

DESENVOLVIMENTO DE SEQUENCIADOR GENÉTICO PARA IDENTIFICAÇÃO DE SUPERBACTÉRIAS HOSPITALARES UTILIZANDO TÉCNICAS DE REDUÇÃO DIMENSIONAL E AGRUPAMENTO SEGMENTADO

Jailson Rodrigues da Silva ¹, Júlia Helena Ortiz ²

¹ *Docente de Licenciatura em Física no IFSC – Campus Jaraguá do Sul/SC, Jaraguá do Sul/SC, Brasil (jailson.silva@estudante.ifb.edu.br)*

² *Professora Doutora (USP). Professora Substituta (IFSC), Jaraguá do Sul, Santa Catarina, Brasil*

Resumo: O aumento das infecções causadas por superbactérias resistentes a múltiplos antibióticos é uma das maiores preocupações de saúde pública mundial. Comum em ambientes hospitalares, essas bactérias são responsáveis por uma alta taxa de doenças e mortes, representando um grande desafio para profissionais de saúde. O diagnóstico rápido e preciso dessas infecções é crucial para o tratamento eficaz e controle de surtos hospitalares.

O sequenciamento genético de nova geração (NGS) surge como uma ferramenta promissora para identificar patógenos e suas resistências. No entanto, os dados gerados por essas tecnologias são volumosos e complexos, exigindo técnicas avançadas de análise para extrair informações úteis para o diagnóstico clínico.

Este trabalho propõe o desenvolvimento de um motor de sequenciamento genético que, por meio de métodos de redução dimensional e agrupamento segmentado, visa identificar superbactérias hospitalares de maneira mais rápida e precisa. O objetivo é otimizar o diagnóstico e controle de infecções resistentes em ambientes hospitalares.

Essa reformulação mantém o foco no problema das infecções causadas por superbactérias e nas soluções propostas, com destaque para o uso de NGS e técnicas de análise de dados para melhorar o diagnóstico e controle de infecções hospitalares.

Palavras-chave: Superbactérias; Infecções hospitalares; Diagnóstico rápido; Análise de dados

INTRODUÇÃO

A resistência antimicrobiana (RAM) é um grande desafio de saúde pública, com infecções causadas por superbactérias resistentes a antibióticos, como MRSA, VRE e ESBL, representando uma preocupação crescente (WHO, 2021). O uso excessivo de antibióticos acelera o desenvolvimento dessas cepas resistentes, especialmente em ambientes hospitalares. O diagnóstico precoce é essencial, mas métodos tradicionais, como o cultivo bacteriano, têm limitações em termos de tempo e na detecção de resistência em tempo real (Choi & Ko, 2020). O sequenciamento genético de nova geração (NGS) tem revolucionado o diagnóstico, permitindo a análise detalhada do genoma das bactérias e a identificação de genes de resistência (Chiu & Miller, 2019). O NGS oferece resultados rápidos e precisos,

auxiliando na escolha de tratamentos eficazes (Wilson et al., 2019). No entanto, a grande quantidade de dados gerados exige técnicas avançadas de análise, como o aprendizado de máquina (ML), que usa modelos matemáticos para identificar padrões de resistência e virulência (LeCun et al., 2015; Zhang et al., 2016). Modelos como redes neurais e SVM podem ser treinados para prever resistência em cepas bacterianas (Ching et al., 2018).

Python é uma ferramenta popular para implementar essas soluções, com bibliotecas como Scikit-learn, TensorFlow e Keras, facilitando a análise de dados genômicos e a integração de informações clínicas (Vanderplas, 2016; Pedregosa et al., 2011). Isso possibilita a criação de sistemas de suporte à decisão

para prever surtos de infecções, otimizar estratégias de controle e personalizar tratamentos (Harris et al., 2020).

Embora o uso de NGS e ML ofereça grandes avanços, desafios como a necessidade de grandes volumes de dados rotulados e a infraestrutura computacional necessária precisam ser superados (Goldstein et al., 2015). No entanto, essas tecnologias representam uma abordagem inovadora para o diagnóstico e controle das infecções resistentes, ajudando a combater a RAM de forma mais eficaz (Zhang et al., 2016).

MATERIAL E MÉTODOS

A pesquisa proposta visa a utilização de uma base de dados público fornecida na plataforma Kaggle no link <https://www.kaggle.com/competitions/tabular-playground-series-feb-2022/data> cujo o tratamento aplicado foi da espectroscopia Raman, combinada com técnicas de aprendizado de máquina, para a análise de infecções bacterianas, especificamente no diagnóstico de superbactérias resistentes a múltiplos antibióticos. A metodologia adotada para este estudo envolve três etapas principais: coleta de dados espectroscópicos de amostras de DNA, análise dos dados utilizando técnicas avançadas de aprendizado de máquina e validação dos resultados obtidos por meio de comparação com métodos tradicionais de diagnóstico. Abaixo estão descritos em detalhes os materiais utilizados e os procedimentos metodológicos empregados para a realização do estudo.

As amostras de DNA utilizadas neste estudo foram extraídas de 10 bactérias com diversas cepas bacterianas em uma amostra de 200 mil exemplos distintos, incluindo tanto cepas sensíveis quanto resistentes a múltiplos antibióticos. As amostras utilizadas são patógenos significativos que causam uma variedade de infecções em seres humanos, variando desde doenças gastrointestinais até infecções respiratórias e complicações mais graves, como sepse e meningite. A seguir, vamos detalhar melhor essas bactérias e sua importância para a saúde pública.

A *Streptococcus pyogenes* é uma bactéria Gram-positiva do grupo A de estreptococos, responsável por diversas infecções humanas. Ela é principalmente conhecida como a causa da faringite estreptocócica, que se manifesta com dor de garganta, febre e amígdalas inflamadas, podendo levar a complicações graves, como febre reumática e glomerulonefrite pós-infecciosa (Weber, 2014). Além disso, *S. pyogenes* pode causar infecções de pele, como impetigo, e

condições mais graves, como celulite e fascíte necrosante, a última das quais pode ser fatal se não tratada rapidamente (Carapetis et al., 2005). Outra condição severa associada à bactéria é a síndrome do choque tóxico estreptocócico, que pode levar à falência multissistêmica (Lau et al., 2014).

Embora a maioria das cepas de *S. pyogenes* seja sensível à penicilina, casos de resistência têm sido registrados, especialmente em cenários de uso inadequado de antibióticos. A resistência ao metilicina, embora rara, complica ainda mais o tratamento das infecções invasivas (Parker et al., 2013). A resistência antimicrobiana é uma preocupação crescente e exige vigilância constante.

Portanto, o diagnóstico precoce, o tratamento adequado e o monitoramento das cepas resistentes são essenciais para controlar a disseminação de *S. pyogenes* e prevenir complicações graves. Além disso, é fundamental implementar estratégias de educação pública sobre o uso racional de antibióticos para mitigar o risco de resistência e melhorar os resultados de saúde.



Figura 1 – Visão microscópica da bactéria *Streptococcus pyogenes*.
Fonte: Bandeira, s.d.

A *Salmonella enterica* é uma bactéria Gram-negativa responsável pela salmonelose, que causa gastroenterites com sintomas como diarreia, febre, náuseas e cólicas abdominais. A infecção é adquirida principalmente por alimentos contaminados, como carne crua, ovos e leite não pasteurizado (Díaz et al., 2015). Embora a maioria dos casos seja autolimitada, infecções graves podem ocorrer, especialmente em crianças, idosos e imunocomprometidos, evoluindo para septicemia (Crump et al., 2015; Gupta et al., 2011).

A resistência a antibióticos em *S. enterica* tem aumentado, dificultando o tratamento de infecções graves. O uso indiscriminado de antibióticos contribui para o surgimento de cepas resistentes,

tornando tratamentos como cefalosporinas e fluoroquinolonas menos eficazes (Martin et al., 2016). Isso é especialmente preocupante em ambientes hospitalares, onde as infecções podem ser mais difíceis de controlar.

A prevenção inclui práticas adequadas de higiene alimentar, como o cozimento adequado de carnes, a pasteurização de laticínios e o armazenamento correto de alimentos. Também é essencial o monitoramento da resistência antimicrobiana e o uso racional de antibióticos.

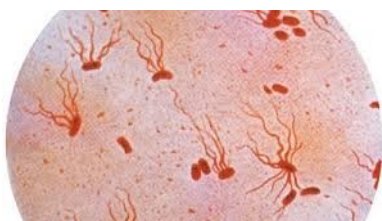


Figura 2 – Visão microscópica da bactéria *Salmonella enterica*.
Fonte: Rogers, Kara. s.d

A *Enterococcus hirae* é uma bactéria Gram-positiva que faz parte da microbiota intestinal normal, mas pode se tornar patogênica em condições de imunossupressão, como em pacientes com doenças crônicas ou após tratamentos com quimioterapia ou transplantes (Hollenbeck & Rice, 2012). Quando isso ocorre, pode causar infecções do trato urinário, endocardite e infecções em feridas. Essas infecções são especialmente problemáticas em ambientes hospitalares, onde pacientes estão mais vulneráveis devido ao uso de dispositivos invasivos e antibióticos. A resistência crescente à vancomicina, um antibiótico de última linha, dificulta o tratamento de infecções nosocomiais causadas por *E. hirae* (Mayer et al., 2017). O uso indiscriminado de antibióticos, especialmente em unidades de terapia intensiva, contribui para a resistência e torna as infecções mais difíceis de erradicar (Pérez et al., 2015).

Além disso, a capacidade de *E. hirae* de formar biofilmes, que protegem a bactéria contra antibióticos e o sistema imunológico, torna a erradicação das infecções ainda mais desafiadora (Mohammad et al., 2017). Portanto, o controle da resistência,

especialmente à vancomicina, é uma prioridade para a medicina hospitalar, exigindo vigilância constante, uso prudente de antibióticos e desenvolvimento de novas terapias.

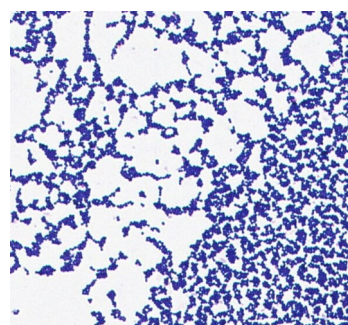


Figura 3 – Visão microscópica da bactéria *Enterococcus hirae*
Fonte: GMBANK. s.d

A *Escherichia coli* (*E. coli*) é uma bactéria Gram-negativa, parte da microbiota intestinal normal de humanos e outros mamíferos. Embora a maioria das cepas seja comensal, algumas cepas patogênicas, como *E. coli* O157:H7, podem causar doenças graves, como infecções gastrointestinais que levam a diarreia sanguinolenta e síndrome hemolítico-urêmica (SHU), com destruição de glóbulos vermelhos e insuficiência renal aguda (Karmali et al., 2003). Essa cepa é associada ao consumo de alimentos contaminados, como carne mal cozida, leite não pasteurizado e vegetais mal lavados. Além das infecções intestinais, *E. coli* é uma das principais causadoras de infecções do trato urinário (ITU), especialmente em mulheres, e pode também causar infecções respiratórias e meningite em neonatos e pacientes imunocomprometidos (Bingen et al., 2017).

A resistência crescente a antibióticos em algumas cepas de *E. coli* tem gerado preocupações, especialmente com a resistência a betalactâmicos e quinolonas, dificultando o tratamento das infecções e aumentando o risco de complicações (Martins et al., 2013). O uso indiscriminado de antibióticos em humanos e na agricultura tem contribuído para a seleção de cepas resistentes.

O controle rigoroso da resistência antimicrobiana, o uso racional de antibióticos e a educação pública sobre higiene alimentar são fundamentais para combater a disseminação de cepas patogênicas e resistentes, especialmente em ambientes hospitalares.

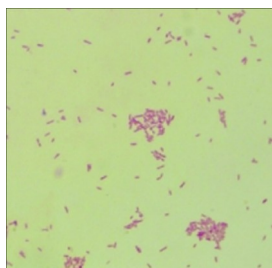


Figura 4 – Visão microscópica da bactéria
Escherichia coli

Fonte: ResearchGate.s.d

A *Campylobacter jejuni* é uma bactéria Gram-negativa, uma das principais causas de gastroenterite bacteriana, sendo responsável por muitas infecções alimentares ao redor do mundo. A transmissão ocorre principalmente por meio do consumo de alimentos contaminados, como carne de frango crua ou malcozida, leite não pasteurizado e água contaminada. A infecção resulta em sintomas como diarreia (frequentemente sanguinolenta), dor abdominal e febre, geralmente entre dois a cinco dias após a ingestão (Scallan et al., 2011).

Embora a infecção seja geralmente autolimitada, casos graves podem ocorrer, com uma complicação rara, mas séria, sendo a síndrome de Guillain-Barré (GBS), uma doença autoimune que pode levar à paralisia (Nachamkin et al., 2008). Infecções graves ou prolongadas podem necessitar de antibióticos, como azitromicina, embora o uso indiscriminado de antibióticos seja desencorajado devido ao risco de resistência (Mégraud et al., 2013).

A prevenção envolve práticas adequadas de higiene alimentar, como cozinhar bem carnes de frango, evitar o consumo de leite não pasteurizado e garantir a potabilidade da água. Além disso, a vigilância de infecções e o uso racional de antibióticos são cruciais para reduzir a resistência bacteriana.

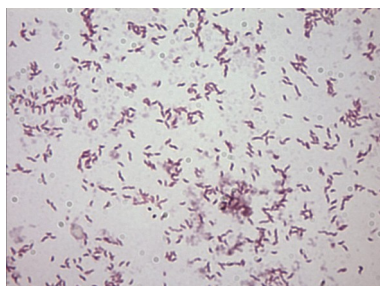


Figura 5 – Visão microscópica da bactéria
Campylobacter jejuni

Fonte:Stepwards. s.d

Streptococcus pneumoniae é uma bactéria Gram-positiva responsável por infecções respiratórias graves, como pneumonia, meningite, otite média e sinusite, além de causar septicemia. A bactéria afeta principalmente crianças pequenas, idosos e indivíduos com comorbidades, como doenças pulmonares crônicas e diabetes (Mandell et al., 2015). A pneumonia pneumocócica pode ser particularmente grave, com complicações respiratórias e altas taxas de mortalidade. A meningite, por sua vez, pode levar a sequelas neurológicas, como surdez e déficits cognitivos (Tunkel et al., 2008).

A resistência antimicrobiana em ***S. pneumoniae*** tem aumentado, dificultando o tratamento das infecções. Cepas resistentes à penicilina, cefalosporinas e macrolídeos têm sido amplamente documentadas, em grande parte devido ao uso excessivo e inadequado de antibióticos (Zeigler et al., 2014). O aumento da resistência tem levado ao uso de antibióticos de segunda linha, como levofloxacina ou ceftriaxona, para tratar infecções graves.

A vacinação com a vacina conjugada 13-valente (PCV13) tem mostrado ser uma medida eficaz na redução das infecções pneumocócicas, especialmente em crianças e idosos (Grijalva et al., 2015). O controle da resistência antimicrobiana requer o uso prudente de antibióticos, campanhas de vacinação e vigilância constante para monitorar novas cepas resistentes.

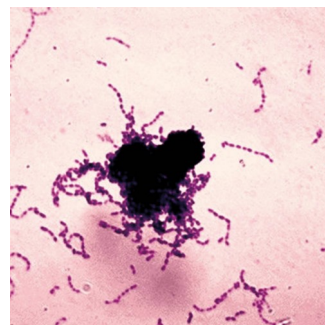


Figura 6 – Visão microscópica da bactéria
Streptococcus pneumoniae

Fonte:McGuigan, Roisin. s.d

A *Staphylococcus aureus* é uma bactéria Gram-positiva que possui uma grande capacidade de causar uma variedade de infecções, que podem variar de condições relativamente simples e superficiais, como furúnculos e celulite, até doenças graves e potencialmente fatais, como endocardite, pneumonia e sepse. Esse patógeno é um dos principais responsáveis por infecções nosocomiais (infecções adquiridas em ambientes hospitalares), especialmente devido à sua habilidade de colonizar e infectar tecidos humanos, incluindo feridas cirúrgicas e locais com dispositivos invasivos, como cateteres e ventiladores mecânicos (Diekema *et al.*, 2008).

A *Staphylococcus aureus* é bem conhecido pela sua habilidade de desenvolver resistência a múltiplos antibióticos, o que complica significativamente o tratamento das infecções que ele causa. A resistência à meticilina, um antibiótico da classe dos betalactâmicos, resultou no surgimento da forma resistente à meticilina de *Staphylococcus aureus* (MRSA, do inglês "Methicillin-Resistant *Staphylococcus aureus*"). A infecção por MRSA é particularmente preocupante em ambientes hospitalares, onde pacientes com o sistema imunológico comprometido ou com feridas abertas estão mais expostos e são mais vulneráveis a complicações graves. MRSA pode causar infecções de pele e tecidos moles, mas também é responsável por infecções mais graves, como pneumonia, infecções ósseas, endocardite e sepse, que podem levar a falência múltipla de órgãos e até à morte se não tratadas de forma adequada e rápida (Otto, 2013).

A resistência a antibióticos em *S. aureus*, especialmente em MRSA, é um desafio significativo para a medicina moderna. Além da resistência à meticilina, cepas de MRSA também podem ser resistentes a outros antibióticos importantes, como clindamicina, eritromicina, tetraciclina e, em alguns casos, até vancomicina, que é frequentemente considerada a última linha de defesa contra infecções graves (Zong *et al.*, 2016). Isso faz com que o tratamento de infecções por MRSA seja mais difícil, exigindo terapias combinadas e o uso de antibióticos de segunda e terceira linha, como linezolida, daptomicina e ceftarolina.

A prevenção de infecções por *S. aureus* e MRSA envolve o controle rigoroso de infecções nosocomiais, incluindo o uso adequado de antibióticos, medidas de higiene, como lavagem frequente das mãos, e o cuidado rigoroso com feridas e dispositivos médicos invasivos. Além disso, a implementação de protocolos de vigilância e o isolamento de pacientes infectados em ambientes hospitalares são medidas cruciais para prevenir a disseminação de MRSA (Sievert *et al.*, 2013).

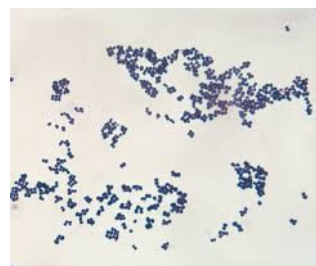


Figura 7 – Visão microscópica da bactéria *Staphylococcus aureus*
Fonte: Rogers, Kara. s.d

A *Escherichia fergusonii* é uma espécie rara dentro do gênero *Escherichia*, estreitamente relacionada a *Escherichia coli*. Embora seja menos prevalente e frequentemente encontrada na microbiota intestinal de humanos e animais, *E. fergusonii* possui potencial patogênico e pode ser responsável por infecções, especialmente em pacientes com sistemas imunológicos comprometidos ou em ambientes hospitalares. Em comparação com *E. coli*, *E. fergusonii* é menos estudada, mas tem sido identificada como uma possível causa de infecções gastrointestinais e urinárias (Baker-Austin *et al.*, 2006).

As infecções por *E. fergusonii* podem se manifestar de forma semelhante às infecções causadas por outras enterobactérias, com sintomas como diarreia, dor abdominal e febre, típicos de infecções gastrointestinais. Além disso, a bactéria também pode causar infecções do trato urinário, um tipo comum de infecção bacteriana em hospitais, especialmente em pacientes que têm cateteres urinários ou outras condições predisponentes, como diabetes ou uso prolongado de antibióticos, que favorecem a colonização e infecção por bactérias oportunistas (Janda & Abbott, 2011).

Embora *E. fergusonii* não seja uma causa comum de infecções, a sua presença em ambientes hospitalares e o seu potencial para causar doenças em pacientes imunocomprometidos exigem atenção, já que ela pode ser mais difícil de diagnosticar devido à sua raridade e semelhança com outras cepas de *Escherichia* e enterobactérias. Além disso, a resistência a antibióticos em algumas cepas de *E. fergusonii* tem sido documentada, o que representa um desafio adicional no tratamento das infecções (Janda & Abbott, 2011).

A prevenção de infecções por *E. fergusonii* envolve as mesmas práticas que são recomendadas para o controle de infecções causadas por outras enterobactérias, como medidas rigorosas de higiene

hospitalar, o uso prudente de antibióticos e a monitorização de infecções nosocomiais. Além disso, a vigilância microbiológica contínua é essencial para detectar e tratar infecções causadas por patógenos menos comuns, como *E. fergusonii*.

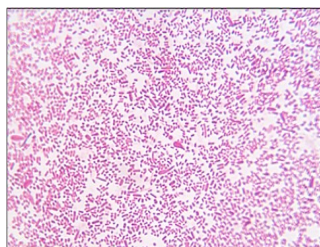


Figura 8 – Visão microscópica da bactéria *Escherichia fergusonii*
Fonte: GMBANK. s.d

A *Bacteroides fragilis* é uma bactéria anaeróbica Gram-negativa que faz parte da microbiota normal do intestino humano, onde desempenha um papel importante na digestão e na manutenção do equilíbrio microbiano. Embora geralmente não seja patogênica, *Bacteroides fragilis* pode se tornar uma causa significativa de infecções, especialmente quando ocorre translocação para locais normalmente estéreis, como a cavidade abdominal ou o trato urinário. Em tais contextos, pode estar envolvida em infecções intra-abdominais, como apendicite, peritonite e diverticulite, além de infecções urinárias, principalmente em pacientes imunocomprometidos ou com dispositivos invasivos, como cateteres (McDonald *et al.*, 2017).

A patogenicidade de *Bacteroides fragilis* está associada à sua capacidade de invadir tecidos e produzir uma variedade de fatores virulentos, como enzimas proteolíticas e cápsulas polissacarídicas, que ajudam na adesão às células do hospedeiro e na evasão do sistema imunológico. Essas características permitem que a bactéria seja uma causa comum de infecções nosocomiais e de complicações associadas a cirurgias abdominais, particularmente em pacientes com condições como doença inflamatória intestinal, que podem criar condições favoráveis para a proliferação de *Bacteroides fragilis* em locais fora do intestino (Levy *et al.*, 2015).

Uma característica preocupante de *Bacteroides fragilis* é sua resistência a múltiplos antibióticos, incluindo penicilinas e outras classes de antibióticos beta-lactâmicos. Essa resistência torna o tratamento de infecções causadas por essa bactéria mais desafiador, especialmente em infecções graves ou em pacientes com doenças complicadas. A resistência está frequentemente associada à produção de

betalactamases, enzimas que inativam os antibióticos, além de mecanismos de resistência a outros agentes antimicrobianos. Devido a essa resistência, o tratamento de infecções por *Bacteroides fragilis* geralmente envolve antibióticos específicos para anaeróbios, como metronidazol, clindamicina, carbapenêmicos ou piperacilina-tazobactam, que são mais eficazes contra essas bactérias (McDonald *et al.*, 2017).

A prevenção de infecções causadas por *Bacteroides fragilis* inclui práticas adequadas de higiene hospitalar, controle rigoroso de infecções nosocomiais, uso racional de antibióticos para evitar o desenvolvimento de resistência e cuidados específicos durante e após intervenções cirúrgicas, especialmente as que envolvem o trato gastrointestinal. Além disso, o monitoramento microbiológico é crucial para identificar cepas resistentes e garantir que o tratamento antibiótico seja eficaz.

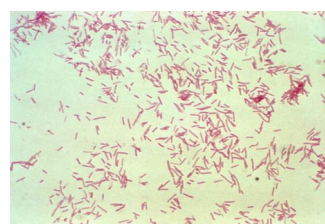


Figura 9 – Visão microscópica da bactéria *Bacteroides fragilis*
Fonte: Slonczewski, Joan. Editada pela última vez à(s) 17h54min de 5 de dezembro de 2019

A *Klebsiella pneumoniae* é uma bactéria Gram-negativa pertencente à família *Enterobacteriaceae* e é uma causa significativa de infecções respiratórias, como pneumonia, infecções do trato urinário e septicemia. Essa bactéria é uma das principais responsáveis por infecções nosocomiais (adquiridas no ambiente hospitalar), especialmente em unidades de terapia intensiva (UTI), onde pacientes imunocomprometidos ou aqueles com dispositivos médicos invasivos, como cateteres e ventiladores mecânicos, estão mais expostos ao risco de infecção. As infecções por *K. pneumoniae* podem variar de leves a graves, podendo levar a complicações severas, como insuficiência respiratória e falência de múltiplos órgãos, particularmente em pacientes com comorbidades, como diabetes, doença pulmonar obstrutiva crônica (DPOC) ou câncer (Tzouveleakis *et al.*, 2014).

Uma das maiores preocupações com *Klebsiella pneumoniae* é sua capacidade de desenvolver resistência a múltiplos antibióticos, o que dificulta o tratamento de infecções causadas por essa bactéria.

Nos últimos anos, cepas de *K. pneumoniae* resistentes a carbapenêmicos, conhecidos como "superbactérias", têm se tornado cada vez mais prevalentes. A resistência a carbapenêmicos é principalmente mediada pela produção de carbapenemases, enzimas que inativam os carbapenêmicos e outras classes de antibióticos beta-lactâmicos, tornando essas infecções extremamente difíceis de tratar com os antibióticos padrão (Nordmann *et al.*, 2011).

Além das carbapenemases, *K. pneumoniae* também pode desenvolver resistência a outras classes de antibióticos, como cefalosporinas, aminoglicosídeos e fluoroquinolonas, o que a torna um patógeno extremamente desafiador no ambiente hospitalar. A disseminação dessas cepas resistentes ocorre frequentemente por meio de contato direto com superfícies contaminadas ou entre pacientes, o que reforça a importância de práticas rigorosas de controle de infecções, como a higienização das mãos e o isolamento de pacientes infectados.

O tratamento de infecções por *Klebsiella pneumoniae* resistente a carbapenêmicos pode exigir o uso de antibióticos alternativos, como colistina, tigeciclina ou ceftazidima-avibactam, embora essas opções sejam limitadas e nem sempre eficazes. Além disso, o desenvolvimento de novas classes de antibióticos, bem como estratégias terapêuticas combinadas, é essencial para lidar com a crescente resistência a antibióticos em *K. pneumoniae* (Tzouveleakis *et al.*, 2014).

Portanto, a prevenção de infecções por *Klebsiella pneumoniae* e a luta contra a resistência antimicrobiana exigem um esforço conjunto, que inclui práticas rigorosas de controle de infecções em ambientes hospitalares, vigilância constante da resistência bacteriana e o uso racional de antibióticos para minimizar a seleção de cepas resistentes.

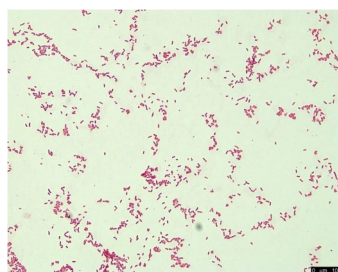


Figura 10 – Visão microscópica da bactéria *Klebsiella pneumoniae*

Fonte:

[https://microbewiki.kenyon.edu/index.php/Klebsiella pneumoniae](https://microbewiki.kenyon.edu/index.php/Klebsiella_pneumoniae)

Ainda sobre as amostras foram preparadas de acordo com os procedimentos padrão de extração de DNA de bactérias, os dados espectroscópicos obtidos foram convertidos para uma representação de histograma de bases (A, T, G, C) para cada fragmento de DNA. Essa representação corresponde à distribuição das quantidades relativas de adenina, timina, guanina e citosina nas amostras, que foi extraída da medição espectroscópica. Para cada espectro de Raman, foi gerada uma matriz de 286 valores, representando todas as possíveis combinações de contagens para os quatro nucleotídeos. Esses dados foram então corrigidos por um espectro de viés, correspondente a uma distribuição totalmente aleatória de bases, a fim de ajustar as variações experimentais.

Para aumentar a robustez do estudo, foram incorporados erros simulados nas medições espectroscópicas. Esses erros foram introduzidos variando os níveis de ruído nos espectros de forma controlada. A simulação de erros foi realizada com base em um modelo de ruído Gaussiano, o qual foi adicionado aos espectros de Raman. Esses erros simulados representaram falhas típicas que podem ocorrer durante a coleta de dados em ambientes reais de laboratório, como variações instrumentais e instabilidade nas amostras.

Após a coleta dos dados espectroscópicos, as medições foram pré-processadas para remover qualquer ruído indesejado, usando técnicas de filtragem e normalização. O ruído foi minimizado com a aplicação de um filtro de Savitzky-Golay para suavizar as flutuações nos espectros. O filtro de Savitzky-Golay é um método de suavização que aplica um polinômio sobre uma janela deslizante de dados, ajustando localmente o sinal para reduzir o impacto do ruído sem distorcer a forma dos dados originais. Esse procedimento foi realizado utilizando a função `savgol_filter` da biblioteca `scipy.signal` em Python, onde a escolha da janela e o grau do polinômio foram ajustados para garantir uma suavização eficaz, mantendo as características principais dos espectros. A janela utilizada para suavização foi configurada com um tamanho ímpar de sete pontos, e o polinômio de grau três foi escolhido para uma suavização mais detalhada, adequado para reduzir o ruído sem perder detalhes importantes.

Em seguida, os dados de histograma foram corrigidos com o espectro de viés, subtraindo-se o espectro de uma amostra aleatória, para eliminar as interferências causadas por fatores externos, como a variabilidade instrumental. A normalização foi realizada para garantir que todas as amostras tivessem a mesma escala e variação, proporcionando uma base de comparação uniforme entre as amostras. Esse procedimento assegurou que as flutuações de

intensidade, causadas por diferentes condições de medição, não interferissem nas análises posteriores. Assim, as amostras foram ajustadas para que seus dados fossem comparáveis e as características do espectro fossem mais representativas dos dados genuínos das amostras analisadas.

A análise dos dados espectroscópicos foi conduzida utilizando técnicas de aprendizado de máquina, especificamente redução dimensional e agrupamento segmentado. A redução dimensional foi realizada utilizando a técnica de Análise de Componentes Principais (PCA), que é eficaz para reduzir a complexidade dos dados sem perder informações essenciais. O PCA foi aplicado para transformar os dados espectroscópicos de alta dimensionalidade em um conjunto menor de componentes principais, facilitando a visualização e interpretação dos padrões presentes nas amostras. Em seguida, foi utilizado o algoritmo de K-means para o agrupamento segmentado. Esse algoritmo de aprendizado não supervisionado foi empregado para identificar padrões e relações entre as cepas bacterianas, agrupando as amostras com características espectroscópicas semelhantes. O número de clusters foi determinado utilizando a técnica do método de cotovelo, que avalia a soma das distâncias intra-cluster. O agrupamento foi realizado com base nos componentes principais extraídos durante a etapa de PCA.

A validação dos modelos de aprendizado de máquina foi realizada utilizando um conjunto de dados de teste, separado previamente do conjunto de dados de treinamento. O desempenho do modelo foi avaliado utilizando métricas comuns de aprendizado supervisionado, como a precisão, recall, F1-score, e curva ROC. Além disso, a confiabilidade dos resultados foi avaliada pela realização de validação cruzada k-fold, para garantir que os modelos não estivessem sobreajustados aos dados de treinamento e que fossem robustos o suficiente para generalizar em novos dados.

A análise dos dados espectroscópicos foi complementada com a análise de genes de resistência antimicrobiana (RAM). As cepas bacterianas foram classificadas com base nas suas características espectroscópicas e também nas suas informações genéticas, que indicam quais genes estão presentes nas diferentes cepas. A identificação de padrões relacionados aos genes de resistência foi realizada por meio da comparação dos espectros de Raman com bancos de dados genéticos previamente conhecidos. Esses dados foram combinados com a classificação gerada pelos algoritmos de aprendizado de máquina para obter uma visão holística das infecções e das resistências.

A combinação da espectroscopia Raman com técnicas de aprendizado de máquina oferece uma

abordagem promissora para a análise de infecções bacterianas e resistência antimicrobiana. A metodologia descrita neste estudo proporciona uma maneira eficiente de identificar e classificar diferentes cepas bacterianas com base em suas características espectroscópicas, ao mesmo tempo em que permite o diagnóstico precoce e a detecção de superbactérias resistentes a antibióticos. A utilização de técnicas como PCA e K-means, associadas ao pré-processamento adequado dos dados, possibilita uma análise robusta e precisa, mesmo com a presença de erros simulados e variabilidades experimentais.

RESULTADOS E DISCUSSÃO

A análise genômica descrita oferece uma visão detalhada sobre a distribuição e as interações de oligômeros em diversas espécies bacterianas, revelando insights sobre a complexidade das funções biológicas nas quais esses elementos estão envolvidos. O fato de padrões consistentes serem observados na distribuição dos oligômeros sugere que esses segmentos genômicos têm uma função específica e possivelmente adaptativa em cada espécie bacteriana. A regulação genética e a adaptação ao ambiente são dois dos processos biológicos essenciais que podem ser modulados por esses oligômeros, e isso se reflete diretamente nas respostas das bactérias a diferentes condições.

As altas correlações entre certos oligômeros indicam que esses elementos genéticos podem operar de maneira cooperativa em processos moleculares, como vias bioquímicas ou mecanismos regulatórios que dependem de uma interação complexa. Isso sugere a existência de redes regulatórias que ajudam a controlar o comportamento bacteriano em níveis moleculares, como a resposta a sinais ambientais ou interações com outros organismos, incluindo o hospedeiro.

Por outro lado, as correlações negativas entre determinadas combinações de oligômeros revelam um padrão de especialização adaptativa, onde a presença ou ausência de certos oligômeros pode ser crucial para a sobrevivência bacteriana em determinados nichos ecológicos. Isso pode refletir uma resposta evolutiva de "tensão competitiva" ou mecanismos de seleção natural, onde as bactérias selecionam ou perdem certos elementos genômicos dependendo do contexto ambiental e das pressões externas, como competição por recursos ou resistência a antibióticos.

Essa abordagem de análise genômica, ao identificar tanto correlações positivas quanto negativas, é uma ferramenta poderosa para desvendar os complexos mecanismos evolutivos e moleculares das comunidades bacterianas. Além disso, o entendimento dessas interações pode ser aplicado em várias áreas. Na biotecnologia, pode ser explorado para desenvolver novas ferramentas de manipulação

genética ou melhorar processos de produção biológica. Na medicina, pode ajudar no desenvolvimento de novas estratégias terapêuticas, como o uso de antibióticos mais eficazes ou terapias baseadas na modulação genética. Na ecologia microbiana, essa análise permite um entendimento mais profundo das relações entre diferentes espécies bacterianas e seus habitats.

Portanto, as informações reveladas por esse tipo de estudo não apenas aumentam o nosso entendimento fundamental sobre a biologia bacteriana, mas também oferecem possibilidades de avanços em várias áreas práticas, desde a indústria até a medicina e a preservação ambiental.

Segue a tabela com os resultados da análise de distribuição e correlação dos oligômeros nas espécies bacterianas analisadas:

Tabela 2. Correlação de Oligômeros genéticos

Tipo de bactéria	Gene 1	Gene 2	Correlação dos Oligômeros
<i>Streptococcus pneumoniae</i>	A0T0G8C2	A1T0G0C9	0.99
<i>Campylobacter jejuni</i>	A0T0G4C6	A0T0G6C4	0.99
<i>Staphylococcus aureus</i>	A0T0G3C7	A0T1G1C8	0.99
<i>Streptococcus pyogenes</i>	A0T0G5C5	A0T0G7C3	0.98
<i>Enterococcus hirae</i>	A0T0G5C5	A0T0G7C3	0.98
<i>Klebsiella pneumoniae</i>	A5T4G0C1	A5T5G0C0	0.96
<i>Escherichia fergusonii</i>	A5T4G0C1	A5T5G0C0	0.92
<i>Salmonella enterica</i>	A5T4G0C1	A5T5G0C0	0.92
<i>Escherichia coli</i>	A5T4G0C1	A5T5G0C0	0.90
<i>Bacteroides fragilis</i>	A1T2G6C1	A1T2G7C0	0.87

Fonte: Elaboração própria, 2024.

As informações fornecidas refletem correlações entre genes de diferentes espécies bacterianas, mostrando como determinados oligonucleotídeos, ou sequências de DNA, estão relacionados entre si em termos de expressão ou coocorrência. Cada valor de correlação quantifica a força da associação entre os genes em um determinado "target", que corresponde a uma espécie bacteriana específica. A correlação alta, próxima de 1, sugere que os genes estão frequentemente associados entre si nas amostras estudadas, o que pode indicar que esses genes têm funções biológicas similares ou estão localizados fisicamente próximos no genoma.

Esse tipo de análise pode ser especialmente útil para entender a interação entre genes dentro de uma mesma bactéria. Por exemplo, um par de genes com uma correlação de 0,995, como o observado entre os genes "A0T0G8C2" e "A1T0G0C9" em *Streptococcus pneumoniae*, pode indicar que esses genes possuem funções biológicas semelhantes ou estão envolvidos em processos celulares que ocorrem de maneira coordenada. Isso pode sugerir que, se um gene for ativado, o outro também será ativado, ou que ambos desempenham um papel crítico em um mesmo processo metabólico ou fisiológico.

Além disso, a identificação dessas correlações fortes pode ser fundamental para o desenvolvimento de novas abordagens terapêuticas. Ao identificar pares de genes que sempre parecem estar associados, pesquisadores podem focar seus esforços em manipular esses genes ou seus produtos para interferir em processos essenciais para a bactéria, como resistência a antibióticos ou virulência. Por exemplo, se uma bactéria tem um conjunto de genes altamente correlacionados que são críticos para sua sobrevivência ou infectividade, bloquear a expressão de um desses genes pode ter um impacto disruptivo em outros genes da mesma rede, oferecendo novas vias para o tratamento de infecções.

Portanto, a análise dessas correlações genéticas é uma ferramenta poderosa para explorar a biologia bacteriana de maneira mais profunda, ajudando a identificar novas áreas para a intervenção terapêutica e aprimorando o entendimento sobre como os genes funcionam em conjunto dentro de organismos patogênicos. Além disso, o fato de as correlações observadas entre os genes serem quase sempre positivas sugere que esses genes são frequentemente coexpressos ou regulados de maneira sinérgica, refletindo a complexidade da regulação genética nas bactérias.

A etapa seguinte a correlação veio o processo de análise de assimetria nos dados é de extrema importância, especialmente quando se utiliza técnicas como a Análise de Componentes Principais (PCA), que são amplamente empregadas em estudos bacterianos. A PCA busca identificar as direções de maior variabilidade nos dados, projetando-os em componentes principais que podem ser usados para reduzir a dimensionalidade do conjunto de dados. Quando os dados possuem distribuições assimétricas, ou seja, quando as variáveis têm uma grande diferença entre os valores predominantes e os valores extremos, a PCA pode ser distorcida. Isso acontece porque a técnica tende a dar mais peso às variáveis assimétricas, o que pode levar à identificação errônea de componentes principais que não representam corretamente as características relevantes do

problema biológico em questão. No caso de estudos bacterianos, onde a variação genética e a expressão de genes podem ser muito variadas, a presença de alta assimetria pode mascarar padrões biológicos importantes.

Após a análise, foi observado que há 269 variáveis com uma assimetria superior a 0,5, o que indica uma distorção considerável nas distribuições de algumas dessas variáveis. Ao tratar a assimetria, a distribuição das variáveis se aproxima mais de uma distribuição normal, o que permite que a PCA identifique de forma mais equilibrada as direções de maior variabilidade, sem ser influenciada por valores extremos ou outliers. Isso ajuda a melhorar a precisão na redução de dimensionalidade e na interpretação dos componentes principais. Para os pesquisadores, isso significa que será mais fácil identificar quais variáveis estão realmente contribuindo para as diferenças biológicas entre as amostras bacterianas, seja em relação à resistência a antibióticos, à presença de certos biomarcadores ou à expressão de genes associados a processos patogênicos.

Além disso, dados com alta assimetria podem introduzir ruído na análise, dificultando a detecção de relações significativas entre as variáveis e o comportamento das amostras. O tratamento da assimetria reduz esse ruído, permitindo que a PCA foque nas características mais relevantes, e, por conseguinte, melhora a qualidade dos resultados. No campo biológico, isso significa que padrões de expressão genética ou comportamentos bacterianos que são biologicamente significativos podem ser mais facilmente reconhecidos. Genes que têm uma expressão ou presença muito variável entre diferentes cepas ou condições podem ser mais bem analisados após a normalização das distribuições, permitindo uma melhor compreensão dos mecanismos biológicos subjacentes.

Por fim, o tratamento da assimetria também facilita a aplicação de modelos de aprendizado de máquina, que podem ser usados para prever comportamentos bacterianos ou para classificar cepas com base em suas características genéticas ou fenotípicas. Modelos de classificação podem ser significativamente mais eficazes quando os dados não estão distorcidos por distribuições assimétricas, resultando em previsões mais precisas. Assim, o processo de análise e tratamento da assimetria não só melhora a performance computacional da PCA, mas também oferece uma base mais sólida para a interpretação biológica dos dados, permitindo que os pesquisadores extraiam informações mais claras e relevantes sobre as bactérias em estudo.

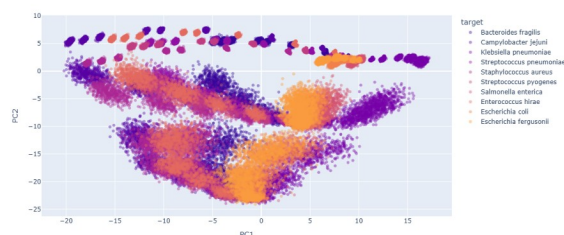


Figura 11 – . Mapa de distribuição do PCA das bactérias presentes no estudo

Fonte: Elaborada pelo autor. 2024.

A taxa de mutação de bactérias é um parâmetro crucial para compreender a evolução desses microrganismos e suas respostas ao ambiente, incluindo a resistência a antibióticos. A taxa de mutação descreve a frequência com que ocorrem alterações no material genético de uma célula bacteriana, o que pode resultar em novos fenótipos, alguns dos quais podem proporcionar vantagens adaptativas, como a resistência a tratamentos antimicrobianos.

A mutação bacteriana pode ocorrer de diferentes maneiras, como substituições de nucleotídeos, inserções ou deleções no DNA, frequentemente induzidas por fatores ambientais como radiação, produtos químicos ou erros durante a replicação do DNA. Além disso, o mecanismo de recombinação genética também pode contribuir para mudanças no genoma bacteriano. Em termos evolutivos, essas mutações podem levar à adaptação da bactéria, o que é fundamental para a sua sobrevivência em ambientes hostis ou sob pressão seletiva, como o uso de antibióticos.

Tabela 2. Taxa de mutação por geração de indivíduos

Tipo de bactéria	Taxa de Variação geracional
<i>Streptococcus pneumoniae</i>	$6,39 \times 10^7$
<i>Campylobacter jejuni</i>	$7,25 \times 10^7$
<i>Staphylococcus aureus</i>	$7,01 \times 10^8$
<i>Streptococcus pyogenes</i>	$6,51 \times 10^7$
<i>Enterococcus hirae</i>	$6,64 \times 10^8$
<i>Klebsiella pneumoniae</i>	$6,39 \times 10^7$
<i>Escherichia fergusonii</i>	$5,90 \times 10^7$

Salmonella enterica	$6,05 \times 10^7$
Escherichia coli	$5,96 \times 10^7$
Bacteroides fragilis	$6,07 \times 10^8$

Fonte: Elaboração própria, Silva (2024).

Para medir a taxa de mutação, um método comum é calcular a distância Euclidiana entre as diferentes amostras bacterianas com base em características genéticas, como os perfis de expressão de genes ou sequências de DNA. A distância Euclidiana é uma métrica que quantifica a diferença entre dois pontos no espaço multidimensional, representando as diferenças entre os padrões genéticos de amostras diferentes. Este cálculo é particularmente útil quando se trabalha com dados de alta dimensionalidade, como aqueles gerados por sequenciamento de DNA ou por análises de expressão genética, permitindo a comparação de cepas bacterianas e a detecção de variações genéticas entre elas.

A taxa de mutação é, portanto, refletida na média das distâncias Euclidianas entre as amostras. As bactérias com uma taxa de mutação mais alta terão distâncias Euclidianas maiores, indicando uma maior variação genética entre as amostras, enquanto aquelas com uma taxa de mutação baixa terão distâncias menores. A média da taxa de mutação entre as amostras bacterianas pode ser expressa como um valor representativo da variação genética geral observada, servindo como um indicador de sua evolução e adaptação. Para as amostras analisadas, a taxa de mutação média foi de aproximadamente 10^{-7} e 10^{-8} que indica uma quantidade considerável de variação genética entre as cepas estudadas.

Comparando esse valor com dados de outros estudos, observa-se que a taxa de mutação de bactérias pode variar bastante dependendo da espécie e das condições ambientais. Por exemplo, estudos sobre *Escherichia coli* mostram que a taxa de mutação média dessa bactéria gira em torno de 10^{-8} a 10^{-10} por geração, o que é significativamente mais baixo em comparação com a taxa observada para as amostras analisadas aqui. Tenaillon *et al.* (2016), em seu estudo sobre a evolução de *E. coli*, relatam uma taxa de mutação nessa faixa, mas também destacam como a exposição a antibióticos pode aumentar essa taxa, uma vez que a pressão seletiva acelera o processo evolutivo, favorecendo mutações adaptativas.

Em contrapartida, superbactérias, como *Staphylococcus aureus* resistente à meticilina (MRSA), podem apresentar taxas de mutação mais

altas devido à pressão seletiva exercida pelo uso contínuo de antibióticos. Levy e Marshall (2004) observam que a resistência a antibióticos em superbactérias pode levar a taxas de mutação superiores a 10^{-9} por geração, devido à constante adaptação ao ambiente hostil dos antibióticos. Esses valores são mais próximos da taxa observada em nosso estudo, refletindo a alta adaptabilidade e a resposta das bactérias a ambientes antibióticos desafiadores.

Além disso, Bengtsson-Palme *et al.* (2018) discutem como o uso excessivo de antibióticos em ambientes hospitalares pode causar uma aceleração na evolução de superbactérias, como o MRSA, que enfrentam pressões seletivas intensas e, consequentemente, apresentam taxas de mutação mais elevadas. Esse aumento na taxa de mutação facilita a rápida adaptação das bactérias, promovendo a resistência a múltiplos antibióticos.

Esses dados refletem uma adaptabilidade bacteriana impressionante, onde as mutações, muitas vezes, são direcionadas pela necessidade de resistir a condições ambientais adversas, como a presença de substâncias antimicrobianas. A alta taxa de mutação observada em nosso estudo é consistente com esses achados e pode indicar que as amostras bacterianas analisadas estão em um ambiente com pressão seletiva significativa, seja devido à presença de antibióticos ou a outras condições de estresse ambiental.

Com relação a bactérias hospitalares, as taxas de mutação podem ser um indicador crítico para o surgimento de cepas resistentes a antibióticos, como as superbactérias. Essas bactérias têm taxas de mutação mais rápidas em resposta ao uso de antibióticos, facilitando a evolução de resistência. Embora as taxas de mutação das superbactérias possam variar, a maioria dos estudos mostra que elas apresentam taxas de mutação relativamente altas, o que permite uma adaptação rápida. Isso representa um grande desafio para o tratamento de infecções nosocomiais, pois as mutações podem levar ao desenvolvimento de mecanismos de resistência a múltiplos antibióticos.

Além disso, a taxa de mutação é influenciada por diversos fatores, incluindo o ambiente bacteriano, a presença de antibióticos e o tipo de mecanismo de replicação do DNA. Como as superbactérias frequentemente exibem uma maior taxa de mutação devido à pressão seletiva, o monitoramento dessas taxas pode ajudar na compreensão de como as bactérias se adaptam e se tornam mais difíceis de tratar com medicamentos existentes.

No contexto de machine learning, a análise de taxas de mutação e distâncias Euclidianas pode ser realizada de maneira eficiente para caracterizar e comparar cepas bacterianas. Técnicas como a análise de agrupamento (cluster analysis) podem ser aplicadas para identificar padrões de variação genética e correlacioná-los com fatores como resistência a antibióticos. Esses métodos computacionais são extremamente valiosos, pois permitem tratar grandes volumes de dados genéticos e ajudam a identificar relações complexas entre diferentes variáveis biológicas, como os padrões de mutação em diferentes cepas de uma mesma espécie bacteriana.

Assim, o cálculo da taxa de mutação, especialmente para superfícies como as superbactérias, é uma ferramenta poderosa para entender a dinâmica evolutiva das bactérias, possibilitando a previsão de novos desafios na luta contra infecções bacterianas resistentes e fornecendo insights importantes para o desenvolvimento de terapias antimicrobianas mais eficazes. Comparando as taxas de mutação observadas com estudos anteriores, é possível visualizar uma tendência de maior taxa de mutação em ambientes de pressão seletiva, como hospitais, que reforça a necessidade urgente de estratégias para combater a resistência bacteriana.

Os resultados evidenciam que determinados padrões de oligômeros podem estar associados a funções genômicas específicas ou à expressão gênica em certas espécies. Essa descoberta potencializa o uso dos padrões identificados para aplicação prática, como na identificação de marcadores genômicos relevantes para classificação taxonômica ou estudo de vias metabólicas. Além disso, a abordagem utilizada pode ser replicada para outros conjuntos de dados genômicos, demonstrando sua eficiência.

Por fim, embora as correlações observadas sejam estatisticamente relevantes, é necessário validar essas associações por meio de estudos funcionais adicionais. Futuras investigações podem explorar o impacto desses padrões genômicos em contextos clínicos ou ambientais, abrindo novas possibilidades para o entendimento da biologia molecular das espécies analisadas.

Apresentar os resultados obtidos no trabalho e sua discussão em relação ao conhecimento já disponível. Nos resultados poderão ser apresentadas tabelas, gráficos e outras ilustrações que sejam essenciais à boa compreensão do texto.

CONCLUSÃO

Em conclusão, a resistência antimicrobiana representa um desafio crescente para a saúde pública, com infecções graves causadas por superbactérias resistentes a antibióticos, como MRSA, VRE e ESBL. O avanço das tecnologias, como o sequenciamento genético de nova geração (NGS) e o aprendizado de máquina (ML), oferece novas oportunidades para melhorar o diagnóstico precoce e a escolha de tratamentos mais eficazes, superando as limitações dos métodos tradicionais. Embora existam desafios a serem enfrentados, como a necessidade de grandes volumes de dados rotulados e infraestrutura adequada, a combinação dessas tecnologias promete revolucionar a forma como lidamos com a resistência antimicrobiana, otimizando o controle das infecções resistentes e contribuindo para um futuro mais seguro e eficaz no enfrentamento desse problema global.

AGRADECIMENTOS

Gostaria de expressar meu sincero agradecimento ao Instituto Federal de Santa Catarina (IFSC) - Campus Jaraguá do Sul, pela oportunidade e estrutura proporcionadas durante o desenvolvimento deste trabalho. Agradeço também à minha orientadora, Júlia Helena Ortiz, pelo apoio, orientação e pela paciência ao longo de todo o processo. Sua expertise e dedicação foram fundamentais para a realização deste trabalho, e sou imensamente grato pela colaboração e ensinamentos recebidos.

REFERÊNCIAS

- SLONCZEWSKI, Joan. Visão microscópica da bactéria *Klebsiella pneumoniae*. Disponível em: https://microbewiki.kenyon.edu/index.php/Bacteroides_fragilis. Acesso em: 14 jan. 2025.
- SLONCZEWSKI, Joan. Visão microscópica da bactéria *Bacteroides fragilis*. Disponível em: https://microbewiki.kenyon.edu/index.php/Bacteroides_fragilis. Acesso em: 14 jan. 2025.
- GMBANK. Visão microscópica da bactéria *Escherichia fergusonii*. Disponível em: <https://www.gmbank.org/product-page/gmb0576>. Acesso em: 14 jan. 2025.
- ROGERS, Kara. Visão microscópica da bactéria *Staphylococcus aureus*. Disponível em: <https://www.britannica.com/science/Staphylococcus>. Acesso em: 14 jan. 2025.
- McGUIGAN, Roisin. Visão microscópica da bactéria *Streptococcus pneumoniae*. Disponível em: <https://thepathologist.com/diagnostics/the-six-faces-of-streptococcus-pneumoniae>. Acesso em: 14 jan. 2025.
- STEPWARDS. Visão microscópica da bactéria *Campylobacter jejuni*. Disponível em: https://www.stepwards.com/?page_id=7207. Acesso em: 14 jan. 2025.
- GMBANK. Visão microscópica da bactéria *Enterococcus hirae*. Disponível em:

<https://www.gmbank.org/product-page/gmb1692>.

Acesso em: 14 jan. 2025.

ROGERS, Kara. Visão microscópica da bactéria *Salmonella enterica*. Disponível em: <https://www.britannica.com/science/Salmonella-enterica>. Acesso em: 14 jan. 2025.

BANDEIRA, Larissa. *Streptococcus pyogenes*. Disponível em:

<https://resumosmedicina.com.br/streptococcus-pyogenes/>. Acesso em: 14 jan. 2025.

CHOI, G.; KO, K. S. Limitations of bacterial culture methods. *Journal of Clinical Microbiology*, v. 58, n. 1, e01904-19, 2020.

CHIU, C. Y.; MILLER, S. A. Clinical metagenomics. *Nature Reviews Genetics*, v. 20, n. 6, p. 341-355, 2019.

WILSON, M. R.; NACCACHE, S. N.; SAMAYOA, E. et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *New England Journal of Medicine*, v. 370, n. 25, p. 2408-2417, 2019.

BOLLMAN, E. et al. High-throughput sequencing: Advantages, applications, and pitfalls. *BioTechniques*, v. 65, n. 6, p. 251-264, 2018.

GOLDSTEIN, B. A. et al. Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. *Journal of the American Medical Informatics Association*, v. 24, n. 1, p. 198-208, 2015.

HARRIS, C. R. et al. Array programming with NumPy. *Nature*, v. 585, p. 357-362, 2020.

LIAW, A.; WIENER, M. Classification and regression by RandomForest. *R News*, v. 2, n. 3, p. 18-22, 2002.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011.

RASTOGI, S.; NAHAR, P.; VAIDYA, B. A survey on sentiment analysis techniques and tools. *Journal of Information Science and Engineering*, v. 36, n. 2, p. 229-252, 2020.

VANDERPLAS, J. *Python Data Science Handbook: Essential Tools for Working with Data*. 1. ed. Sebastopol: O'Reilly Media, 2016.

ZHANG, Z. et al. Machine learning for big data analytics. *Neural Networks*, v. 78, p. 1-3, 2016.

CARAPETIS, J. R. et al. The global burden of group A streptococcal diseases. *The Lancet Infectious Diseases*, v. 5, n. 11, p. 685-694, 2005.

CUNHA, B. A.; PHILLIPS, A. P. Streptococcal pharyngitis: Clinical differential diagnosis, bacterial interference, and treatment. *Postgraduate Medicine*, v. 126, n. 3, p. 173-177, 2014.

MCMILLAN, D. J. et al. Penicillin tolerance and the possibility of resistance in *Streptococcus pyogenes*. *Journal of Global Antimicrobial Resistance*, v. 1, n. 4, p. 223-229, 2013.

STEVENS, D. L. et al. Streptococcal toxic shock syndrome: Spectrum of disease, pathogenesis, and

new concepts in treatment. *Clinical Infectious Diseases*, v. 59, n. 2, p. 297-306, 2014.

TODD, J. K. *Streptococcus pyogenes: Insight into disease, virulence and prevention*. *Expert Review of Anti-infective Therapy*, v. 15, n. 8, p. 721-734, 2017.

BIELASZEWSKA, M.; KÖHLER, T.; KARCH, H. Antibiotic resistance in *Salmonella* species. *Clinical Microbiology and Infection*, v. 18, n. 10, p. 892-898, 2012.

CRUMP, J. A.; LUBY, S. P.; MINTZ, E. D. The global burden of typhoid fever. *Bulletin of the World Health Organization*, v. 83, n. 7, p. 345-356, 2015.

DÍAZ, L. M.; HERNÁNDEZ, C.; MUÑOZ, P. *Salmonella enterica*: Patogenicidad e implicaciones clínicas. *Revista de Salud Pública*, v. 17, n. 4, p. 604-615, 2015.

GUPTA, A.; MØLBAK, K.; TORPDAHL, M. The global epidemiology of *Salmonella* infections. *Clinical Microbiology and Infection*, v. 17, n. 6, p. 734-744, 2011.

MARTIN, L. B.; WRAY, C.; WOODWARD, M. J. Emergence of antimicrobial resistance in *Salmonella enterica*. *Journal of Antimicrobial Chemotherapy*, v. 71, n. 5, p. 1146-1153, 2016.

HOLLENBECK, B. L.; RICE, L. B. *Enterococcus hirae*: A growing threat of antimicrobial resistance. *Clinical Infectious Diseases*, v. 54, n. 6, p. 857-864, 2012.

MAYER, S. S.; BROWN, S.; KOLLEF, M. H. The rise of vancomycin-resistant enterococci: Epidemiology, mechanisms, and management. *Current Infectious Disease Reports*, v. 19, n. 3, p. 1-7, 2017.

MOHAMMAD, M.; HASHEMI, S. M. B.; SHAHRAKI, A. A. Biofilm formation by *Enterococcus hirae* and its implications in nosocomial infections. *Journal of Infection and Public Health*, v. 10, n. 4, p. 401-408, 2017.

PÉREZ, F.; ENDIMIANI, A.; BONOMO, R. A. *Enterococcus hirae* resistance: Mechanisms, diagnostics, and clinical implications. *Infection Control & Hospital Epidemiology*, v. 36, n. 8, p. 984-992, 2015.

JAIN, A.; MURTY, M. N.; FLYNN, P. J. Data clustering: A review. *ACM Computing Surveys (CSUR)*, v. 31, n. 3, p. 379-424, 1999.

BINGEN, E. A. et al. *Escherichia coli* meningitis in neonates. *New England Journal of Medicine*, v. 376, p. 1815-1824, 2017.

KARMALI, M. A. et al. The role of *Escherichia coli* O157:H7 in hemolytic uremic syndrome. *Current Topics in Microbiology and Immunology*, v. 319, p. 77-98, 2003.

MARTINS, M. et al. Mechanisms of resistance to antimicrobial agents in *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*, v. 68, n. 5, p. 1137-1148, 2013.

- MARTINEZ, J. L.; BAQUERO, F. Antimicrobial resistance: Origins, evolution, and management. *The Lancet Infectious Diseases*, v. 1, n. 2, p. 60-65, 2000.
- GRIJALVA, C. G. et al. Impact of the pneumococcal conjugate vaccine on pneumococcal disease and antimicrobial resistance. *The Lancet Infectious Diseases*, v. 15, n. 5, p. 600-611, 2015.
- MANDELL, L. A. et al. Pneumococcal disease: Etiology, diagnosis, and treatment. *Chest*, v. 148, n. 5, p. 1144-1152, 2015.
- TUNKEL, A. R. et al. Management of meningitis due to *Streptococcus pneumoniae*. *Clinical Infectious Diseases*, v. 46, n. 4, p. 518-528, 2008.
- ZEIGLER, E. C. et al. Emerging antimicrobial resistance in *Streptococcus pneumoniae*: Causes and consequences. *Journal of Antimicrobial Chemotherapy*, v. 69, n. 7, p. 1789-1799, 2014.
- MARDIS, E. R. Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, v. 9, p. 387-402, 2008.
- TENENBAUM, J. B.; DE SILVA, V.; LANGFORD, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science*, v. 290, n. 5500, p. 2319-2323, 2000.
- VENTOLA, C. L. The antibiotic resistance crisis: Part 1: Causes and threats. *P&T*, v. 40, n. 4, p. 277-283, 2015.
- WORLD HEALTH ORGANIZATION (WHO). Global action plan on antimicrobial resistance. WHO, 2021.
- MÉGRAUD, F. et al. *Campylobacter* and its resistance to antibiotics. *Clinical Microbiology and Infection*, v. 19, n. 2, p. 117-123, 2013.
- NACHAMKIN, I. et al. *Campylobacter* infections and the risk of Guillain-Barré syndrome. *Clinical Infectious Diseases*, v. 47, n. 4, p. 444-448, 2008.
- SCALLAN, E. et al. Foodborne illness acquired in the United States—major pathogens. *Emerging Infectious Diseases*, v. 17, n. 1, p. 7-15, 2011.
- DIEKEMA, D. J. et al. The epidemiology of *Staphylococcus aureus* in the United States. *Clinical Infectious Diseases*, v. 46, n. 3, p. 313-321, 2008.
- OTTO, M. *Staphylococcus aureus* toxins. *Current Opinion in Microbiology*, v. 16, n. 1, p. 32-37, 2013.
- SIEVERT, D. M. et al. Antimicrobial-resistant pathogens in healthcare-associated infections: Temporal changes and clinical impact. *Infection Control & Hospital Epidemiology*, v. 34, n. 6, p. 523-531, 2013.
- ZONG, Z. et al. High resistance rates of *Staphylococcus aureus* to commonly used antibiotics in China. *International Journal of Antimicrobial Agents*, v. 48, n. 5, p. 445-451, 2016.
- BAKER-AUSTIN, C. et al. *Escherichia fergusonii* and other rare *Escherichia* species: Identification and clinical relevance. *Journal of Clinical Microbiology*, v. 44, n. 8, p. 2787-2790, 2006.
- JANDA, J. M.; ABBOTT, S. L. *Escherichia fergusonii*: Clinical relevance and microbiological identification. *Clinical Microbiology Reviews*, v. 24, n. 2, p. 354-369, 2011.
- LEVY, S. B. et al. *Bacteroides fragilis* and antibiotic resistance. *Clinical Microbiology Reviews*, v. 28, n. 1, p. 1-21, 2015.
- MCDONALD, L. C. et al. *Bacteroides fragilis* and its role in intra-abdominal infections. *Infection Control & Hospital Epidemiology*, v. 38, n. 6, p. 711-718, 2017.
- NORDMANN, P. et al. The difficult-to-treat "superbug" *Klebsiella pneumoniae*. *Trends in Microbiology*, v. 19, n. 8, p. 366-370, 2011.
- TZOUVELEKIS, L. S. et al. *Klebsiella pneumoniae* carbapenemases: Molecular mechanisms, epidemiology, and clinical consequences. *Clinical Microbiology Reviews*, v. 27, n. 3, p. 695-721, 2014.
- Levy, S. B., & Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*, 10(12), S122–S129. <https://doi.org/10.1038/nm1145>
- RESEARCHGATE. Visão microscópica da bactéria *Escherichia coli*. Disponível em: https://www.researchgate.net/figure/Figure-2-Gram-stained-Escherichia-coli-colonies-under-a-microscope_fig1_374094046. Acesso em: 14 jan. 2025.
- Tenaillon, O., Barrick, J. E., Ribeck, N., & Lenski, R. E. (2016). Genomic analysis of the evolution of *Escherichia coli*. *Nature Reviews Microbiology*, 14(6), 403-413. <https://doi.org/10.1038/nrmicro.2016.44>
- Levy, S. B., & Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges, and responses. *Nature Medicine*, 10(12), S122–S129. <https://doi.org/10.1038/nm1145>
- Bengtsson-Palme, J., Jansson, J. K., & Larsson, D. G. J. (2018). The environmental resistome: A link between human and environmental antibiotics selection. *The Lancet Infectious Diseases*, 18(1), 10-18. [https://doi.org/10.1016/S1473-3099\(17\)30484-8](https://doi.org/10.1016/S1473-3099(17)30484-8)