1019, 2003.
- TOUCHON, M.; ROCHA, E.P.C. Causes of insertion sequences abundance in prokaryotic genomes. **Mol. Biol. Evol.**, v. 24, p. 969-981, 2007.
- TREMBLAY, C.L. et al. Characterization of hospital-associated lineages of ampicillin-resistant *Enterococcus faecium* from clinical cases in dogs and humans. **Front. Microbiol.**, v. 4, n. 245, 2013. DOI: 10.3389/fmicb.2013.00245.
- TURNER, K.M.E. et al. Assessing the reliability of eBURST using simulated populations with known ancestry. **BMC Microbiol.**, v. 7, n. 30, 2007. DOI: <https://doi.org/10.1186/1471-2180-7-30>.
- TUTAR, Y. Pseudogenes. **Comp. Funct. Genomics**, v. 2012, p. 424526, 2012.
- TYSON, G.H. et al. WGS accurately predicts antimicrobial resistance in *Escherichia Coli*. **Antimicrob. Agents Chemother.**, v. 70, p. 2763-2769, 2015.
- ULRICH, N.; VONBERG, R.P.; GASTMEIER, P. Outbreaks caused by vancomycin-resistant *Enterococcus faecium* in hematology and oncology departments: A systematic review. **Heliyon**, v. 3, n. 12, p. e00473, 2017.
- URUSHIBARA, N. et al. Contribution of type II topoisomerase mutations to fluoroquinolone resistance in *Enterococcus faecium* from Japanese clinical setting. **Microb. Drug Resist.**, v. 24, n. 1, p. 1-7, 2018.
- VAKULENKO, S.B. et al. Multiplex PCR for detection of aminoglycoside resistance genes in enterococci. **Antimicrob. Agents Chemother.**, v. 47, p.1423–1426, 2003.
- VALDEZATE, S. et al. Clonal outbreak of ST17 multidrug-resistant *Enterococcus faecium* harbouring an Inc18-like::Tn1546 plasmid in a haemo-oncology ward of Spanish hospital. **J. Antimicrob. Chemother.**, v. 67, n. 4, p. 832-836, 2012.
- VAN ERT, M. et al. Global genetic population structure of *Bacillus anthracis*. **PLoS ONE**, v. 2, n. 5, p. e461, 2007.
- VAN HAL, S.J. et al. Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. **Microb. Genom.**, v. 2, n. 1, 2016. DOI: 10.1099/mgen.0.000048.
- VAN HAL, S.J. Relentless spread and adaptation of non-typeable *vanA* vancomycin-resistant *Enterococcus faecium*: a genome-wide investigation. **J. Antimicrob. Chemother.**, v. 73, n. 6, p. 1487-1491, 2018.
- VAN SCHAIK, W. et al. Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. **BMC Genomics**, v. 11, n. 239, 2010. DOI: <https://doi.org/10.1186/1471-2164-11-239>.

VAN TYNE, D.; GILMORE, M.S. Friend turned foe: Evolution of enterococcal virulence and antibiotic Resistance. **Annu. Rev. Microbiol.**, v. 68, p. 337–356, 2014.

VANKERCKHOVEN, V. et al. Development of a multiplex PCR for the detection of *asaI*, *gelE*, *cylA*, *esp*, and *hyl* genes in enterococci and survey for virulence determinants among European hospital isolates of *Enterococcus faecium*. **J. Clin. Microbiol.**, v. 42, p. 4473-4479, 2004.

VIGNAROLI, C. et al. Antibiotic and heavy metal resistance in enterococci from coastal marine sediment. **Environ. Pollut.**, v. 237, p. 406-413, 2018.

VILLA T.G. et al. Horizontal gene transfer in bacteria, an overview of the mechanisms involved. In: VILLA, T.; VIÑAS, M. (ed.) **Horizontal Gene Transfer**. [S. l.]: Springer, 2019.

VOGWILL, T.; MACLEAN, R.C. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. **Evol. Appl.**, v. 8, n. 3, p. 284-295, 2015.

VOS, M. Why do bacteria engage in homologous recombination? **Trends Microbiol.**, v. 17, n. 6, p. 226-232, 2009.

VOS, M.; DIDELOT, X. A comparison of homologous recombination rates in bacteria and archaea. **ISME J.**, v. 3, p. 199-208, 2009.

WAGNER, A.; LA CHAUX, N. Distant horizontal gene transfer is rare for multiple families of prokaryotic insertion sequences. **Mol. Genet. Genomics**, v. 280, p. 397-408, 2008.

WAGNER, A.; LEWIS, C.; BICHSEL, M. A survey of bacterial insertion sequences using IScan. **Nucleic Acid Res.**, v. 35, p. 5284-5293, 2007.

WAKSMAN, S.A.; WOODRUFF, H.B. Streptothricin, a new selective bacteriostatic and bactericidal agent, particularly active against Gram-Negative bacteria. **Exp. Biol. Med.**, v. 49, n. 2, p. 207-210, 1942.

WANG, G. et al. Evolution and mutation predisposing to daptomycin resistance in vancomycin-resistant *Enterococcus faecium* ST736 strains. **PLoS ONE**, v. 13, n. 12, p. e0209785, 2018.

WANG, S. et al. Characteristic of *Enterococcus faecium* clinical isolates with quinupristin/dalfopristin resistance in China. **BMC Microbiol.**, v. 16, n. 246, 2016. DOI: <https://doi.org/10.1186/s12866-016-0863-8>.

WARDAL, E. et al. Diversity of plasmids and Tn1546-type transposons among *vanA* *Enterococcus faecium* in Poland. **Eur. J. Clin. Microbiol.**, v. 36, p. 313-328, 2017.

WARDENBURG, K.E. et al. Phenotypic and genotypic characterization of linezolid-resistant *Enterococcus faecium* from the USA and Pakistan. **J. Antimicrob. Chemother.**, v. 74, p. 3445-3452, 2019.

WASS, M.N.; KELLEY, L.A.; STERNBERG, M.J.E. 3DLigandSite: predicting ligand-binding sites using similar structures. **Nucleic Acids Res.**, v. 38, p. W469-W473, 2010.

WASSILEW, N. et al. Outbreak of vancomycin-resistant *Enterococcus faecium* clone ST796, Switzerland, December 2017 to April 2019. **Euro Surveill.**, v. 23, n. 29, p. 1800351, 2018.

WATANABE, S. et al. Genetic diversity of enterococci harboring the high-level gentamicin resistance gene *aac(6')-Ie-aph(2'')*-Ia or *aph(2'')*-Ie in Japanese hospital. **Microb. Drug Resist.**, v. 15, n. 3, p. 185-194, 2009.

WATTAM, A.R. et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. **Nucleic Acids Res.**, v. 44, n. D1, p. D535-D542, 2017.

WEAVER, K.E. et al. The RepA_N replicons of Gram-positive bacteria: A family of broadly distributed but narrow host range plasmids. **Plasmid**, v. 61, n. 2, p. 94-109, 2009.

WEIDENMAIER, C.; PESCHEL, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. **Nat. Rev. Microbiol.**, v. 6, p. 276-287, 2008.

WEINER, L.M. et al. Antimicrobial-resistant pathogens associated with healthcare-associated infections: Summary of data reported to the National Healthcare safety network at the Centers for Disease Control and Prevention, 2011-2014. **Infect. Control. Hosp. Epidemiol.**, v. 37, n. 11, p. 1288-1301, 2016.

WERNER, G. et al. Emergence and spread of vancomycin resistance among enterococci in Europe. **Euro Surveill.**, v. 13, p. 1-11, 2008.

WERNER, G. et al. High-level ciprofloxacin resistance among hospital-adapted *Enterococcus faecium* (CC17). **Int. J. Antimicrob. Agents**, v. 35, p. 119-125, 2010.

WERNER, G.; HILDEBRANDT, B.; WITTE, W. Linkage of *erm(B)* and *aadE-sat4-aphA-3* in multiple-resistant *Enterococcus faecium* isolates of different ecological origins. **Microb. Drug Resist.**, v. 9, n. 1, p. 9-16, 2003.

WERNER, G.; KLARE, I.; WITTE, W. The current MLVA typing scheme for *Enterococcus faecium* is less discriminatory than MLST and PFGE for epidemic-virulent, hospital-adapted clonal types. **BMC Microbiol.**, v. 7, n. 28, 2007. DOI: 10.1186/1471-2180-7-28.

WORLD HEALTH ORGANIZATION. **Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics.** Geneva, [2017].

Disponível em: <http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25FebET_NM_WHO>. Acesso em: 18 maio 2020.

WILLEMS, R.J.L. et al. Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. **Emerg. Infect. Dis.**, v. 11, p. 821-828, 2005.

WILLEMS, R.J.L. et al. Mutations in the DNA mismatch repair proteins MutS and MutL of oxazolidinone-resistant or -susceptible *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 47, n. 10, p. 3061-3066, 2003.

WILLEMS, R.J.L. et al. Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. **FEMS Microbiol. Rev.**, v. 35, n. 5, p. 872-900, 2011.

WILLEMS, R.J.L. Population genetics of *Enterococcus*. In: ROBINSON, D.A.; FALUSH, D.; FEIL, E.J. (ed). **Bacterial population genetics in infectious diseases**. Hoboken, NJ : John Wiley & Sons, 2010. p 195–216.

WILLEMS, R.J.L. et al. Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. **MBio**, v. 3, n. 4, p. e00151-12, 2012.

WILLEMS, R.J.L. et al. Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. **Lancet**, v. 357, p. 853-855, 2001.

WILLEMS, R.J.L.; VAN SCHAİK, W. Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. **Future Microbiol.**, v. 4, n. 9, p. 1125-1135, 2009.

WILLIAMS, A.B. Genome instability in bacteria: Causes and consequence. In: KOVALCHUK, I.; KOVALCHUK, O. (ed). **Genome Stability**. [S. l.]:Academic Press, 2016. p. 69-85.

WIRTH, T. et al. The rise and spread of a new pathogen: seroresistant *Moraxella catarrhalis*. **Genome Res.**, v. 17, p. 1647-1656, 2007.

WISELL, K.T.; KAHLMETER, G.; GISKE, C.G. Trimethoprim and enterococci in urinary tract infections: new perspectives on an old issue. **J. Antimicrob. Chemother.**, v. 62, n. 1, p. 35-40, 2008.

WITTE, W. et al. Antibiotics in animal feed. **Acta Vet. Scand.**, v. 93, p. 37-45, 2000.

WOJCIECHOWSKI, D.; CHANDRAN, S. Effect of ciprofloxacin combined with sulfamethoxazole-trimethoprim prophylaxis on the incidence of urinary tract infections after kidney transplantation. **Transplantation**, v. 96, n. 4, p. 400-405, 2013.

WOLF, Y.I.; KOONIN, E.V. Genome reduction as the dominant mode of evolution. **Bioessays**, v. 35, n. 9, p. 829-837, 2013.

WOLF, Y.I. et al. Genome trees and the tree of life. **Trends Genet.**, v. 18, p. 472-479, 2002.

WOODFORD, N. et al. Comparison of high-level gentamicin-resistant *Enterococcus faecium* isolates from different continents. **Antimicrob. Agents Chemother.**, v. 37, n. 4, p. 681-684, 1993.

WRIGHT, G.D. The antibiotic resistome. **Expert Opin. Drug Discov.**, v. 5, n. 8, p. 779-788, 2010.

WRIGHT, G.D.; POINAR, H. Antibiotic resistance in ancient: implications for drug discovery. **Trends Microbiol.**, v. 20, p. 157-159, 2012.

XANTHOPOULOU, K. et al. Vancomycin-resistant *Enterococcus faecium* colonizing patients on hospital admission in Germany: prevalence and molecular epidemiology. **J. Antimicrob. Chemother.**, v. 75, n. 10, p. 2743-2751, 2020.

XU, H.T. et al. Nosocomial spread of hospital-adapted CC17 vancomycin-resistant *Enterococcus faecium* in a tertiary-care hospital of Beijing, China. **Chin. Med. J.**, v. 124, n. 4, p. 489-503, 2011.

XU, X. et al. *vanM*, a new glycopeptide resistance gene cluster found in *Enterococcus faecium*. **Antimicrob. Agents Chemother.**, v. 54, p. 4643-4647, 2010.

YAHIA, H.B. et al. Antimicrobial resistance and genetic lineages of faecal enterococci of wild birds: Emergence of *vanA* and *vanB2* harbouring *Enterococcus faecalis*. **Int. J. Antimicrob. Agents**, v. 52, n. 6, p. 936-941, 2018.

YAMAGUCHI, Y.; PARK, J.H.; INOUE, M. Toxin-antitoxin systems in bacteria and archaea. **Annu. Rev. Genet.**, v. 45, p. 61-79, 2011.

YANG, J. et al. Prevalence of diverse clones of vancomycin-resistant *Enterococcus faecium* ST78 in a Chinese Hospital. **Microb. Drug Resist.**, v. 22, n. 4, p. 294-300, 2016.

YANG, J. et al. Selected mutation by nemonoxacin and fluoroquinolone exposure among relevant gram-positive bacterial strains in Taiwan. **Microb. Drug Resist.**, v. 26, n. 2, p. 110-117, 2020.

YU, Z. et al. Antimicrobial resistance and its association with tolerance to heavy metals in agriculture production. **Food Microbiol.**, v. 64, n. 23-32, 2017.

ZAHEER, R. et al., Surveillance of *Enterococcus* spp. reveals distinct species and antimicrobial resistance diversity across a one-health continuum. **Sci. Rep.**, v. 10, n. 1, p. 3937, 2020.

ZANELLA, R.C. et al. First confirmed case of a vancomycin-resistant *Enterococcus faecium* with VanA phenotype from Brazil: isolation from a meningitis case in São Paulo. **Microb. Drug Resist.**, v. 5, p. 159-161, 1999.

ZANELLA, R.C. et al. Phenotypic and genotypic characterization of *vanA* *Enterococcus* isolated during the first nosocomial outbreak in Brazil. **Microb. Drug Resist.**, v. 9, p. 283-291, 2003.

ZANKARI, E. et al. Identification of acquired antimicrobial resistance genes. **J. Antimicrob. Chemother.**, v. 67, n. 11, p. 2640-2644, 2012.

ZENG, D. et al. Approved glycopeptide antibacterial drugs: Mechanism of action and resistance. **Cold Spring Harb. Perspect. Med.**, v. 6, n. 12, p. a026989, 2016.

ZERVOS, M.J.; SCHABERG, D.R. Reversal of the in vitro susceptibility of enterococci to trimethoprim-sulfamethoxazole by folinic acid. **Antimicrob. Agents Chemother.**, v. 28, p. 446-448, 1985.

ZHENG, B. et al. Molecular characterization of vancomycin-resistant *Enterococcus faecium* isolates from mainland China. **J. Clin. Microbiol.**, v. 45, p. 2813-2818, 2007.

ZHONG, Z. et al. Comparative genomic analysis of the genus *Enterococcus*. **Microbiological Res.**, v. 196, p. 95-105, 2017.

ZHONG, Z. et al. Comparative genomic analysis revealed great plasticity and environmental adaptation of the genomes of *Enterococcus faecium*. **BMC Genomics.**, v. 20, n. 602, 2019.
DOI: <https://doi.org/10.1186/s12864-019-5975-8>.

ZHU, X. et al. Molecular characterization of outbreak-related strains of vancomycin-resistant *Enterococcus faecium* from an intensive care unit in Beijing, China. **J. Hosp. Infect.**, v. 72, n. 2, p. 147-154, 2009.

ZIELENKIEWICZ, U.; CEGLOWSKI, P. The toxin-antitoxin system of the streptococcal plasmid pSM19035. **J. Bacteriol.**, v. 187, p. 6094-6105, 2005.

APÊNDICE - Dados e Métricas do Genoma de 74 Amostras de *Enterococcus faecium*

Amostra	Ano	Hospital	ST	N. Reads	Tamanho das Reads	Cobertura	Contigs	N50	G+C	tRNA	rRNA	CDS	Tamanho do Genoma
CL6583	2004	XVI	78	1.295.445	125	55x	197	39,854	37,7	48	3	2799	2,8
CL6596	2004	XVII	78	248.499	125	10x	370	14,769	37,8	33	3	2582	2,7
CL6670	2004	IV	78	1.437.200	125	61x	189	37,451	37,7	48	3	2811	2,8
CL6684	2004	I	78	778.43	125	33x	180	39,855	37,7	50	3	2805	2,9
CL6729	2005	I	78	1.429.426	125	61x	222	35,965	37,7	47	3	2804	2,8
CL6787	2005	XVIII	78	868.493	151	44x	197	50,067	37,7	62	8	2792	2,8
CL6900	2005	I	78	1.953.414	126	84x	272	21,507	37,4	19	3	2747	2,8
CL6941	2005	VI	78	977.868	151	50x	188	55,414	37,7	61	8	2794	2,8
CL6959	2005	I	78	949.552	151	49x	186	51,559	37,4	62	8	2845	2,9
CL7045	2006	I	78	1.061.343	151	54x	190	55,414	37,7	64	8	2771	2,8
CL7087a	2006	VII	78	575.387	151	29x	191	36,097	38	60	7	2514	2,6
CL7093	2006	XIX	78	258.848	151	13x	169	49,709	37,8	56	8	2697	2,7
CL7133	2006	XX	78	1.676.204	126	72x	127	56,501	38	48	3	2498	2,6
CL7153	2006	XXI	78	655.772	151	33x	190	49,709	37,7	63	8	2794	2,8
CL7197	2006	IV	78	152.177	151	7x	138	34,074	37,6	39	7	1745	1,8
CL7435	2007	I	78	935.527	151	48x	193	51,559	37,7	61	8	2763	2,8
CL7448	2007	VII	78	434.801	151	22x	175	49,893	37,8	59	8	2669	2,7
CL7611	2007	XXII	78	919.888	151	47x	227	41,041	37,7	64	7	2845	2,9
CL7720b	2007	XXIII	78	747.849	151	38x	196	49,461	37,8	62	8	2685	2,8
CL8055	2008	NI	78	441.263	151	22x	244	32,622	37,8	59	8	2718	2,8
CL8183	2008	IX	412	494.663	151	25x	227	31,158	37,7	58	6	2761	2,8
CL8216	2008	I	412	112.009	151	5x	389	18,462	37,7	62	8	2752	2,9

Amostra	Ano	Hospital	ST	N. Reads	Tamanho das Reads	Cobertura	Contigs	N50	G+C	tRNA	rRNA	CDS	Tamanho do Genoma
CL8226	2009	II	192	1.366.242	126	58x	222	44,996	37,7	46	3	2733	2,8
CL8267	2009	III	412	1.218.363	126	52x	191	47,273	37,7	50	3	2777	2,8
CL8308	2009	III	412	478.038	151	24x	198	43,451	37,7	63	5	2781	2,8
CL8324	2009	V	412	1.366.089	126	58x	196	47,274	37,7	50	3	2773	2,8
CL8439	2009	II	192	1.047.782	126	45x	219	41,418	37,7	57	3	2734	2,8
CL8450	2009	III	412	963.850	126	41x	198	39,222	37,7	57	3	2777	2,8
CL8476	2009	IV	78	1.153.551	126	49x	164	38,84	37,8	55	3	2694	2,7
CL8527	2009	I	412	1.768.966	126	76x	194	49,769	37,7	58	3	2801	2,9
CL8579	2009	VII	1626	1.210.602	126	52x	185	34,979	37,8	49	3	2666	2,7
CL8655	2009	I	412	840.997	151	43x	209	50,137	37,7	60	8	2811	2,9
CL8759A	2010	III	412	914.531	151	47x	181	50,931	37,7	58	8	2744	2,8
CL8780	2010	X	412	898.765	151	46x	181	49,749	37,7	56	7	2742	2,8
CL8790	2010	II	1628	1.415.593	126	61x	76	124,315	38	57	3	2492	2,5
CL8793	2010	X	412	1.194.469	126	51x	227	36,213	37,7	50	3	2740	2,8
CL8924	2011	IX	963	1.508.367	126	65x	194	42,436	37,6	55	3	2778	2,8
CL8986	2011	IV	412	1.396.509	126	60x	208	32,633	37,7	56	3	2797	2,8
CL8988	2011	I	203	1.287.743	126	55x	180	40,989	37,7	50	3	2800	2,9
CL9162	2011	II	412	1.310.312	126	56x	215	36,681	37,7	54	3	2809	2,8
CL9166	2011	II	412	955.012	126	41x	237	31,011	37,7	52	3	2835	2,9
CL9218	2012	II	412	1.201.432	126	51x	209	32,633	37,7	58	3	2792	2,8
CL9263b	2012	VI	412	193.735	151	9x	229	29,492	37,7	54	8	2848	2,9
CL9313	2013	XXIV	963	143.943	151	7x	174	31,086	37,8	54	8	2168	2,2
CL9361	2013	V	963	1.205.229	126	52x	190	37,234	37,7	57	3	2678	2,8
CL9373	2013	II	885	1.341.255	126	57x	111	66,275	37,9	59	3	2595	2,6

Amostra	Ano	Hospital	ST	N. Reads	Tamanho das Reads	Cobertura	Contigs	N50	G+C	tRNA	rRNA	CDS	Tamanho do Genoma
CL9386	2013	V	963	1.113.632	126	48x	183	45,687	37,8	49	3	2695	2,8
CL9388	2013	V	844	1.147.555	126	49x	213	50,121	37,8	56	3	2742	2,8
CL9405	2013	XI	844	2.253.450	126	97x	226	53,968	37,8	49	3	2755	2,8
CL9406	2013	XI	844	1.178.088	126	50x	238	50,121	37,8	49	3	2758	2,8
CL9458	2013	II	253	2.595.203	126	111x	74	141,77	37,9	52	8	2633	2,7
CL9540	2014	VI	963	1.214.543	126	52x	199	35,057	37,8	50	3	2649	2,7
CL9583	2014	II	963	1.235.153	126	53x	229	33,209	37,8	48	3	2713	2,8
CL9725	2015	VIII	ST novo	405.179	151	20x	191	44,634	37,8	59	7	2702	2,8
CL9734	2015	XII	412	761.332	151	39x	217	31,158	37,8	63	7	2799	2,8
CL9735	2015	XXV	896	464.093	151	23x	216	31,208	38	60	8	2729	2,8
CL9747	2015	XIII	963	676.526	151	34x	199	38,096	37,7	59	7	2792	2,8
CL9756	2015	XXVI	896	1.785.588	126	77x	229	30,465	37,8	53	3	2752	2,8
CL9759	2015	XXVII	412	738.992	151	38x	228	30,561	37,7	61	8	2804	2,8
CL9762	2015	XV	963	1.036.570	151	53x	252	34,094	37,8	55	8	2712	2,8
CL9785	2015	XII	963	1.357.328	151	70x	213	33,984	37,8	60	7	2741	2,8
CL9789	2015	XXVIII	896	724.498	151	37x	250	31,204	37,7	51	8	2800	2,9
CL9800	2015	XIV	896	510.815	151	26x	221	31,707	37,8	58	6	2778	2,8
CL9804	2015	XIV	896	1.137.385	151	58x	276	29,962	37,9	58	8	2749	2,8
CL9826	2015	VIII	ST novo	1.384.655	126	59x	198	42,868	37,8	57	3	2693	2,8
CL9839	2015	II	97	1.555.869	126	67x	72	75,318	38,1	60	3	2332	2,4
CL9848	2015	XXIX	896	1.589.348	126	68x	230	31,547	37,8	53	3	2771	2,8
CL9856	2015	VIII	896	618.350	151	31x	222	28,841	37,9	57	7	2733	2,8
CL9882	2015	XV	896	1.546.628	126	66x	233	31,548	37,8	51	3	2761	2,8
CL9947	2016	XIII	963	1.252.566	126	54x	207	32,443	37,8	56	3	2654	2,7

Amostra	Ano	Hospital	ST	N. Reads	Tamanho das Reads	Cobertura	Contigs	N50	G+C	tRNA	rRNA	CDS	Tamanho do Genoma
CL10055	2016	XXX	963	1.168.081	126	50x	231	35,057	37,8	57	3	2697	2,8
CL10189	2016	XXXI	963	1.139.342	126	49x	240	32,44	37,8	53	3	2697	2,8
CL10395	2016	XV	896	1.278.745	126	55x	233	30,465	37,8	56	3	2820	2,9
CL10484	2016	III	412	1.034.019	126	44x	244	30,465	37,6	58	3	2909	2,9

Legenda: *sequence type* (ST); número (N); *coding sequence* (CDS).