



**FACULDADE DE MEDICINA DE RIBEIRÃO PRETO - FMRP
UNIVERSIDADE DE SÃO PAULO
PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR**



EDSON ALEXANDRE DO NASCIMENTO SILVA

**Caracterização funcional de potenciais genes de resistência a antibióticos identificados
por um algoritmo de *deep learning***

VERSÃO CORRIGIDA

**RIBEIRÃO PRETO - SP
2023**

EDSON ALEXANDRE DO NASCIMENTO SILVA

Caracterização funcional de potenciais genes de resistência a antibióticos identificados por um algoritmo de *deep learning*

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo como parte dos requisitos necessários para obtenção do título de Mestre em Ciências.

Área de Concentração: Biologia Celular e Molecular.

Orientadora: Prof^ª Dra. María Eugenia Guazzaroni.

**RIBEIRÃO PRETO - SP
2023**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Silva, Edson Alexandre do Nascimento

Caracterização funcional de potenciais genes de resistência a antibióticos identificados por um algoritmo de *Deep Learning*.
Ribeirão Preto, 2023. 63 p. : il.

Dissertação de Mestrado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Biologia Celular e Molecular.

Orientador: Prof^a. Dra. María Eugenia Guazzaroni

Palavras-chave: 1. Resistência antimicrobiana. 2. bactéria. 3. genes de resistência antimicrobiana 4. Aprendizado profundo

Versão corrigida. A versão original encontra-se disponível tanto na Biblioteca da Unidade que aloja o Programa, quanto na Biblioteca Digital de Teses e Dissertações da USP (BDTD).

Discente: SILVA, Edson Alexandre do Nascimento

Título: Caracterização funcional de potenciais genes de resistência a antibióticos identificados por um algoritmo de *Deep Learning*

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo como parte dos requisitos necessários para obtenção do título de Mestre em Ciências.

Aprovado em: 09/03/2023

Banca Examinadora

Prof(a). Dr(a). _____

Instituição: _____

Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____

Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____

Assinatura: _____

DEDICATÓRIA

Dedico este trabalho à Deus, amigos e a minha família, em especial meus pais Djalma Vicente da Silva e Maria das Graças do Nascimento Silva.

AGRADECIMENTOS

Primeiramente e sempre, agradeço ao meu Deus todo Poderoso pelas minhas conquistas, pois sem Ele não estaria aqui escrevendo esta dissertação. Tive experiências que eu levo para vida, o que me fez aproximar e entendê-lo mais. De quem virá meu socorro? Do Senhor dos exércitos, socorro presente na escuridão. Obrigado, YESHUA. Sem ti não sou nada, só tenho a agradecer pela misericórdia e graça concedida, mesmo que eu não mereça.

Gostaria de dedicar também um parágrafo especial para uma pessoa que me espelho no ser de cada dia, meu pai. Este tópico de Agradecimentos é muito pouco para expressar com estas letras escritas em fonte Time New Roman o prazer de tê-lo não só como pai, mas como meu melhor amigo aqui neste planeta terráqueo. Mas gostaria de destacar aquele fatídico dia em que me levantei da minha cama às 02:30 da madrugada, ansioso por conta da faculdade e pensava que tudo iria dar errado no futuro para mim. O senhor que sai para trabalhar quase todos os dias às 04:00, se levantou e no descanso me deu conforto naquele momento em que não enxergava um mísero fóton de luz e me falou algo que carrego até hoje: “Onde há vida, há esperança”. Obrigado, papaiô... papaiê.

Tive que lembrar, pois no meu TCC esqueci de colocar você, meu irmão. Mas agradeço o companheirismo em sempre estar comigo e ser um excelente irmão. Agradeço à minha família, em especial, minha mãe pelos cuidados e preocupações diárias com seu “Bom dia”, “Está chovendo aí?”.

Aos meus colegas de laboratório Metagen, agradeço o suporte em cada dia e a paciência em me instruir. Além disso, foram verdadeiros amigos por sempre se preocuparem com meu bem-estar e por darem um tempo para me escutar nas minhas dificuldades. Obrigado Thalita pela ajuda com os materiais, pela atenção e o amor. Agradeço também aos amigos de outros países e os colegas do laboratório do CCBL pelo o apoio para a minha pesquisa. Aos meus amigos desde a graduação, meu muito obrigado pelo apoio quando eu cheguei, em especial, Matias que me hospedou em sua casa e me ajudou em diversas outras coisas.

Agradeço aos irmãos do Ministério Casa de Honra, pela disponibilidade em ajudar e pela comunhão. Que nosso Deus abençoe-os infinitamente.

Agradeço à Universidade de São Paulo, em especial, o programa de Pós-Graduação em Biologia Celular e Molecular e a Faculdade de Medicina de Ribeirão Preto por terem me aceito como discente do programa e pelo suporte institucional que foram essenciais para a continuação do meu mestrado. Agradeço à Faculdade de Filosofia, Ciências e Letras de

Ribeirão Preto pelo espaço para a realização dos meus experimentos. Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de mestrado que deram um suporte essencial para a minha manutenção nesta cidade, Ribeirão Preto.

Por fim, agradeço pelos cuidados de “mãe” da minha orientadora, a Professora Maria Eugenia Guazzaroni. Primeiramente por ter me aceito e também por acreditar no meu potencial. Agradeço por apontar meus erros com suas instruções as quais que vou levar para vida e me esforçar para não repeti-las.

**“Se eu morri, digam que foi pela minha pesquisa”
-Cientista caricato do League of Legends no ARAM**

**“Onde há vida, há esperança”
- Djalma Vicente (Meu pai)**

FONTES DE FINANCIAMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 33002029. Além disso, o laboratório contou com o apoio financeiro da Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP).

Caracterização funcional de potenciais genes de resistência a antibióticos identificados por um algoritmo de *Deep Learning*

RESUMO

A resistência antimicrobiana é um problema de saúde em escala mundial com a morte anual de mais de 1 milhão de pessoas. Entretanto, a situação pode se agravar caso novos antimicrobianos não estejam disponíveis para uso clínico. Em diversos estudos, a presença de organismos resistentes vem sendo identificada em múltiplos ambientes. Dentre eles, o ambiente hospitalar tem sido uma fonte de proliferação de microrganismos super-resistentes a diversos antibióticos e, logo, a transferência de seus genes de resistência a antimicrobianos tem preocupado as instituições de saúde. Portanto, é de total importância o rastreamento das linhagens microbianas e seus genes de resistência nestes locais. As abordagens tradicionais de bioinformática em bancos de dados são baseadas na estratégia de identificação de melhores *hits*. Uma das desvantagens nesta metodologia é o descarte de potenciais genes que estejam abaixo do limiar considerado. Contudo, novas abordagens como a aprendizagem profunda têm sido propostas como uma solução por serem mais sensíveis e precisas. Posteriormente, genes de resistência de microrganismos isolados podem ser identificados por esta metodologia. Portanto, este trabalho teve o objetivo de identificar genes de resistência a antimicrobianos pela ferramenta DeepARG baseada em aprendizagem profunda em genomas de patógenos bacterianos isolados em um hospital. Além da identificação dos potenciais genes de resistência a antibióticos, o estudo realizou uma caracterização funcional dos clones portadores de plasmídeos contendo as sequências selecionadas, na qual não foi observado nenhum sinal de resistência a antibióticos nas condições experimentais utilizadas.

Palavras chave: Resistência antimicrobiana, bactéria, genes de resistência antimicrobiana, Aprendizado profundo

Functional characterization of potential antibiotic resistance genes identified by a Deep Learning algorithm

ABSTRACT

The antimicrobial resistance is a health problem worldwide with the annual deaths of more than 1 million people. However, the situation can aggravate in the case that new antimicrobials don't be available for clinic's purposes. In many studies, the presence of resistance microorganisms had been detected in multiple environments. Among them, the clinical environment has been a source of propagation of superresistance microorganisms to many antibiotics and then, the transference of their antimicrobial resistance genes has been warning out health care organizations. Therefore, it's total importance the tracking of microbial strains and their resistance genes in these places. The traditional approaches of bioinformatic in databases are based on the strategy by best hit identification. One of the disadvantages in this approach is the retreat of potential genes that are below the cutoff considerate. However, novel approaches as the Deep learning have been proposed as a solution for being more sensitive and precise in their analyses. Afterwards, isolated resistance genes from isolated microorganisms can be identify by this methodology. Therefore, this project aimed to identify resistance genes by a Deep Learning based tool, the DeepARG in genomes of pathogenic bacterias isolated from hospital. Besides the identification of potential antibiotic resistance genes, the study carried out a functional characterization of clones carrier of plasmids having the selected sequences, in which it was not identified any signal of antibiotic resistance in experimental conditions used.

Keywords: Antimicrobial resistance, bacteria, antimicrobial resistance genes, Deep learning

LISTA DE ILUSTRAÇÕES

Figura 1 - Linha do tempo da descoberta da resistência de patógenos a diversos antibióticos.	16
Figura 2 - Mecanismos de resistência a antibióticos.	22
Figura 3 - Mecanismos de disseminação da resistência a antibióticos entre bactérias.	25
Figura 4 - Arquitetura geral das redes neurais.	30
Figura 5 - Fluxograma de funcionamento do DeepARG.	33
Figura 6 - Esquema geral da metodologia do projeto.	36
Figura 7 - Esquematização da construção gênica realizada no pUC19 para a caracterização funcional.	38
Figura 8 - Esquema geral do ensaio de gotejamento.	41
Figura 9 - Metodologia da avaliação do <i>burden</i> metabólico para os clones contendo as sequências selecionadas.	41
Figura 10 - Genes de resistência a antibióticos identificados pelo DeepARG.	42
Figura 11 - Panorama geral de um potencial ARG falso positivo identificado pelo DeepARG.	43
Figura 12 - Avaliação da resistência antimicrobiana do vetor pUC19 frente a antibióticos beta-lactâmicos.	47
Figura 13 - Antibiograma dos clones contendo as sequências selecionadas.	48
Figura 14 - Resultados do ensaio de gotejamento para as sequências selecionadas.	49
Figura 15 - Curva de crescimento dos clones contendo as sequências de interesse.	50

LISTA DE TABELAS

Tabela 1 - Resultados do BLAST das sequências pré-selecionadas.	43
Tabela 2 - Características gerais das sequências de genes selecionadas.	44
Tabela 3 - Resultado da caracterização <i>in silico</i> das proteínas de resistência a antibióticos	46

SUMÁRIO

1 INTRODUÇÃO	15
1.1 Resistência antimicrobiana: Passado, presente e reflexões para o futuro	15
1.2 Mecanismos de resistência antimicrobiana	21
1.2.1 Classificação geral da resistência antimicrobiana	21
1.2.2 Redução intracelular dos antimicrobianos na célula bacteriana	22
1.2.3 Alteração do alvo do antimicrobiano	23
1.2.4 Modificação enzimática ou inativação do agente antimicrobiano	24
1.2.5 Mecanismos de disseminação da resistência antimicrobiana	25
1.3 Resistência antimicrobiana: inimiga à espreita, novos métodos para vigiá-la	27
1.4 Deep Learning: debugando a inteligência artificial	29
1.4.1 Estrutura das redes neurais artificiais	30
1.4.2 Aplicação do <i>Deep Learning</i> em problemas de biologia computacional	31
1.4.2.1 O caso DeepARG	32
2 OBJETIVOS	35
2.1 Objetivo geral	35
2.2 Objetivos específicos	35
3 MATERIAIS E MÉTODOS	36
3.1 Panorama geral da metodologia do projeto	36
3.2 Identificação de genes de resistência a antibióticos pelo DeepARG	36
3.3 Estratégia experimental adotada para caracterização funcional dos genes selecionados	37
3.4 Caracterização <i>in silico</i> das sequências selecionadas	39
3.5 Teste de difusão de disco para os clones contendo as sequências selecionadas	39
3.6 Teste de resistência a antibióticos por ensaio de gotejamento (<i>drop assay</i>)	40
3.7 Avaliação do <i>burden</i> metabólico dos clones contendo as sequências selecionadas	41
4 RESULTADOS	42
4.1 Identificação de genes de resistência a antibióticos pelo DeepARG	42
4.2 Caracterização <i>in silico</i> das proteínas de resistência a antibióticos	45
4.3 Teste de difusão de disco para as sequências selecionadas	46
4.4 Teste de resistência a antibióticos por ensaio de gotejamento (<i>drop assay</i>)	48
4.5 Avaliação do <i>burden</i> metabólico dos clones contendo as sequências selecionadas	49
5 DISCUSSÃO	51
6 CONCLUSÕES	55
7 BIBLIOGRAFIA	56
ANEXO A - Resultado da digestão do vetor pUC19 contendo as sequências selecionadas	62
ANEXO B - ARGs identificados pelo programa ABRicate	62
ANEXO C - Antibiograma do clone contendo o gene KPC-2	63

1 INTRODUÇÃO

1.1 Resistência antimicrobiana: Passado, presente e reflexões para o futuro

O conceito de resistência antimicrobiana é dividido em três pontos de vista: epidemiológico, operacional e clínico (MARTÍNEZ; COQUE; BAQUERO, 2014). A resistência antimicrobiana no pensamento **clínico** é baseada no ponto de interrupção do MIC (concentração mínima inibitória), que expressa as concentrações em que o patógeno se torna suscetível, intermediário ou resistente ao antibiótico (MACGOWAN; WISE, 2005). Entretanto, esta visão é inadequada para bactérias não patogênicas a humanos (MARTÍNEZ; COQUE; BAQUERO, 2014).

O conceito **epidemiológico** é determinado pelo Comitê Europeu de Testes de Suscetibilidade Antimicrobiana (EUCAST) com valores de corte de resistência (ECOFFs). O ECOFF é determinado através de um gráfico com valores de MIC para uma população de uma espécie selvagem com atribuição do valor ECOFF para o limite máximo do MIC para a espécie. Logo, uma linhagem isolada da espécie que esteja acima do ECOFF pode ser considerada uma linhagem portadora de genes de resistência (ESPINEL-INGROFF; TURNIDGE, 2016).

Já o conceito **operacional** baseia-se na comparação da espécie selvagem com a espécie mutada ou que tenha genes de resistência adquiridos pelo processo de transferência horizontal (MARTÍNEZ; COQUE; BAQUERO, 2014). Caso a linhagem mutada tenha um MIC maior que a linhagem selvagem da espécie, esta é considerada resistente (MARTÍNEZ; COQUE; BAQUERO, 2014). Este conceito é muito usado em estudos de genômica e metagenômica em que se pode observar a expressão de um gene que torna o hospedeiro mais resistente (MARTÍNEZ; COQUE; BAQUERO, 2014).

A descoberta dos antibióticos a partir da década 20 foi um grande avanço para a humanidade no combate aos microrganismos patogênicos. Entre elas, a descoberta da penicilina teve grande impacto com seu uso durante a Segunda Guerra Mundial. Entretanto, durante este período, casos de isolados bacterianos resistentes à ação da penicilina foram surgindo (PHILLIPS, 1976).

Apesar disso, novos antibióticos da classe dos beta-lactâmicos como meticilina, cefalosporina e carbapenêmicos foram surgindo e contornando o problema. Todavia, não demorou muito para que novos isolados clínicos se tornassem resistentes: em 1962 surgiu o primeiro caso de *Staphylococcus aureus* resistente à meticilina no Reino Unido. Na era de

ouro dos antibióticos, os cientistas partiram para a descoberta de antibióticos naturais e desenvolvimento de antibióticos sintéticos com diferentes mecanismos de ação. Infelizmente, este progresso foi acompanhado pelo surgimento de linhagens resistentes à ação dos antibióticos logo após seu lançamento no mercado (Figura 1) (VENTOLA, 2015).

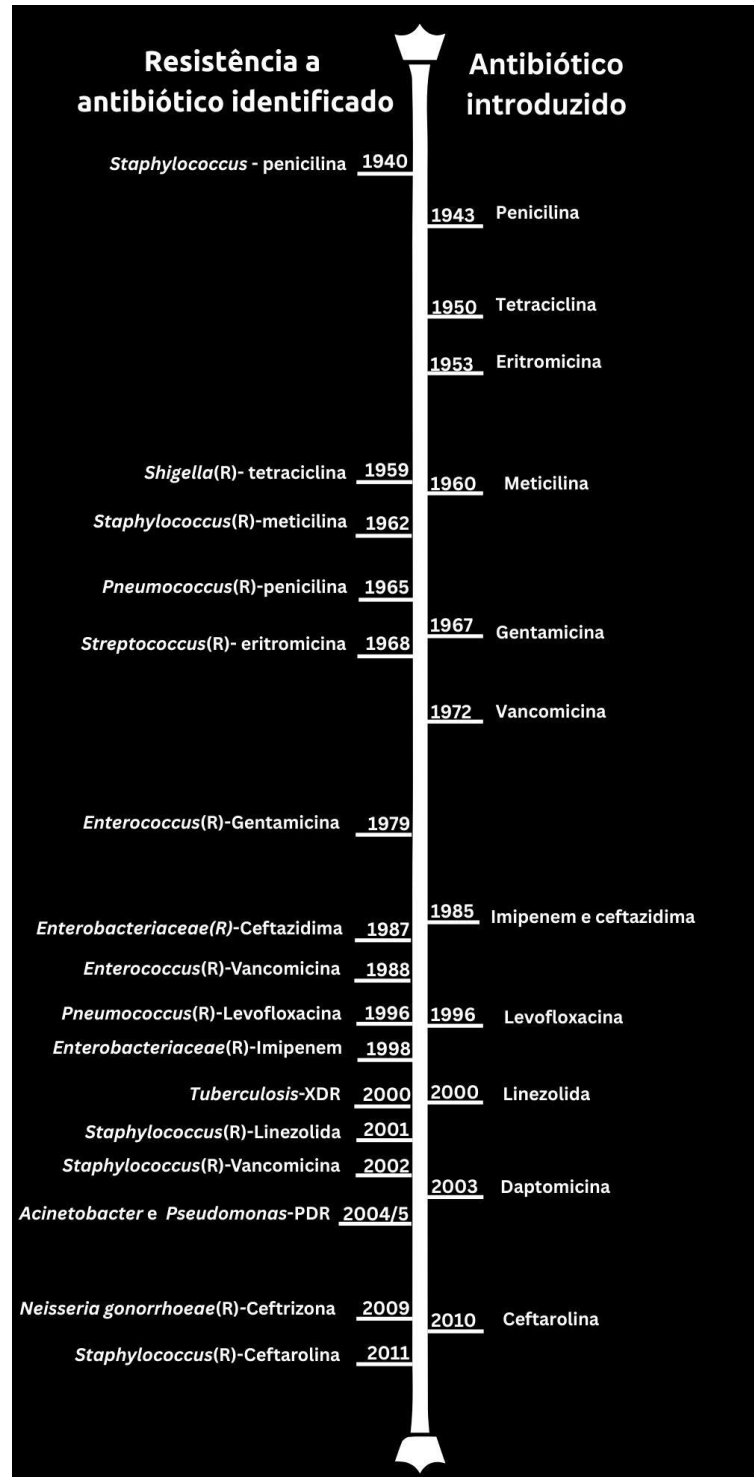


Figura 1 - Linha do tempo da descoberta da resistência de patógenos a diversos antibióticos. Fonte: Adaptado de Ventola (2015). XDR- Resistência extensiva a drogas, PDR- Resistência Pan a drogas, R- Resistente.

Apesar da necessidade de antibióticos com novos mecanismos de ação para combater os microrganismos patogênicos, percebe-se um decaimento da presença de novos agentes antimicrobianos no mercado ao longo dos anos. Em contrapartida, o surgimento de linhagens de bactérias resistentes e multirresistentes a antibióticos têm aumentado consideravelmente (CASSINI *et al.*, 2019; SPELLBERG, 2014).

Neste contexto, a pandemia do coronavírus trouxe uma situação da qual a humanidade não estava habituada com isolamento social e milhões de mortos ao redor do mundo (KIM; JUNG, 2021). Junto à COVID-19, a multirresistência bacteriana se alastrou neste período. Segundo o Centro de Controle e Prevenção de doenças, o CDC apontou que em 2019, 3 milhões de estadunidenses adquiriram alguma infecção de microrganismos resistentes a antimicrobianos (CDC, 2022).

Além do aumento da resistência, outros problemas vieram com a pandemia como dados não disponíveis ou atrasados sobre a resistência antimicrobiana, assim como, a falta de vigilância em hospitais (CDC, 2022). As principais causas a serem apontadas são os esforços de recursos humanos foram voltados mais à COVID e isto inclui os testes voltados para a detecção do vírus SARS-CoV-2 e a sobrecarga da equipe técnica voltada para tratamento de pacientes infectados pelo vírus, o que levou a falhas na prevenção e guia para controle da resistência antimicrobiana em hospitais (GASPAR *et al.*, 2022; KNIGHT *et al.*, 2021).

A maior taxa de resistência pôde ser vista em países subdesenvolvidos e o aumento de linhagens multirresistentes a drogas voltou a crescer. O CDC indicou que pelo menos houve aumento de 15% da presença de linhagens resistentes (CDC, 2022). Em uma pesquisa em Wuhan, o estopim da pandemia, mostrou a presença de 159 linhagens resistentes em 102 pacientes hospitalizados pela COVID-19 (LAI *et al.*, 2021).

Entre as linhagens de espécies bacterianas, *Acinetobacter baumannii* e *K. pneumoniae* foram as mais predominantes; sendo 91,2% e 75%, linhagens resistentes a carbapenêmicos, respectivamente. As razões para esta expansão levam em conta o aumento da prescrição de antibióticos, pois a distinção entre pneumonia bacteriana e coronavírus é difícil, além do que houve o aumento de coinfeções de patógenos (LAI *et al.*, 2021).

Em uma revisão sistemática lançada pela Lancet em 2022 observou que no ano de 2019 cerca de 1.27 milhão de pessoas foram a óbito devido à resistência antimicrobiana, enquanto que em quase 5 milhões de casos, a resistência antimicrobiana esteve associada com a morte de pacientes (MURRAY *et al.*, 2022). Ademais, os dados coletados mostram que entre as bactérias patogênicas mais responsáveis por estes casos estão algumas bactérias do

grupo ESKAPE¹ como *Escherichia coli*, *S. aureus* e *Klebsiella pneumoniae* (MURRAY *et al.*, 2022). Estas bactérias estão na lista de prioridades de bactérias da OMS² em que há várias listas de ocorrências de linhagens multirresistentes a diversos antibióticos, onde se vê a necessidade de novos agentes antimicrobianos capazes de combater estas linhagens resistentes (TACCONELLI *et al.*, 2017).

Em países em desenvolvimento, os fatores para o aumento e dispersão da resistência antimicrobiana são a falta de vigilância, uso equivocado e acessibilidade de antibióticos. Enquanto que em países desenvolvidos, a principal causa está no uso indiscriminado em animais para alimentação e agricultura (CHOKSHI *et al.*, 2019). Neste quesito, a Organização das Nações Unidas para a Alimentação e Agricultura ressalta sua preocupação em torno da maior demanda futura de antibióticos para o crescimento de animais, com destaque para países em desenvolvimento, em especial, os países do bloco econômico BRICS³ (RABELLO *et al.*, 2020).

Em 2015, países membros da OMS formularam o plano global em combate à resistência antimicrobiana. No tocante referente a este tema, o Brasil formalizou em 2018 o seu plano nacional para contenção da resistência. Com a integração dos ministérios do Estado Brasileiro, a PAN-BR⁴ tem 5 objetivos estratégicos, 14 objetivos e 33 intervenções estratégicas que atendam os interesses de organizações mundiais como a FAO⁵ e a OMS (BRASIL, 2018):

I) Promover a conscientização sobre a RAM

- Estabelecer estratégias de comunicação que propiciem a educação da população sobre a resistência antimicrobiana.

II) Fortalecimento do conhecimento através da ciência

- Fomentar a pesquisa brasileira e fortalecer os estudos sobre a resistência antimicrobiana, assim como, estabelecer um sistema nacional de vigilância com a criação de uma rede de laboratórios de vigilância e monitoramento.

III) Redução da incidência de infecções

- Através de ações de saneamento básico, higiene e prevenção, a PAN-BR tem o objetivo de ampliar a cobertura vacinal e de saneamento básico, assim como,

¹ Grupo de bactérias patogênicas que são um alerta para a OMS como *E. coli*, *S. aureus*, *K. pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* e *Enterobacter spp.*

² Organização Mundial da Saúde.

³ Grupo de países emergentes formado pela África do Sul, Brasil, China, Índia e Rússia.

⁴ Plano de Ação Nacional de Prevenção e Controle da Resistência aos Antimicrobianos.

⁵ Organização das Nações Unidas para Alimentação e Agricultura.

estabelecer boas práticas no meio agropecuário e o controle de infecções em comunidades.

IV) Otimização do uso de antimicrobianos

- Estabelecimento do uso consciente de antimicrobianos com a criação de diretrizes visando estabelecer parâmetros para prescrições de medicamentos.
- Promover o acesso a vacinas, antimicrobianos e testes de diagnósticos.
- Gerenciamento adequado de resíduos contendo antimicrobianos de acordo com a Política Nacional de Resíduos Sólidos.

V) Assegurar investimentos na luta contra a resistência antimicrobiana.

- Priorizar o controle da RAM como política de Estado.
- Incentivar o desenvolvimento na produção industrial em medicamentos antimicrobianos e pesquisas envolvendo a RAM.
- Buscar suporte para pesquisas envolvendo resistência antimicrobiana, desenvolvimento de vacinas e antimicrobianos e temas relacionados.

Jim O'Neill (2014) em seu relatório sobre a resistência antimicrobiana reportou que em 2050, 10 milhões de pessoas morrerão anualmente devido à resistência antimicrobiana e que o custo global da resistência levaria cerca de 100 trilhões de dólares. Estes estudos foram comissionados pelos grupos RAND Europe e KPMG que verificaram o impacto da resistência antimicrobiana no futuro.

Entretanto, os dados são uma subestimativa, pois o estudo foi realizado em um subconjunto de microrganismos como *K. pneumoniae* e *E. coli*, assim como, os pesquisadores consideraram apenas o impacto econômico em seus estudos. Logo, caso nada seja feito em relação ao futuro, os números de mortalidade e morbidade podem ser ainda mais agravantes (O'NEILL, 2014).

Porém, é essencial destacar que as previsões relacionadas à resistência antimicrobiana são de difícil acerto, pois planejamentos em políticas públicas de saúde podem flutuar de acordo com os dados em relação à resistência antimicrobiana, como por exemplo as decisões médicas em torno da escolha de um antibiótico a ser prescrito (RABELLO *et al.*, 2020).

Apesar disso, cientistas apontam suas preocupações com o aumento das taxas da resistência antimicrobiana e a ausência de novos antibióticos. Logo, com a ineficiência constatada de diversos antibióticos perante alguns isolados de patógenos resistentes, as instituições de saúde temem uma volta da era pré-antibióticos, onde qualquer infecção pode levar à morte do paciente (ASHKENAZI, 2013).

Nesse ínterim, diversas estratégias alternativas têm surgido como forma de combate aos patógenos resistentes a antimicrobianos. Entre elas, o desenvolvimento de nanopartículas, novas vacinas e o uso de bacteriófagos têm sido destaques como combatentes a infecções de difícil tratamento (FATHIMA; ARCHER, 2021; YOUSAF *et al.*, 2020). Especialistas e organizações da saúde envolvidos na resistência antimicrobiana indicam alguns pontos que devem ajudar na contenção e prevenção do disparate da resistência antimicrobiana. As orientações vão desde o indivíduo até adoções de políticas públicas:

- 1) Uso racional de antimicrobianos: Em tempo e doses corretos, os antimicrobianos devem ser tomados com cautela seguindo as recomendações médicas com o uso do antibiótico correto. Entretanto, o panorama geral é controverso em diversos países em que há o uso exagerado de antimicrobianos, principalmente para outros objetivos como promoção de crescimento animal. Este quadro desditoso pode ser contornado com conscientização da população e medidas de precaução como adoção de vacinas (ALGHAMDI, 2021; BUSCH *et al.*, 2020; LIN *et al.*, 2020).
- 2) Vigilância e controle sanitário: Os países devem tornar ativos os planos governamentais de controle e prevenção da resistência antimicrobiana. Como também, é essencial o fortalecimento de programas de vigilância, assim como adoção de métodos de diagnósticos mais sensíveis e ágeis (OMS, 2015). Ademais, o controle da dispersão de patógenos infecciosos deve ser feito em áreas sensíveis como hospitais e indústrias. Em hospitais, deve-se seguir o controle rigoroso com a limpeza de equipamentos e materiais de uso clínico e adoção de hábitos preventivos como constante higienização das mãos para os frequentadores do ambiente (HARBARTH *et al.*, 2015). Como dito antes, a pandemia demonstrou a falta de controle da resistência em hospitais com o aumento da resistência e casos de infecção bacteriana em consonância com a COVID-19 (GASPAR *et al.*, 2022). Portanto, são necessários a formulação e execução de novas medidas de prevenção e controle para membros do corpo hospitalar como forma de frear a disseminação de patógenos infecciosos. As indústrias farmacêuticas devem ficar cientes do seu papel crítico na disseminação da resistência antimicrobiana com o manuseamento correto dos efluentes, neste caso, os antibióticos (KOTWANI; JOSHI; KALONI, 2021).
- 3) Investimentos em saúde pública: A perspectiva melancólica frente aos investimentos em novos antibióticos tem sido um desânimo. Os motivos

principais são o surgimento rápido de isolados resistentes a antibióticos e a constante replicação dos mesmos antibióticos em estudos de triagem e prospecção (DA CUNHA; FONSECA; CALADO, 2019). Em outro aspecto, a pandemia da COVID no Brasil mostrou o despreparo dos sistemas de saúde em atender uma alta demanda para casos de infecção, assim como, infelizmente a falta de assistencialismo em comunidades mais necessitadas ficou bem evidente (PEREIRA *et al.*, 2020).

1.2 Mecanismos de resistência antimicrobiana

1.2.1 Classificação geral da resistência antimicrobiana

De modo geral, a resistência a antibióticos pode ser dividida em resistência intrínseca e adquirida. A resistência intrínseca é alcançada quando a bactéria naturalmente já possui mecanismos que não permitem a ação ou entrada dos agentes antimicrobianos como a presença ou não de uma estrutura celular como a parede celular (ABUSHAHEEN *et al.*, 2020). A resistência adquirida é observada quando uma bactéria até então suscetível à ação bactericida ou bacteriostática de um antibiótico recebe genes que conferem resistência a este agente (LERMINIAUX; CAMERON, 2019b). Este mecanismo pode ser realizado por transferência horizontal de material genético (THG) entre as bactérias por intermédio de elementos genéticos móveis como plasmídeos e transposons (ARNOLD; HUANG; HANAGE, 2021). Além disso, a resistência bacteriana a antibióticos pode ser dividida de acordo com o mecanismo de ação, dentre elas, destacam-se a redução intracelular dos antimicrobianos, alteração do alvo e modificação enzimática do antibiótico (Figura 2).

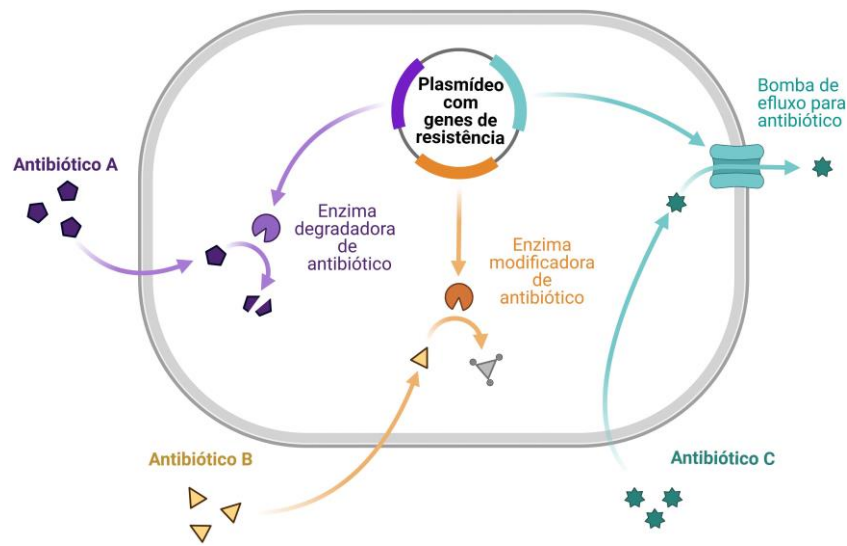


Figura 2 - Mecanismos de resistência a antibióticos. Fonte: A imagem foi gerada no site Biorender.com. Os mecanismos de resistência a antibióticos podem ser por degradação de antibiótico, modificação do antibiótico e efluxo do antibiótico para fora da célula.

1.2.2 Redução intracelular dos antimicrobianos na célula bacteriana

Um antibiótico pode ser impedido de entrar ou pode ser expelido para fora da célula bacteriana, com isso impossibilitando a sua ação. Um dos mecanismos de resistência antimicrobiana ocorre nas porinas, proteínas canais presentes na membrana externa da célula. Esta proteína canal pode ser uma porta de entrada de macromoléculas e drogas como tetraciclina e beta-lactâmicos. Entretanto, mutações pontuais podem alterar o fluxo de entrada com a infra-regulação, modificação estrutural ou deleção dos genes responsáveis por sua síntese. Esta redução do influxo de antibióticos pelas porinas pode ser visto em *E. coli* e outras bactérias Gram negativas (C REYGAERT, 2018).

Por outro lado, sistemas de transportes podem expelir o antibiótico para fora da célula. Entre eles, estão as famílias de regulação-nodulação-divisão celular (RND), cassete de ligação ao ATP (transporte ABC) e extrusão de multidrogas e compostos tóxicos (MATE). Estas famílias de proteínas se caracterizam por serem representantes de sistemas de transporte contra multidrogas, portanto conferem multirresistência à bactéria (ABUSHAHEEN *et al.*, 2020).

A família de transportadores RND é composta por três proteínas: uma proteína presente na membrana interna responsável pelo transporte do antibiótico, uma proteína acessória periplasmática que em conjunto com a proteína canal da membrana externa

transportam o substrato ao realizar a troca pelo antibiótico (NIKAIDO, 2011). Assim sendo, um antibiótico é transferido para fora da célula em troca do transporte interno de íons de hidrogênio (ABUSHAHEEN *et al.*, 2020). Pesquisas relatam a presença deste sistema em isolados de *Mycobacterium tuberculosis* e *Vibrio cholerae* (ALVAREZ-ORTEGA; OLIVARES; MARTÍNEZ, 2013).

O sistema de transporte ABC pode ser responsável pela absorção ou extrusão de macromoléculas como drogas, proteínas e íons. O sistema é caracterizado por usar o ATP como fonte de energia para sua funcionalidade e possui 6 segmentos transmembranares com formas de alfa-hélices e uma ATPase acoplada. Já a família de transportadores MATE usa o gradiente de íons sódio como fonte de energia e permite o efluxo de antibióticos como fluoroquinolonas (C REYGAERT, 2018).

Todavia, as bactérias podem se unir em comunidades e formar o biofilme, uma estrutura constituída por uma rede de proteínas e polissacarídeos pela qual pode variar de acordo com a(s) espécie(s) constituinte(s) (ABE; NOMURA; SUZUKI, 2020). Com a função de comunicação e trocas de macromoléculas, esta larga rede pode impedir a fácil percussão de antibióticos. Com isso, o biofilme pode ser visto como um meio de mecanismo de resistência, assim como, pode fomentar a troca de material genético como genes de resistência (SANTAJIT; INDRAWATTANA, 2016).

1.2.3 Alteração do alvo do antimicrobiano

A resistência molecular a antimicrobianos pode ser destacada pela modificação estrutural do alvo do antibiótico. Este mecanismo tem sido proeminente e tem como consequência a diminuição da afinidade da droga pelo seu alvo. O mecanismo pode ter como origem uma mutação espontânea em um gene responsável pela síntese do alvo do antimicrobiano ou por modificação pós traducional, mas também pode ser adquirido por THG (ABUSHAHEEN *et al.*, 2020; C REYGAERT, 2018; SANTAJIT; INDRAWATTANA, 2016).

O mecanismo pode ser observado em glicopeptídeos, classe de antibióticos que se ligam ao terminal D-Ala-D-ala, um componente da parede celular bacteriano. Com o impedimento da formação da parede celular, a bactéria tem sua consequente morte celular. Entretanto, bactérias Gram-positivas têm formulado mecanismos de combate à ação deste antibiótico com a troca de um resíduo de Alanina por lactato. A mudança provoca a repulsão eletrostática do antibiótico pelo seu alvo e com isso diminui em 1000 vezes a sua afinidade.

Por sua vez, esta modificação é realizada pelo operon de genes *van* (BONECA; CHIOSIS, 2005; STOGIOS; SAVCHENKO, 2020).

A polimixina é uma classe de antibióticos catiônicos que desestabilizam a membrana celular ao ligar-se ao lipopolissacarídeo (LPS). Especificamente, a polimixina se liga a um componente da LPS, o lipídio A, e desloca os íons cálcio e magnésio. Como consequência, este antibiótico provoca a desestabilização da membrana e morte celular com a geração de espécies oxidativas de oxigênio. Entretanto, a bactéria possui um mecanismo de resistência contra a ação desta classe de antibióticos: a modificação do LPS. A adição de grupos funcionais como 4-amino-L-arabinose e galactosamina à porção 1' e 4' do lipídio impedem a ligação da polimixina ao seu alvo (TRIMBLE *et al.*, 2016).

Em antibióticos que se ligam a ribossomos, enzimas chamadas metiltransferases atuam no contra ataque à ação destas classes de antibióticos. Um exemplo são enzimas Erm que modificam o 23S rRNA, parte integrante do ribossomo, e conferem à bactéria resistência a antibióticos que se ligam à porção 50S do ribossomo como macrolídeos, lincosamidas e estreptogramina. Como consequência da adição da metilação ao alvo do antibiótico, o bloqueio estérico ocorre e o antibiótico não consegue se ligar ao seu alvo (WILSON, 2014).

1.2.4 Modificação enzimática ou inativação do agente antimicrobiano

Um mecanismo de resistência a antibióticos peculiar é a inativação do agente antimicrobiano por enzimas pertencentes à bactéria. Este mecanismo pode acontecer por duas vias: modificação enzimática ou degradação do antibiótico. A modificação enzimática ocorre com a adição de um grupo funcional ao antibiótico por uma enzima. A partir desta adição, o antibiótico perde sua funcionalidade. Este mecanismo pode se suceder por acetilação, metilação ou adenilação, sendo a acetilação a mais comum entre as classes de antibióticos (C REYGAERT, 2018).

A acetilação acontece com a adição de um grupo acetil a um antibiótico. Esta reação pode ocorrer em classes de antibióticos como aminoglicosídeos e fluoroquinolonas. Já a metilação e adenilação adicionam um grupo metil ou adenina em um antibiótico, respectivamente. Entretanto, dados literários informam a restrição desta ação apenas em aminoglicosídeos (C REYGAERT, 2018).

A degradação do antibiótico ocorre por uma reação enzimática, onde a estrutura química do antimicrobiano é desfeita e o mesmo perde a sua atividade funcional. Um exemplo

clássico são as beta-lactamases que hidrolisam o anel beta-lactâmico dos antibióticos beta-lactâmicos como penicilina e cefalosporina (BLAIR *et al.*, 2014).

1.2.5 Mecanismos de disseminação da resistência antimicrobiana

A princípio, no “mundo bacteriano”, há três formas canônicas em que o material genético pode ser transmitido entre as espécies de bactérias: transformação, conjugação e transdução. A Figura 3 ilustra como funciona tais mecanismos.

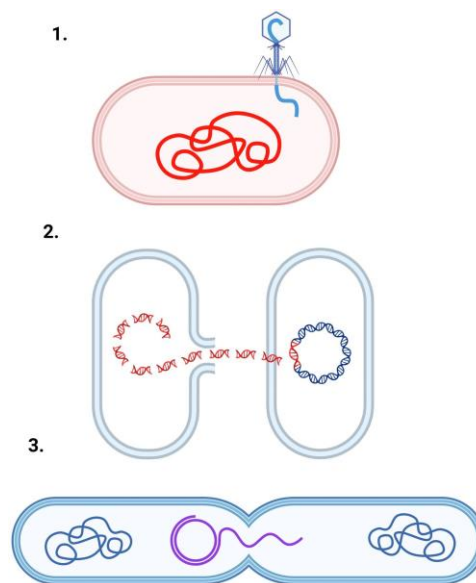


Figura 3 - Mecanismos de disseminação da resistência a antibióticos entre bactérias. Fonte: A Imagem foi criada no site Biorender.com. 1. A transdução ocorre com os seguintes passos: I) Um vírus contendo um gene de resistência a antibióticos (ARG) em seu genoma infecta a bactéria e II) O ARG é incorporado no genoma bacteriano. 2. A transformação ocorre da seguinte maneira: O ARG originado de DNA exógeno é incorporado no genoma bacteriano. 3. A conjugação ocorre com o compartilhamento de ARGs por plasmídeos.

A **transformação** decorre da captação e recombinação de material genético do ambiente para uma bactéria a qual deve ser competente para receber o DNA externo. Este DNA externo tem origem de bactérias que passaram por morte celular e o DNA, assim como os debrís celulares, podem ser encontrados no ambiente externo (LERMINIAUX; CAMERON, 2019b).

A descrição do mecanismo é bem mais descrita em bactérias Gram negativas. Em suma, o tDNA (DNA transformado) é deslocado até a membrana externa pela retração da *pili* e então se liga à proteína de ligação ao tDNA, a ComEA. Em seguida, a proteína canal ComEC realiza o transporte do material genético para o citosol. Nesta fase, apenas uma das

fitas de DNA é transformada, enquanto que a outra fita complementar pode ser fragmentada. O tDNA em forma de fita simples é então recombinada com o DNA cromossomal com o auxílio da proteína de recombinação RecA e proteínas ligantes à fita simples (JOHNSTON *et al.*, 2014).

Este mecanismo é pouco comum, mas pode fornecer ao microrganismo, genes de resistência a antibióticos. Apesar de ser pouco conhecida a transferência em ambiente clínico, acredita-se que se deva à dificuldade na identificação *in vivo* da transformação. Apesar disso, de acordo com Lerminiaux e Cameron (2019b), 11 das 12 bactérias patogênicas que são listadas como preocupantes para a OMS têm a capacidade de serem competentes para a aquisição por este meio.

A **conjugação** se caracteriza pela transferência do DNA originada do plasmídeo de uma bactéria doadora para a bactéria receptora por contato direto. Este mecanismo foi primeiramente relatado por Edward Tatum e Joshua Lederberg que observaram a transferência entre bactérias com plasmídeo fator F. A conjugação é considerada a maior responsável pelas mudanças no genoma bacteriano (VIROLLE *et al.*, 2020).

O início do processo da conjugação se inicia com a expressão de genes chamados *tra* que participam no controle da transferência do material genético para a bactéria receptora. De maneira específica, os genes de transferência codificam os fatores protéicos envolvidos na formação da pilus além do sistema de secreção do tipo IV T4SS requerido para a formação do pareamento entre as bactérias e a transferência do DNA, respectivamente (VIROLLE *et al.*, 2020).

O pilus une as duas células através de um movimento de extensão seguido de uma retração. O início da conjugação necessita da montagem e atividade do complexo proteico relaxossomo, a qual processa o plasmídeo antes da sua passagem para a bactéria receptora (VIROLLE *et al.*, 2020).

O plasmídeo é excisado especificamente pela relaxase TraI e após as duas fitas se dividirem, umas das fitas em forma linear é ligada a proteína TraI, onde o complexo é direcionado para a célula receptora através de poros na membrana. Uma vez presente na célula receptora, o DNA em fita simples (ssDNA) é circularizada pela proteína TraI e a proteína FrPo realiza a conversão da fita dupla da ssDNA (VIROLLE *et al.*, 2020).

A troca de material genético por intermédio da conjugação é observada em ambientes como hospitais: em um trabalho realizado no hospital das clínicas de Ribeirão Preto, Gaspar *et al.* (2022) identificaram linhagens resistentes a carbapenêmicos e polimixinas. Os plasmídeos contendo os genes de resistência a antibióticos eram idênticos a nível genoma à

plasmídeos identificados em linhagens de *K. pneumoniae* em 2018 no mesmo hospital. Este trabalho demonstrou que houve uma transferência horizontal de genes entre diferentes espécies, além da presença intensificada dessas linhagens em diversas alas do hospital (GASPAR *et al.*, 2022). A conjugação não se restringe a ambientes hospitalares, mas pode ser observada em diversos ambientes peculiares como na microbiota humana e biofilmes (ABE; NOMURA; SUZUKI, 2020; NEIL; ALLARD; RODRIGUE, 2021).

A **transdução** é intermediada por um vírus que infecta bactérias, o bacteriófago. Ao infectar a célula bacteriana, o vírus integra seu genoma viral ao genoma do hospedeiro. Além do seu próprio genoma, o vírus pode integrar fragmentos de DNA de uma bactéria infectada previamente. Este mecanismo pode ser uma forma benéfica para a célula bacteriana onde pode haver ganho de fatores de resistência a antibióticos (LERMINIAUX; CAMERON, 2019b).

1.3 Resistência antimicrobiana: inimiga à espreita, novos métodos para vigiá-la

Os métodos tradicionais usados para identificação de microrganismos patogênicos são exemplificados pela microdiluição em caldo, ágar-diluição e difusão de disco. Estes testes são amplamente usados em ambientes clínicos para verificação da susceptibilidade de um microrganismo isolado (KOWALSKA-KROCHMAL; DUDEK-WICHER, 2021).

A **microdiluição em caldo** consiste no teste da concentração mínima inibitória (CIM) do patógeno em uma placa de 96 poços. Após a suspensão bacteriana atingir a turbidez de 0.5 na escala McFarland, segue-se a diluição dobrada do antimicrobiano a ser testado nos poços. Uma grande vantagem para este teste é a utilização de uma grande quantidade de antibióticos (8-12) a serem testados uma única vez (TENOVER, 2019).

O teste de **ágar-diluição** é baseado no uso de diversas placas Petri com meio Mueller Hinton contendo diferentes concentrações do antibiótico a ser testado. Assim como a microdiluição em caldo, a bactéria é suspensa em solução salina em uma turbidez de 0.5 da escala McFarland a qual corresponde a concentração bacteriana de 1×10^8 CFU/mL. Em seguida, uma alíquota da suspensão é despejada na placa. Diferente da microdiluição em caldo, o ágar-diluição é viável para experimentos com o uso de diferentes espécies de patógenos por placa. A bactéria é incubada a 37°C até 24 horas e após isso os resultados podem ser constatados com a ausência de colônias na placa, indicando o CIM da bactéria para determinada concentração de antibiótico (TENOVER, 2019).

O teste de **Kirby-Bauer** ou disco-difusão consiste na aplicação de discos de papel contendo um antibiótico em determinada concentração. Uma suspensão bacteriana preparada na escala padrão de 0.5 McFarland é inoculada uniformemente na placa; após 5 minutos, os discos de antibióticos são aplicados no meio. A bactéria é incubada por 37°C por 24 horas e os resultados da resistência ou susceptibilidade do patógeno são observados através da zona de inibição formada ao redor do disco. Um microrganismo que demonstra susceptibilidade apresentará um perfil com um halo maior, enquanto que a resistência pode ser vista com um decréscimo do halo ao redor do disco. Geralmente, a comparação da resistência ou susceptibilidade do patógeno é observado com as determinações padronizadas do halo do Instituto de Padrões Clínicos e Laboratoriais, o CSLI (EUCAST, 2023).

Entretanto, esses testes apresentam determinadas desvantagens que podem interferir nos resultados como algumas espécies bacterianas que são fastidiosas e requerem nutrientes mais ricos no meio de cultivo para crescer. Além disso, algumas bactérias têm um perfil de crescimento lento, o que pode levar a resultados contraditórios. Os resultados destes testes demoram levando-se em conta o tempo do cultivo para a incubação da bactéria. Logo, em um aspecto clínico, é necessário formular novos experimentos que sejam menos laboriosos e com tempo curto de análise, pois o tempo é essencial para as decisões médicas frente às infecções ocasionadas ao paciente (BANERJEE; HUMPHRIES, 2021; PETERSEN *et al.*, 2020).

Recentemente, os métodos de análise não dependentes de cultivo bacteriano têm sido cada vez mais frequente com decréscimo do custo do sequenciamento. Entre estes, pode-se destacar o sequenciamento do genoma completo que tem ganhado destaque em análises como por exemplo, microbiotas ambientais (RANJAN *et al.*, 2016).

O processo para o WGS segue-se primeiro com a extração do DNA, preparo das bibliotecas para o sequenciamento, filtragem das *reads*, montagem de *contigs* e anotação funcional dos genes em bancos de dados. Esta metodologia tem o potencial para a aplicação em vigilância epidemiológica com detecção de surtos, além de rastrear possíveis rotas de transmissão entre as bactérias com a identificação de elementos genéticos móveis (ONICIUC *et al.*, 2018).

Com isso, o sequenciamento em alta performance fornece um aprofundamento acerca dos determinantes genéticos da resistência a antimicrobianos (ONICIUC *et al.*, 2018). Além do WGS, a inteligência artificial tem inovado a forma de tratamento de dados e com isso a aplicação de *Deep Learning* tem se adentrado na resolução de dados biológicos, entre eles, o DeepARG pode ser uma promissora na identificação de genes de resistência a antibióticos (ARANGO-ARGOTY *et al.*, 2018).

Como exemplo, o uso de *Deep Learning* foi essencial para a avaliação do resistoma em 293 amostras de metagenômica providas do projeto oceano TARA. Para este trabalho, Cuadrat *et al.* (2020) processaram 13.168 janelas de leitura aberta que foram identificadas, sendo a maioria pertencentes a genes com resistência a multidrogas. O trabalho também mostrou a diversidade de ARGs (24.567) em plasmídeos, o que mostra a potencialidade da disseminação destas ARGs em ambiente marinho.

1.4 Deep Learning: debugando a inteligência artificial

A aprendizagem profunda ou *Deep Learning* vem do conhecimento da inteligência artificial, sendo uma subárea do aprendizado de máquina⁶. Diferente do *Machine Learning*, o *Deep Learning* não necessita de interferência humana para processamento e aprendizado dos dados fornecidos. Este novo segmento da inteligência artificial vem inovando diversas áreas da ciência de dados, sendo o estado da arte no reconhecimento de voz e imagens, assim como, na descoberta de drogas e no campo da genômica (RUSK, 2015).

A ideia central do *Deep Learning* está em aprender padrões e estruturas presentes em dados brutos passados para a rede neural (HAO; ZHANG; MA, 2016). Um exemplo está na aplicação no reconhecimento de dígitos numéricos, onde o algoritmo deve distinguir as diferentes formas de escritas de números e diferenciar pequenas diferenças em traços entre os números, como por exemplo, os números 4 e 9 (Figura 4) (NIELSEN, 2015; LECUN; BENGIO; HINTON, 2015). Deste mesmo modo, sequências de ácidos nucléicos podem ser diferenciadas entre si a partir de padrões encontrados em aminoácidos, por exemplo, sequências resistentes a diferentes classes de antibióticos (ARANGO-ARGOTY *et al.*, 2018).

⁶ ou *Machine Learning*.

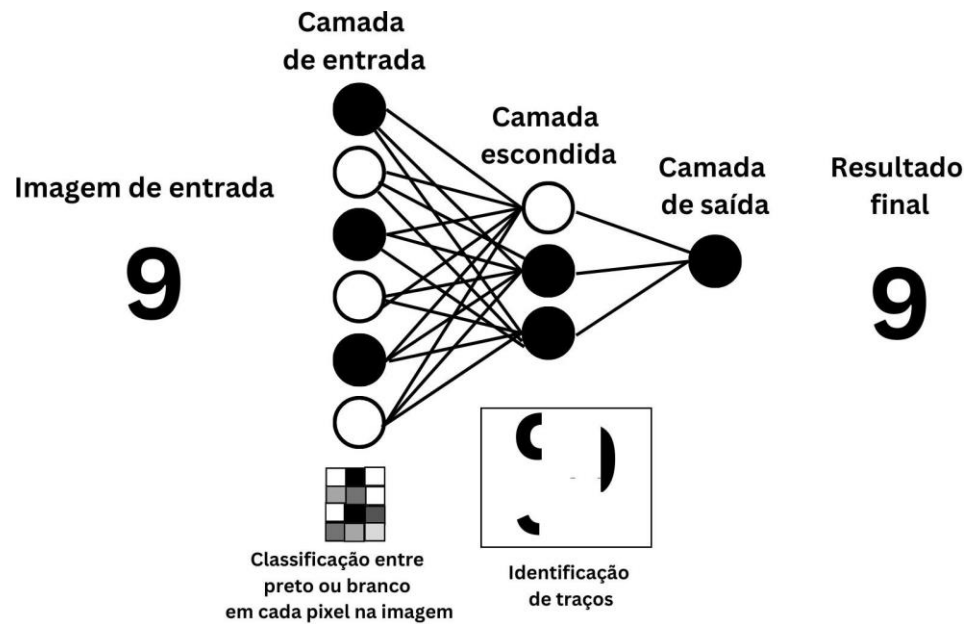


Figura 4 - Arquitetura geral das redes neurais. Na imagem, os dados são apresentados para a rede neural. Neste exemplo, o número 9 é processado pela rede neural. Na primeira camada, a imagem é dividida em pequenas porções ou pixels, onde é então classificado cada pixel em preto ou branco. Na próxima camada, a camada escondida, são identificados traços do número. Esta é uma representação ilustrativa e podem haver mais camadas na rede neural com mais abstrações podendo ser extraídas da imagem. Por fim, a camada de saída dá o resultado final com o número 9. Nós ou neurônios ativados são mostrados em forma circular preta, enquanto que neurônios não ativados em forma circular branca.

1.4.1 Estrutura das redes neurais artificiais

A aprendizagem profunda é inspirada em redes neurais biológicas, sendo composta por multicamadas de nós ou neurônios artificiais. Os neurônios artificiais são ligados entre si por conexões por onde o aprendizado de alguma característica é passado para a próxima camada (NIELSEN, 2015). A estrutura geral das redes neurais segue com camada de entrada, onde os dados são colocados para o algoritmo, camadas escondidas que processam os dados e por fim a camada de saída, onde o resultado final é fornecido pelo modelo (NIELSEN, 2015).

O funcionamento da rede neural pode ser entendido pelo modelo mais simples com o neurônio perceptron. Desenvolvido por Frank Rosenblatt, a rede tem várias entradas binárias com uma única saída binária (TAPPERT, 2019). Neste conceito foi introduzido os pesos que são valores numéricos que mostram o valor ou importância de cada valor de entrada para o neurônio subsequente, neste caso o neurônio de saída. O neurônio é ativado após o valor

ultrapassar o limiar pré-estabelecido, sendo percebido pelo bias do neurônio ao qual indica o quão permissível o neurônio está para ser ativado (NIELSEN, 2015).

No início da aprendizagem, a rede neural recebe o conjunto de dados e passa pela etapa de treinamento, onde o algoritmo ajusta os pesos para cada conexão. O aprendizado da rede neural consiste nas mudanças de pesos e bias a fim de que o modelo aprenda finalmente os resultados esperados para os dados fornecidos. Em seguida, o desenvolvimento da rede neural passa pela etapa de validação para a verificação da robustez de sua aprendizagem, onde outra fonte de dados é fornecida. Com os resultados corretos, os desenvolvedores podem ter a certeza que o algoritmo estará pronto para uso (NIELSEN, 2015).

1.4.2 Aplicação do *Deep Learning* em problemas de biologia computacional

Com a presença de conjunto de dados mais largos e complexos nos dias atuais, a interpretação manual vem necessitando de auxílio de ferramentas mais sofisticadas. O “boom” da inteligência artificial tem entrado na área científica com destaque na genômica, onde a aprendizagem profunda tem sido utilizada para resolução de questões biológicas (MAHMUD *et al.*, 2018).

Desde a elucidação da estrutura proteica, um obstáculo até então era pertinente entre a comunidade científica: a predição da estrutura tridimensional da proteína a partir da informação de sua sequência de aminoácidos. Desde a criação da CASP⁷, os pesquisadores têm utilizado uma diversidade de ferramentas de bioinformática com abordagens diferentes para resolução desta questão como SWISS-MODEL que usa modelagem por homologia e I-TASSER que se baseia no estado mínimo de energia para a conformação da proteína (ROY; KUCUKURAL; ZHANG, 2010; WATERHOUSE *et al.*, 2018).

Um dos triunfos do *deep learning* tem sido o Alpha Fold 2, desenvolvido pela empresa DeepMind que na CASP14 conseguiu com exatidão elucidar as estruturas 3D de proteínas de não conhecimento público. O desempenho do Alphafold na competição pode ser um marco na história, tendo um grande impacto na descoberta de drogas e avanço no tratamento de doenças (JUMPER *et al.*, 2021; PEREIRA *et al.*, 2021).

O *Deep Learning* tem sido proficiente quanto à predição de estruturas secundárias e motivos de proteínas com PsiPred e Prosite, respectivamente (BUCHAN; JONES, 2019; SIGRIST *et al.*, 2010). Além disso, modelos de predição tem se estendido à análise do efeito

⁷ do inglês: Critical Assessment of Techniques for Protein Structure Prediction.

de variantes de nucleotídeo único em elementos regulatórios com o DeepSEA (ZHOU; TROYANSKAYA, 2015).

Em outro aspecto, no campo da descoberta de novas drogas, o lançamento de novos antibióticos tem entrado em queda ao decorrer dos anos. Entretanto, a aprendizagem profunda pode mudar o rumo do mercado de antibióticos: Stokes et. al (2020) demonstraram que a halicina, uma droga envolvida em testes iniciais de efeito contra diabetes, na verdade pode ser um potencial antibiótico com novo mecanismo de ação. A capacidade antimicrobiana da halicina foi descoberta através do uso de *deep learning* que identificou esta molécula entre um repositório de 6.111 compostos químicos. A droga demonstrou efeito bactericida e bacteriostática contra diversas espécies bacterianas, entre elas, *Mycobacterium tuberculosis* e *Staphylococcus aureus*, tanto *in vitro* quanto *in vivo*.

1.4.2.1 O caso DeepARG

A tecnologia do aprendizado profundo pode ser aplicada na predição de genes de resistência. É isto o que propõe o algoritmo DeepARG, desenvolvido por Arango-Argoty *et al.* (2018). O algoritmo é capaz de prever não apenas ARGs já conhecidos, como também, novos genes de resistência que não seriam identificados por métodos convencionais de bioinformática (ARANGO-ARGOTY *et al.*, 2018).

O algoritmo apresenta alto desempenho quando apresentado a um banco de dados amplo com várias categorias de genes de resistência, por onde apresenta uma alta precisão seguida de uma baixa taxa de falsos negativos. Com o aprendizado de padrões dos perfis através das sequências dos genes, o DeepARG é habilitado a identificar genes com sequências incompletas (DeepARG-SS) e sequências completas (DeepARG-LS) (ARANGO-ARGOTY *et al.*, 2018).

Em resumo, o desenvolvimento do banco de dados do DeepARG, DeepARG-DB foi construído a partir das coleções de genes do CARD⁸, ARDB⁹ e UNIPROT (ALCOCK *et al.*, 2023; BATEMAN, 2019; LIU; POP, 2009). Esta coleção de genes passou por processamento, onde sequências idênticas foram retiradas com o auxílio do CD-HIT (FU *et al.*, 2012). As sequências de genes presentes no UNIPROT que são menos curadas, passaram por processo de inspeção com mineração de texto e escolha de genes com alta identidade a nível de DNA

⁸ Banco de Dados Abrangente de Resistência a Antibióticos.

⁹ Banco de Dados de Genes de Resistência a Antibióticos.

com comparação do ARDB e CARD realizada pelo programa DIAMOND (BUCHFINK; XIE; HUSON, 2014).

Portanto, 14.933 genes que compõem o banco de dados foram usados para a construção do modelo, sendo os genes do UNIPROT usados para a validação e treinamento. A construção do algoritmo passou pelas etapas de caracterização, treinamento e validação (Figura 5) (ARANGO-ARGOTY *et al.*, 2018).

Na caracterização, as sequências de DNA foram representadas em formas numéricas e foram classificadas por dissimilaridade: a distância de identidade para ARGs conhecidos do CARD e ARDB, sendo genes atribuídas com valor de 0 representam aquelas com alta identidade e baixa distância e 1 sendo o inverso. O bit-score foi usado como indicador de similaridade e depois foi normalizado para o alcance de pontuação entre 0 e 1 (ARANGO-ARGOTY *et al.*, 2018).

A rede neural do DeepARG contém uma camada de entrada com 4.333 unidades para os ARGs do ARDB e CARD, 4 camadas escondidas com 2000, 1000, 500, 100 unidades e uma camada de saída com 30 unidades para cada classe de antibiótico. Por fim, para a camada de saída foi computado a probabilidade para sequência ser de determinada categoria de ARG. O modelo foi supervisionado, ou seja, as decisões tomadas para identificação de cada sequência pelo algoritmo foram avaliadas passo a passo para a certificação de sua acurácia (ARANGO-ARGOTY *et al.*, 2018).

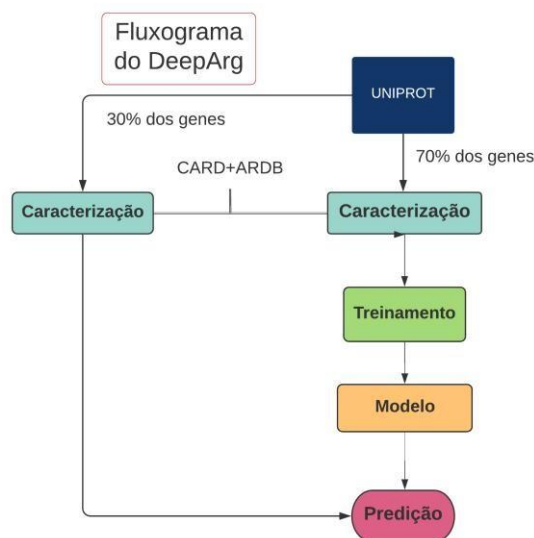


Figura 5 - Fluxograma de funcionamento do DeepARG. A caracterização foi realizada com genes presentes do UNIPROT com os genes do CARD e ARDB representando ARGs positivos. Em relação à validação do modelo, o banco de dados MEGARes foi utilizado.

Pesquisas envolvidas na resistência antimicrobiana têm se utilizado de ferramentas da bioinformática para a identificação de genes de resistência após o sequenciamento genômico (BORELLI *et al.*, 2021). A grande maioria dos softwares e serviços disponíveis na internet é baseada em métodos na comparação das sequências de interesse com as sequências de genes de resistência que foram curadas em bancos de dados (HENDRIKSEN *et al.*, 2019).

As ferramentas se utilizam do BLAST, onde geralmente é realizado um alinhamento entre as sequências e então verificado a homologia (MCGINNIS; MADDEN, 2004). Apesar de a predição possuir uma alta identidade com os parâmetros de corte que vão de 80 a 90%, essa abordagem apresenta desvantagens que residem na falha da detecção de genes de resistência que estejam abaixo do limite estabelecido (ARANGO-ARGOTY *et al.*, 2018).

Logo, potenciais genes de resistência são apontados erroneamente como genes que não conferem resistência, gerando um número maior de falsos negativos (ARANGO-ARGOTY *et al.*, 2018). Outra desvantagem consequente é que o pesquisador terá uma falsa impressão dos resultados, pois não conseguirá ter uma compreensão real do genoma que pode conter novos genes de resistência. Por outro lado, os algoritmos baseados em aprendizado profundo não são baseados na metodologia BLAST, pois se utilizam de muitas características de quantidades enormes de dados para produzir modelos preditivos que se baseiam nos padrões encontrados neles. Isso dispensa a curadoria manual do analista, que é necessária em modelos de aprendizado de máquina, tornando a busca por padrões em dados complexos, como sequências de genes de resistência a antimicrobianos, mais apropriada (ARANGO-ARGOTY *et al.*, 2018).

Segundo Arango-Argoty *et al.* (2018), o algoritmo DeepARG se demonstra ideal para estudos de predição de genes, pois é capacitado a identificar com alta precisão e uma baixa taxa de falsos negativos, novos genes de resistência e genes com baixa homologia. Entretanto, esta abordagem de aprendizado profundo ainda carece de estudos empíricos relacionada à identificação de genes de resistência (HENDRIKSEN *et al.*, 2019). Portanto, este projeto teve o objetivo de identificar potenciais genes de resistência a antibióticos em genomas de microrganismos patogênicos isolados no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HC-RP) mediante o uso do algoritmo de aprendizado profundo DeepArg e da caracterização funcional dos genes selecionados contra diversas classes de antibióticos.

2 OBJETIVOS

2.1 Objetivo geral

Identificar e caracterizar funcionalmente potenciais genes de resistência a antimicrobianos em genomas de patógenos isolados em ambiente hospitalar.

2.2 Objetivos específicos

- I) Selecionar as sequências de nucleotídeos de genes em formato FASTA relacionados com resistência antimicrobiana presentes em plasmídeos com o auxílio do algoritmo DeepARG;
- II) Classificação *in silico* das proteínas selecionadas e comparação dos resultados com a literatura acadêmica relacionada;
- III) Construção de circuitos gênicos no vetor pUC19 mediante síntese das sequências de interesse para a caracterização funcional em *E. coli* DH10B;
- IV) Caracterização funcional dos potenciais genes de resistência a antibióticos em *E. coli* DH10B em ampla diversidade de classes de antibióticos.

3 MATERIAIS E MÉTODOS

3.1 Panorama geral da metodologia do projeto

Na **Figura 6** é mostrada uma contextualização geral da metodologia do projeto. Nos subtópicos subsequentes serão mostrados a metodologia de maneira mais detalhada.

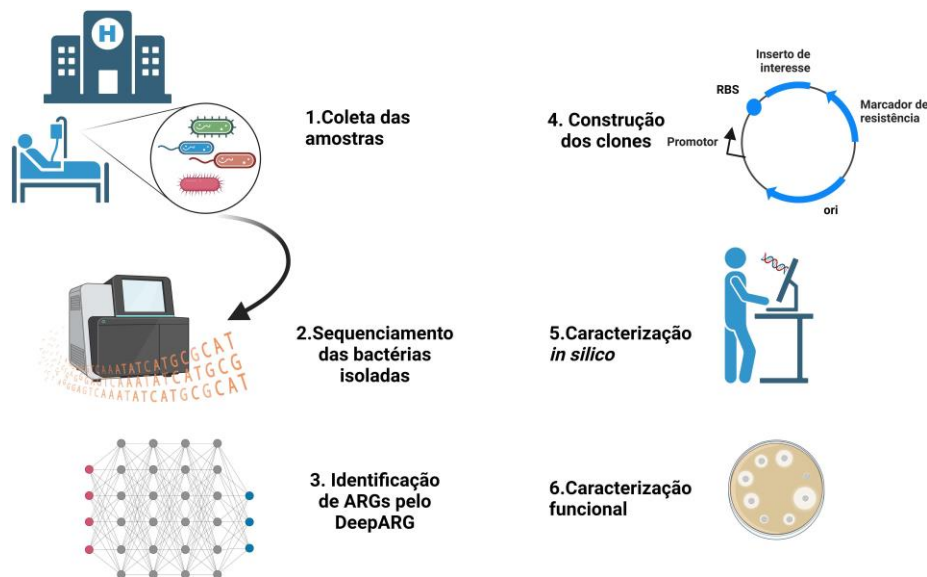


Figura 6 - Esquema geral da metodologia do projeto. (1-2) Em um trabalho anterior, 34 bactérias patógenicas foram isoladas no hospital das clínicas de Ribeirão Preto e em seguida, seus genomas foram sequenciados conforme detalhado em Borelli *et al.* (2021). 3- Após isso, os ARGs foram identificados pelo DeepARG. 4- As sequências mais promissoras foram selecionadas e clonadas em vetores para a caracterização funcional. 5- Uma caracterização *in silico* das sequências selecionadas foi realizada com o auxílio de ferramentas de bioinformática. 6- Para a caracterização funcional dos clones contendo as sequências, teste de disco-difusão e ensaio de gotejamento foram realizados. A imagem foi gerada no site Biorender.com.

3.2 Identificação de genes de resistência a antibióticos pelo DeepARG

A partir da coleção de genomas de bactérias patógenicas, as sequências identificadas em plasmídeos pelo programa Plasflow foram submetidas ao DeepARG para a identificação de potenciais genes resistentes a antibióticos com probabilidade maior de 0.8 em serem classificadas como ARGs para determinada classe de antibiótico (KRAWCZYK; LIPINSKI; DZIEMBOWSKI, 2018). Esta parte do trabalho foi realizada em colaboração com o Informático Biomédico Felipe Marcelo Pereira dos Santos. Logo, os genes identificados pelo DeepARG foram submetidos à ferramenta ORFfinder¹⁰ (NCBI) para identificação de janelas de leitura aberta. Em seguida, realizou-se a busca de sequências idênticas e mais completas no

¹⁰ Disponível em: <https://www.ncbi.nlm.nih.gov/orffinder/>

banco de dados do NCBI pelo programa BLAST no modo BLASTp com os parâmetros padrão em relação às aquelas identificadas pelo DeepARG.

Os potenciais ARGs foram selecionadas com base em baixa identidade a nível de nucleotídeo pelo DeepARG e com alta probabilidade em serem classificadas segundo a resistência a determinada classe de antibiótico. Nos resultados do BLAST, os parâmetros considerados foram a cobertura, identidade, *e-value* e também o tamanho da sequência em aminoácidos, sendo selecionada aquela com o maior número de resíduos de aminoácidos.

Após a pré-seleção dos potenciais ARGs, dados do DeepARG foram comparados com os dados clínicos para a exclusão de potenciais falsos positivos entre as sequências pré-selecionadas.

3.3 Estratégia experimental adotada para caracterização funcional dos genes selecionados

O constructo genético foi esquematizado da seguinte forma: a sequência de interesse foi inserida no vetor pUC19 entre os sítios de restrição EcoRI e HindIII em sentido contrário ao promotor próprio do vetor, o P_{lac} . Esta mudança de sentido foi realizada para que o P_{lac} não tivesse controle da expressão frente as sequências de interesse. Adicionalmente, para um controle maior da expressão foram inseridas sequências do promotor constitutivo P_{j106} com força média de expressão do gene e o sítio de ligação ao ribossomo (RBS) B0034 com força média de ligação (**Figura 7**). Ambas as sequências foram retiradas das partes colecionáveis do iGEM¹¹.

¹¹ Informações sobre o RBS BB0034 disponível em: http://parts.igem.org/Part:BBa_B0034
Informações sobre o promotor P_{j106} disponível em: http://parts.igem.org/Part:BBa_J23106

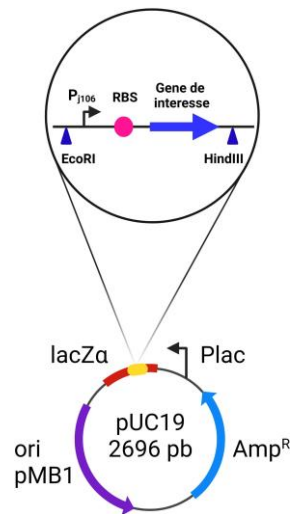


Figura 7 - Esquematização da construção gênica realizada no pUC19 para a caracterização funcional.

Fonte: A imagem foi gerada no site Biorender.com.

Além das sequências selecionadas, o gene KPC-2 que foi identificado pelo DeepARG foi incluído no estudo como um controle positivo para a caracterização funcional e teve a mesma construção gênica que as demais sequências. Em seguida, as construções gênicas contendo as sequências de interesse foram demandadas para a empresa IDT para suas sínteses. Antes da caracterização funcional, os vetores contendo as sequências de interesse foram suspensos em água e seguiu-se a preparação de bactérias *E. coli* DH10B quimiocompetentes com cloreto de cálcio. Após a transformação por choque térmico na bactéria *E. coli* DH10B¹², as bactérias clonadas foram cultivadas em meio Luria Bertani contendo ampicilina (150 µg/mL) e foram estocadas a -80°C em 20% de glicerol. Para a confirmação da presença das sequências no vetor pUC19, inóculos bacterianos contendo as sequências de interesse foram centrifugados e passaram pela mini-preparação de plasmídeo bacteriano conforme o protocolo da empresa Promega¹³ utilizando o kit Wizard® Plus SV. Em seguida, o DNA plasmidial foi quantificado no Nanodrop One^c e a digestão enzimática foi aplicada seguindo as orientações do fabricante. Brevemente, a reação contendo o DNA plasmidial foi incubada em Banho-Maria com as enzimas EcoRI-HF e HindIII por 1 hora a 37 °C, sendo a reação interrompida em seguida a 80 °C durante 15 minutos. Em seguida, as sequências puderam ser constatadas pelo gel de agarose após a eletroforese.

¹² Protocolo disponível em: <https://www.addgene.org/protocols/bacterial-transformation/>

¹³ Protocolo disponível em: https://www.promega.com.br/-/media/files/resources/protcards/wizard-plus-sv-minipreps-dna-purification-system-quick-protocol.pdf?rev=8293c52acfa4de495f1df90a7753d50&sc_lang=pt-br

3.4 Caracterização *in silico* das sequências selecionadas

Para um maior entendimento da função celular das sequências de proteínas selecionadas, as proteínas foram submetidas a análise no PFAM — Banco de dados de famílias de proteínas¹⁴ e no CDD *domain search* (Busca de domínios conservados do NCBI) para a identificação de domínios, a busca de motivos proteicos e peptídeo sinal foi realizada pelo Prosite e pela ferramenta SignalP 5.0, respectivamente (ALMAGRO ARMENTEROS *et al.*, 2019; FINN *et al.*, 2014; LU *et al.*, 2020). A partir das análises de bioinformática, buscava-se avaliar as funções das proteínas de acordo com os resultados das predições da estrutura secundária.

3.5 Teste de difusão de disco para os clones contendo as sequências selecionadas

Antes da avaliação da capacidade de conferir resistência das quatro sequências selecionadas aos antibióticos, a possível interferência do gene de resistência a ampicilina presente intrinsecamente no vetor pUC19 foi avaliada. Para isso, o teste de disco de difusão foi executado para comparar o perfil de resistência entre a bactéria *E. coli* DH10B sem vetor, *E. coli* DH10B com o vetor pUC19 vazio e a mesma bactéria com o gene KPC-2 inserido no vetor pUC19.

Para os clones bacterianos contendo as sequências selecionadas, o teste de difusão de disco foi realizado de acordo com as orientações do Comitê Brasileiro de Teste de Sensibilidade aos Antimicrobianos — BrCAST (BRCAST, 2021). Em suma, um cultivo de 24 horas das bactérias foi inoculado em solução salina de NaCl a 0,85%. Em seguida, as bactérias foram plaqueadas em meio Mueller-Hinton com ampicilina (150 µg/mL) após a verificação da densidade óptica (DO) a 625 nm no intervalo de 0.08-0.13 e então os discos de antibióticos foram inseridos em contato com o meio de cultura.

A observação da resistência aos antibióticos das bactérias portadoras dos vetores com o gene de interesse foi feita em comparação com a bactéria com vetor pUC19 vazio. O antibiograma com a medição dos halos de inibição foi realizado 24 horas após a incubação das bactérias a 37°C. O desenho experimental do teste de difusão de disco para as sequências selecionadas teve um amplo espectro de antibióticos de uso clínico para serem testados incluindo beta-lactâmicos (Cefepime, Ceftriaxona, Aztreonam, Imipenem e Meropenem),

¹⁴ O site Pfam agora está sendo hospedada no site da InterPro.

fluoroquinolonas (Ciprofloxacina e levofloxacina), tetraciclina (tetraciclina e doxiciclina), aminoglicosídeos (gentamicina), polipeptídeos (polimixina B), fosfomicina (fosfomicina) e sulfonamidas (Bactrim). O teste de disco de difusão foi realizado com mais de três réplicas biológicas. As imagens do antibiograma foram geradas pelo equipamento ChemiDoc™ com bandeja branca (590/110) e tempo de exposição automático.

3.6 Teste de resistência a antibióticos por ensaio de gotejamento (*drop assay*)

O ensaio de gotejamento foi baseado na metodologia de McGarvey, Queitsch e Fields (2012) com algumas modificações sendo realizadas. O inóculo bacteriano foi cultivado *overnight* a 37°C e 220 rpm em meio LB com adição de ampicilina (150 µg/mL).

Em seguida, o crescimento do inóculo foi avaliado por DO a 600 nm e realizada uma diluição para uma DO de 0,05 ao que se remete a fase lag. Então seguiu-se uma diluição seriada até 10^{-5} após a bactéria chegar a uma DO de 0,5 nas mesmas condições de meio, temperatura e agitação que o cultivo *overnight*.

Em uma placa Petri de 120 mm com meio LB adicionado com antibiótico teste foram gotejadas 10 µL de cada diluição, assim como, a cultura não diluída. Os resultados foram analisados após 24 horas de incubação a 37°C (**Figura 8**). Este ensaio foi realizado com duas réplicas biológicas e foram considerados a concentração inibitória mínima (CIM) para cada antibiótico testado nos trabalhos realizados anteriormente no laboratório e por McGarvey, Queitsch e Fields (2012). Os antibióticos testados e as respectivas concentrações utilizadas foram: Ampicilina (150 µg/mL), Imipenem (0.5 µg/mL), Gentamicina (10 µg/mL), Colistina (0.125 µg/mL), Canamicina (20 µg/mL), Levofloxacina (0.03 µg/mL), Ciprofloxacina (0.016 µg/mL), Cloranfenicol (10 µg/mL) e Sulfadiazina (128 µg/mL). As imagens foram geradas pelo equipamento ChemiDoc™ com bandeja branca (590/110) e tempo de exposição automático.

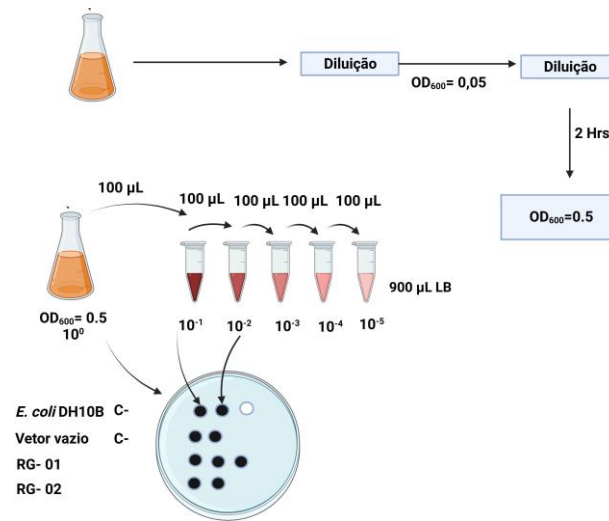


Figura 8 - Esquema geral do ensaio de gotejamento. Fonte: Imagem gerada no site Biorender.com

3.7 Avaliação do *burden* metabólico dos clones contendo as sequências selecionadas

A avaliação de um possível *burden* metabólico nos clones foi observado pelo crescimento bacteriano. A partir da preparação de inóculos *overnight* em meio M9, os clones bacterianos foram diluídos até uma DO de 0.5. Em seguida, os clones foram plaqueados em uma placa de 96 poços até a DO de 0.05 com acréscimo de ampicilina (150 µg/mL), exceto para a *E. coli DH10B* que não teve nenhuma adição de antibiótico. O crescimento bacteriano durante 8 horas foi observado no equipamento Victor Nivo e o experimento foi realizado com 4 réplicas biológicas e 3 réplicas técnicas. Para este teste foram considerados os clones contendo as sequências de interesse, incluindo o clone bacteriano contendo o gene KPC-2, assim como, *E. coli DH10B* sem e com vetor pUC19 (**Figura 9**).

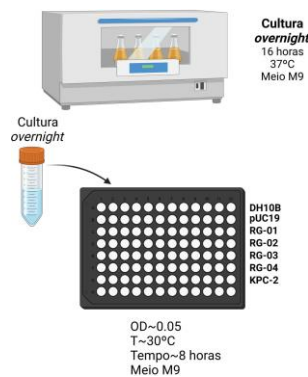


Figura 9 - Metodologia da avaliação do *burden* metabólico para os clones contendo as sequências selecionadas.

Fonte: Imagem gerada no site Biorender.com

4 RESULTADOS

4.1 Identificação de genes de resistência a antibióticos pelo DeepARG

Em um panorama geral, o DeepARG identificou uma maior quantidade de genes referentes a multidrogas entre a coleção de genomas bacterianos, seguida de genes resistentes a beta-lactâmicos (**Figura 10**).

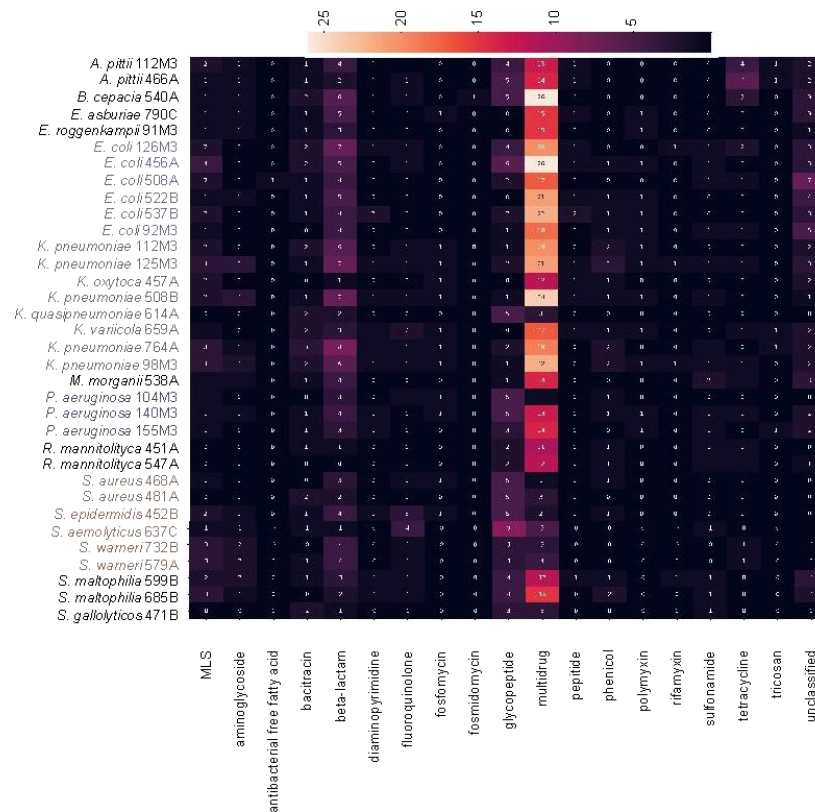


Figura 10 - Genes de resistência a antibióticos identificados pelo DeepARG. Fonte: Imagem retirada de Borelli *et al.* (2021). O mapa de calor mostra a quantidade de ARGs de cada classe identificado pelo algoritmo DeepARG a partir das 35 linhagens de patógenos incluídos no estudo.

Após a identificação de sequências de genes em plasmídeos, seguiu-se a pré-seleção das sequências. Entre a coleção de 30 genes identificados pelo DeepARG, cinco sequências com baixa identidade (30.0-41.1) foram pré-selecionadas com base no resultado do BLAST (**Tabela 1**). Em relação ao resultado do BLAST, as sequências apresentam boa identidade e cobertura, assim como um valor de *e-value* nos parâmetros aceitáveis. Entretanto, uma das sequências que codifica uma proteína resistente a cobre foi excluída após a avaliação do

último critério: o sentido biológico dos resultados dado pelo DeepARG comparados com os dados clínicos.

Tabela 1 - Resultados do BLAST das sequências pré-selecionadas.

Sequência escolhida	Cobertura	Identidade	<i>e-value</i>	Acesso NCBI
RG-01	72%	95%	9e-37	NKA09117.1
RG-02	99%	100%	3e-84	SWP25475.1
RG-03	99%	93.51%	1e-32	AEQ15826.1
RG-04	98%	95,35%	1e-19	AFC60898.1
Sequência X	100%	100%	1e-42	WP_063394441.1

Fonte: Autoral.

Como visto na **Figura 11**, o potencial ARG presente na *R. mannitolilytica* 451A foi atribuída como resistente a glicopeptídeo, uma classe de antibiótico efetiva contra bactérias Gram-positivas, mas não a Gram-negativas. Como identificada pelo sequenciamento de seu genoma, a bactéria *R. mannitolilytica* é uma bactéria Gram-negativa, logo, a presença de um gene resistente a glicopeptídeo em seu genoma não teria sentido biológico.

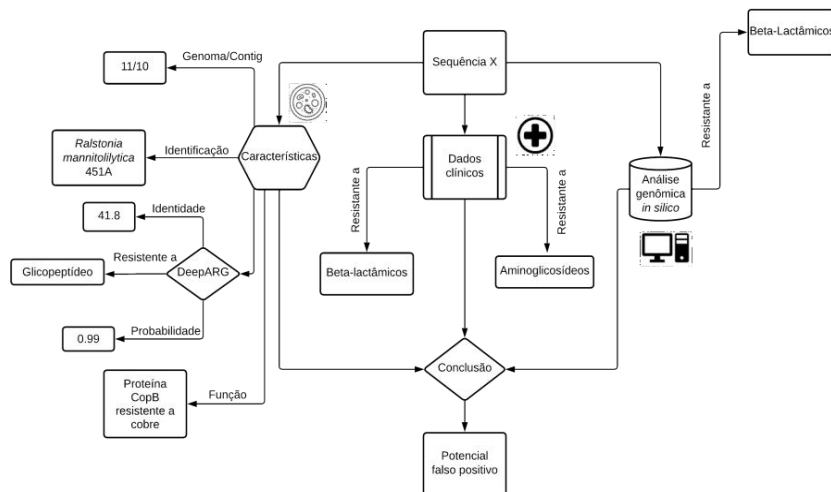


Figura 11 - Panorama geral de um potencial ARG falso positivo identificado pelo DeepARG. A sequência intitulada Sequência X foi excluída após a provável conclusão de se tratar de um falso positivo nas análises. A figura mostra os resultados apresentados pelo DeepARG e BLAST, dados clínicos e análise genômica realizada pelo ABRicate.

Ao analisar os dados clínicos, a bactéria se mostrou suscetível a vancomicina, um antibiótico representante do grupo dos glicopeptídeos, o que não era de se esperar pelo ARG identificado pelo DeepARG resistente a esta classe de antibióticos. Ao analisar o artigo dos desenvolvedores do DeepARG, a classe de genes resistentes a glicopeptídeos tem uma representação baixa de genes para o treinamento do modelo (ARANGO-ARGOTY *et al.*, 2018). Devido a estas questões, a sequência foi compreendida como um potencial ARG falso positivo e, portanto, foi excluída da caracterização funcional.

Apesar da seleção de sequências presentes em plasmídeos, o que reduz o número de potenciais ARGs a serem identificados, esperava-se a seleção de pelo menos 10 potenciais ARGs. Entretanto, o objetivo não foi alcançado devido a alguns fatores a serem destacados como: presença de sequências identificadas com alta identidade a ARGs conhecidas, identificação de não ARGs como transposases e fatores de virulência, sequências com baixa cobertura e/ou baixa identidade e por fim, sequências identificadas pelo BLAST com resistência a antibióticos já estabelecida na literatura como bomba de efluxo. Além do gene positivo verdadeiro (KPC-2) escolhido como controle positivo, 4 sequências foram selecionadas para a caracterização funcional (**Tabela 2 e Anexo A**).

Tabela 2 - Características gerais das sequências de genes selecionadas.

Gene	Identificação da espécie (Dados clínicos)	Classe de antibiótico (DeepARG)	Identificação do gene (BLAST)	Identificação da espécie (NCBI)	Tamanho da proteína (Query/Subject)
RG-01	<i>R. mannitolilytica</i> 451A	Multidrogas	Proteína hipotética	<i>R. solanacearum</i>	93aa/480aa*
RG-02	<i>E. coli</i> 659A	Fluoroquinolona	Proteína da família da Fosfolipase D	<i>Klebsiella pneumoniae</i>	126aa/186aa
RG-03	<i>E. coli</i> 508A	Multidrogas	Proteína hipotética	<i>E. coli</i>	77aa/177aa
RG-04	<i>E. coli</i> 537B	Não classificado	Proteína hipotética	<i>E. coli</i>	43aa/85aa

*aa, aminoácidos. Fonte: Autoral.

Dentre as sequências selecionadas, o DeepARG identificou duas sequências resistentes a categoria multidroga, uma sequência resistente a fluoroquinolona e uma sequência não foi atribuída a nenhuma classe dos 30 antibióticos testados no modelo do DeepARG. As sequências selecionadas no NCBI são dos mesmos gêneros que os identificados pelos dados genômicos: *Ralstonia*, *Klebsiella* e *Escherichia*. Dentre as quatro sequências, três são proteínas hipotéticas (RG-01, RG-03 e RG-04), ou seja, com necessidade de caracterização funcional e uma proteína, a RG-02, foi atribuída a uma proteína da família da Fosfolipase D.

Apesar da proteína ser bastante caracterizada na literatura como membros com funcionalidade de endonucleases ou proteínas de membrana com função na sinalização celular, não há dados literários atribuindo sua funcionalidade com resistência a antibióticos (FROHMAN, 2015). Portanto, esta sequência foi selecionada para a caracterização funcional.

4.2 Caracterização *in silico* das proteínas de resistência a antibióticos

Com o objetivo de conhecer melhor as sequências selecionadas, a caracterização *in silico* foi realizada com a identificação de domínios ou motivos protéicos das sequências selecionadas. Com a identificação de motivo, esperava-se algum indício que pudesse identificar a família da proteína e com isso estabelecer a sua função protéica.

Entretanto, para a maioria das sequências, exceto a RG-02, não identificado nenhum domínio ou motivo protéico, assim como para o peptídeo sinal (**Tabela 3**). A sequência RG-02 representa a proteína Nuc, membro da família Fosfolipase D. A proteína possui um domínio Pld com sítio ativo catalítico com um motivo HKD. Apesar de possuir um peptídeo sinal que pode indicar sua localização na superfície celular da bactéria, membros desta família podem exercer atividade de endonuclease com cortes não específicos no DNA (SELVY *et al.*, 2011).

Tabela 3 - Resultado da caracterização *in silico* das proteínas de resistência a antibióticos

Proteína	Domínio protéico	Motivo protéico	Peptídeo sinal	Acesso NCBI
RG-01	Não encontrado	Não encontrado	Não encontrado	NKA09117.1
RG-02	Domínio Pld	Perfil sítio ativo Phospholipase D phosphodiesterase	Peptídeo sinal (Sec/SP1)	SWP25475.1
RG-03	Não encontrado	Não encontrado	Não encontrado	AEQ15826.1
RG-04	Não encontrado	Não encontrado	Não encontrado	AFC60898.1

Fonte: Autoral.

4.3 Teste de difusão de disco para as sequências selecionadas

Antes da realização do teste de difusão de disco com as sequências selecionadas, o efeito do gene *bla* presente no vetor pUC19 foi avaliado frente a antibióticos beta-lactâmicos. De acordo com a literatura, este gene presente em uma bactéria confere resistência a ampicilina, um antibiótico do subgrupo aminopenicilina. Entretanto, o mesmo é ineficaz contra antibióticos de subgrupos mais potentes como cefalosporinas e carbapenêmicos (GOH; GOOD, 2008).

Como observado na **Figura 12**, linhagens *E. coli* DH10B sem vetor e com vetor vazio (pUC19) tiveram o perfil de resistência similar para os antibióticos beta-lactâmicos testados. Porém, o clone de *E. coli* DH10B contendo o gene de resistência KPC-2 obteve um decréscimo expressivo da zona de inibição. O gene KPC-2 é uma beta-lactamase que cliva a estrutura química do beta-lactâmico e, portanto, confere resistência a subgrupos como cefalosporinas e carbapenêmicos (BARNES *et al.*, 2017).

Desse modo, os resultados demonstram que o gene *bla* do vetor pUC19 não confere resistência aos grupos de antibióticos cefalosporina, monobactâmicos e carbapenêmicos. Este resultado testifica que se fosse constatada a resistência a beta-lactâmicos com as sequências

selecionadas não seria devido ao gene *bla*, mas a própria sequência de interesse inserida no vetor pUC19.

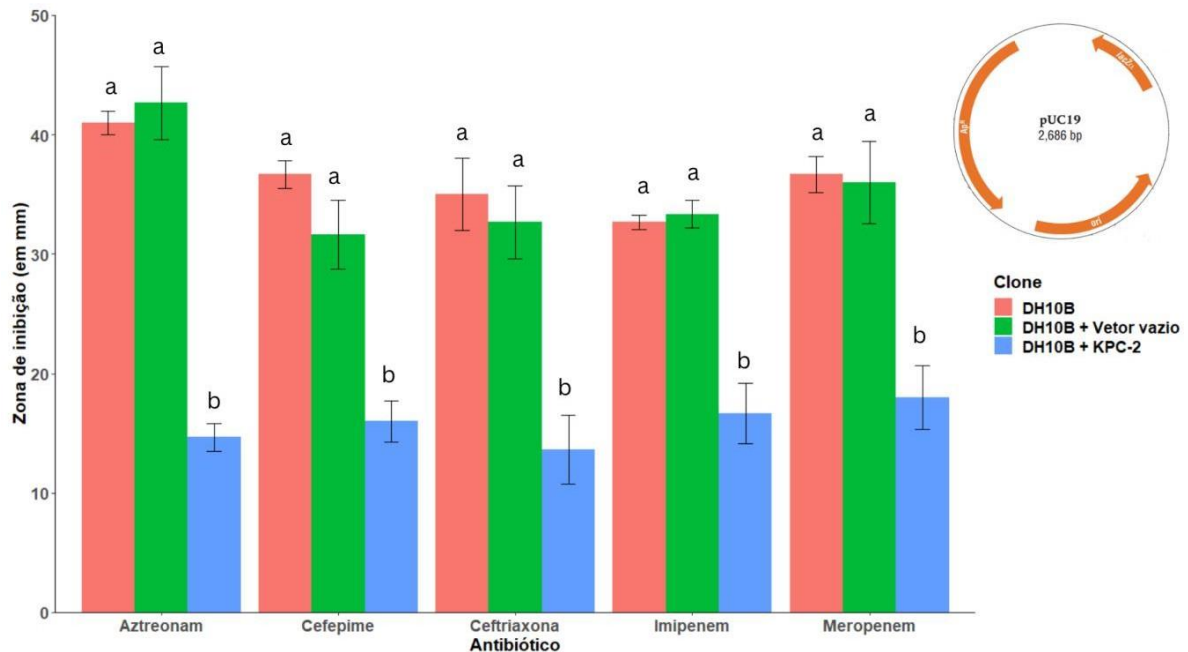


Figura 12 - Avaliação da resistência antimicrobiana do vetor pUC19 frente a antibióticos beta-lactâmicos. Teste estatístico considerando o valor de $p \leq 0,05$ foi performed com teste ANOVA com teste Tukey post-hoc. O experimento foi executado com o mínimo de 3 réplicas biológicas. A imagem do plasmídeo foi retirada do site: <https://www.neb.com>. (a-b) - Significância estatística entre os clones em cada antibiótico testado.

Como mostrado na **Figura 13**, quando comparadas com o controle negativo, ou seja, o clone *E. coli* DH10B com vetor vazio, as sequências não conferem resistência à linhagem bacteriana para os antibióticos testados. Apesar de em alguns casos, clones com sequência conferirem estatisticamente intolerância a determinados antibióticos, o valor de p ficou no limite para atribuição da significância.

Como exemplo, a sequência RG-02 quando exposta ao antibiótico Aztreonam mostrou diferença estatística para o controle negativo com a redução da zona de inibição. Entretanto, o valor do p ficou no limite: 0.047. Além disso, vale ressaltar que em um sentido biológico, as sequências não demonstraram visualmente resistência quando comparadas com o perfil demonstrado pelo gene KPC-2 (**ANEXO C**). Ademais, em alguns casos, as sequências tiveram um perfil de zona de inibição contraditório ao esperado com a diminuição da resistência em comparação com o controle negativo.

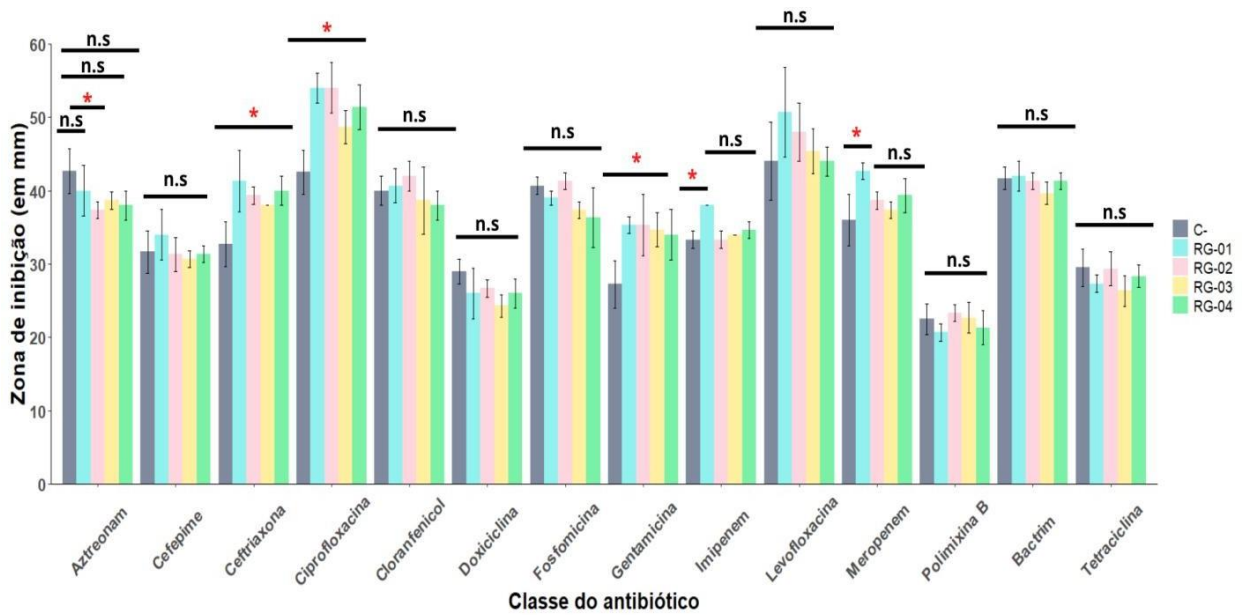


Figura 13 - Antibiograma dos clones contendo as sequências selecionadas. Teste estatístico considerando o valor de $p \leq 0,05$ foi performed com teste t. O experimento foi executado com o mínimo de 3 réplicas biológicas. n.s - Não significante, * - significante.

4.4 Teste de resistência a antibióticos por ensaio de gotejamento (*drop assay*)

Com os resultados negativos do teste de difusão de disco para a resistência a antibióticos dos potenciais ARGs, seguiu-se para o ensaio de gotejamento. Entretanto, as sequências não demonstraram nenhum sinal de resistência à bactéria em comparação com o controle negativo, *E. coli* DH10B com vetor ausente (**Figura 14**).

Apesar disso, o teste se mostrou válido para indicar a resistência a antibióticos, como visto pelo controle positivo com o gene KPC-2. Como também, a presença do vetor foi garantida com o teste para o controle com ampicilina. Para o teste com alguns antibióticos, por exemplo levofloxacina, a presença de colônias foi observada nas primeiras diluições. Porém, este fenômeno provavelmente deve-se à concentração exacerbada de colônias de bactérias, em camadas da superfície onde o antibiótico não consegue alcançar.

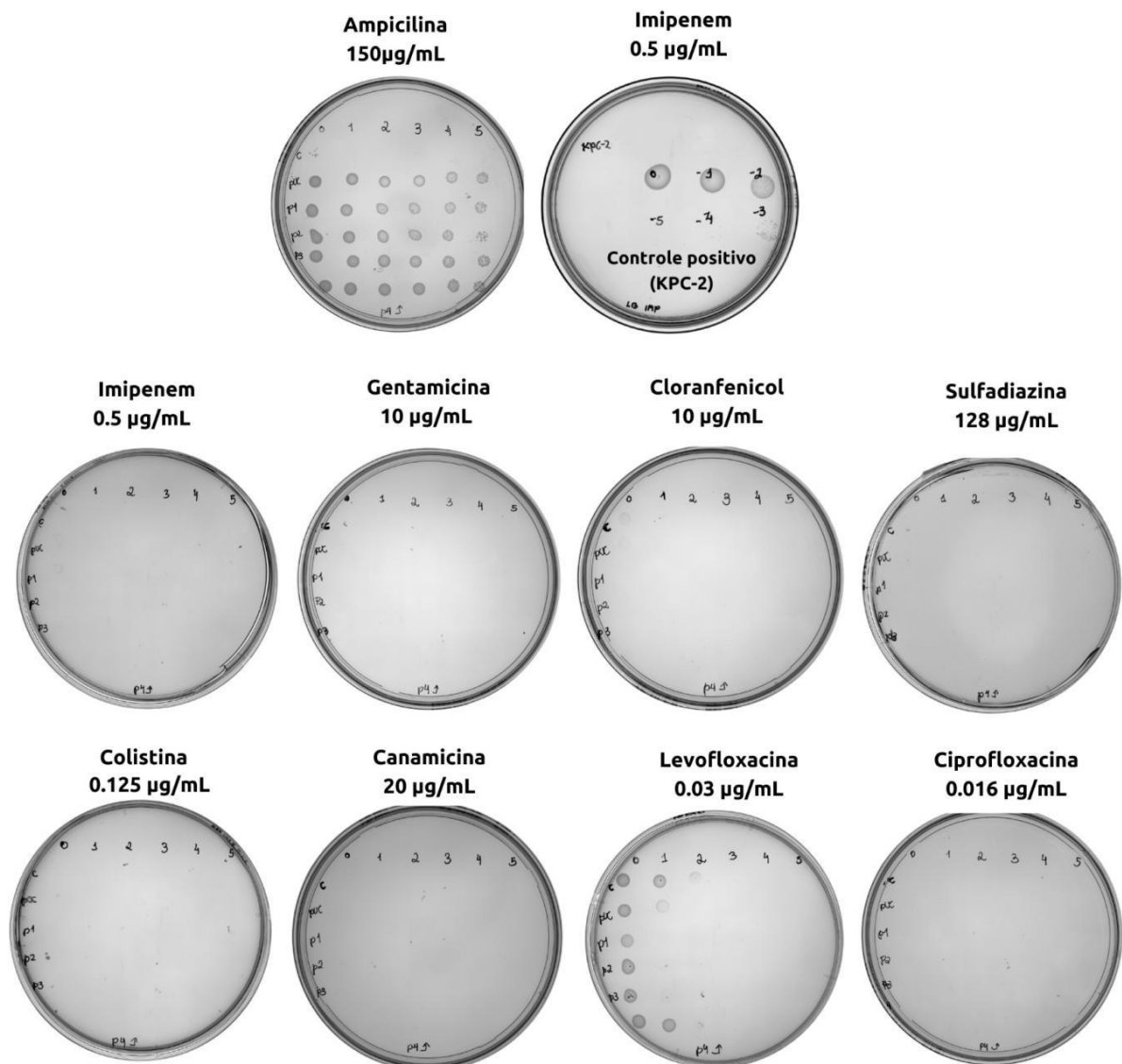


Figura 14 - Resultados do ensaio de gotejamento para as sequências selecionadas. As imagens foram geradas pelo Sistema de Imagens ChemiDoc. Os testes foram realizados com duas réplicas biológicas. Abaixo do nome de cada antibiótico está indicada a concentração ($\mu\text{g/mL}$) utilizada para o teste. C- Controle negativo (*E. coli* DH10B), pUC - *E. coli* DH10B com pUC19 vazio, (p1-p4) - clones bacterianos contendo as sequências de interesse, 0-5 gotejamento de cada diluição seriada em cada linhagem utilizada.

4.5 Avaliação do *burden* metabólico dos clones contendo as sequências selecionadas

A hipótese principal para o resultado observado com as sequências no teste de antibiograma foi o *burden* metabólico. Segundo Wu *et al.* (2016), o *burden* metabólico acontece quando os recursos do metabolismo da célula como moléculas tais como ATP e NADH, os quais são utilizados para a manutenção e crescimento celular são sequestrados para a manutenção do constructo genético engenheirado. Consequentemente, este desvio dos recursos metabólicos leva ao estresse celular. As razões que levam a este pensamento vêm primeiramente do uso do vetor pUC19 que possui um alto número de cópias, assim como, a

inserção e contínua expressão do gene podem ser fatores que implicam no desgaste metabólico para a manutenção celular (AMIRIE *et al.*, 2014; TOH, 2013). Para a elucidação desta hipótese, o crescimento dos clones bacterianos contendo as sequências foi observado ao longo do tempo (**Figura 15**).

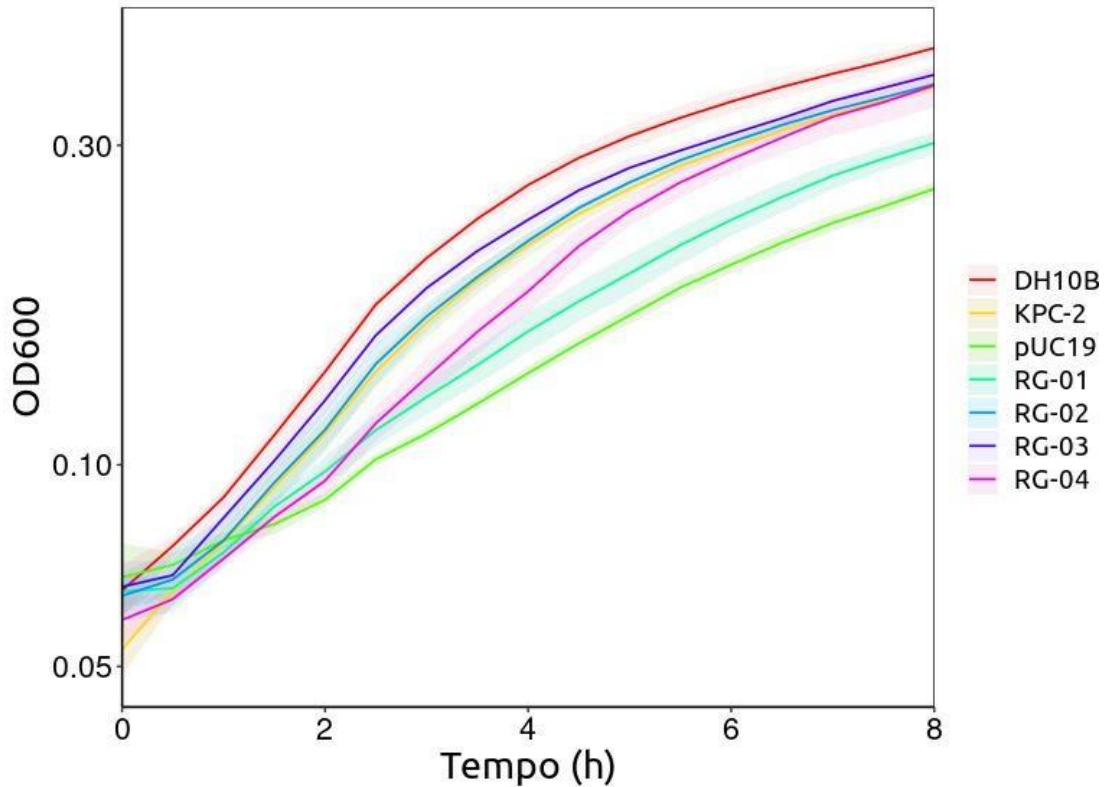


Figura 15 - Curva de crescimento dos clones contendo as sequências de interesse. Na legenda: DH10B - *E. coli* DH10B, pUC19 - *E. coli* DH10B com pUC19 vazio, KPC-2 - clone bacteriano com o gene KPC-2, (RG-01a RG-04) - clone bacteriano com as sequências de interesse.

Entretanto, como observado nos gráficos, os clones contendo as sequências de interesse obtiveram um crescimento similar com o controle *E. coli* DH10B. Com exceção para o cultivo bacteriano contendo a sequência RG-01 que teve o perfil de crescimento similar com *E. coli* DH10B com pUC19 vazio.

5 DISCUSSÃO

Com o objetivo de identificar potenciais genes de resistência em genomas de microrganismos patogênicos isolados em um hospital, o projeto utilizou o DeepARG, um algoritmo de aprendizado profundo para identificação de genes de resistência. Com os 35 patógenos selecionados, o projeto buscou identificar ARGs presentes em plasmídeos, pois a sua identificação nestes elementos genéticos poderia indicar um risco para a disseminação de ARGs entre patógenos no ambiente hospitalar.

Um dos objetivos do projeto era de reportar para bancos de dados como o CARD, a existência destes genes de resistência a antibióticos. Como também, o projeto planejava prosseguir com a caracterização funcional dos genes em outras espécies de bactérias patogênicas, tais como *P. aeruginosa* e *Salmonella typhimurium*. A possível verificação da expressão e atribuição da resistência destes ARGs poderia indicar o risco da disseminação entre patógenos de diferentes espécies. Porém, como observado nos resultados, nenhuma das sequências foram funcionais e, portanto, tais objetivos não foram alcançados.

Em relação às sequências selecionadas: como pode ser observado na Tabela 2, o tamanho das proteínas identificadas pelo DeepARG é menor que as proteínas identificadas pelo NCBI. Este resultado pode ser em decorrência de um dos fatores limitantes do algoritmo que identifica apenas uma parcela do tamanho real do ARG, como também, os genomas não são completos (*draft genomes*), sendo possível então a presença de apenas uma parcela do gene em um *contig*.

Segundo os dados clínicos apresentados pelo Vitek 2, a bactéria *R. mannitolilytica* 451A foi resistente à antibióticos beta-lactâmicos (Piperacilina/Tazobactam, Ceftazidima, Imipenem e Meropenem) e aminoglicosídeos (Gentamicina, Amicacina e Tobramicina), enquanto que a *E. coli* 508A foi resistente a beta-lactâmicos (Ampicilina, Ampicilina/Sulbactam, Cefuroxima e Cefuroxima acetil) e fluoroquinolonas (Ciprofloxacina), além de uma resistência intermediária para cefoxitina, um antibiótico do grupo das cefamicinas. Para estas bactérias, o DeepARG identificou sequências resistentes a multidrogas (RG-01 e RG-03, respectivamente). Como visto em Borelli *et al.* (2021), a resistência a beta-lactâmicos nestas bactérias pode ser atribuída a genes identificados por sequenciamento como genes *bla* e não às supostas sequências identificadas pelo algoritmo.

Diferente de métodos tradicionais de identificação de ARGs baseado em homologia, o DeepARG se utiliza do treinamento em redes neurais a partir de sequências já conhecidas. A diferença pode ser constatada com o panorama geral entre as duas metodologias, em que o

DeepARG mostrou uma maior diversidade de sequência resistentes a multidrogas, enquanto que a metodologia baseada em melhor *hit* teve destaque com a maior presença de sequências resistentes a beta-lactâmicos (Figura 10 e ANEXO B).

De acordo com Arango-Argoty *et al.* (2018), as sequências resistentes a multidrogas são difíceis de identificar por conta das várias informações contidas nos genes desta classe. Além disso, os desenvolvedores apontam as dificuldades do modelo quanto ao processo da validação com o banco de dados MEGARes.

Os resultados mostram que esta classe teve uma boa precisão, mas uma revocação baixa. Isto pode indicar uma maior presença de falsos negativos, ou seja, sequências com resistência a multidrogas foram colocadas como pertencentes a outra categoria. Ademais, isto pode indicar uma necessidade na curadoria de genes de resistência a antibióticos, assim como, uma maior quantidade de genes de resistência para o treinamento: genes de resistência a beta-lactâmicos tiveram o maior percentual com 34%, enquanto que genes de resistência a multidrogas tiveram participação de 5,8% no banco de dados, DeepARG-DB.

No que se refere à bactéria *E. coli* 659A, o Vitek apontou que a espécie foi resistente apenas à ampicilina. Os dados do DeepARG mostraram que havia uma sequência resistente à fluoroquinolona, a RG-02. Todavia, os dados do DeepARG entraram em contradição com os dados clínicos, pois a bactéria foi suscetível à ciprofloxacina. Apesar do erro humano na seleção desta sequência para a caracterização funcional, todavia pôde-se demonstrar a falha do DeepARG em apontar esta sequência como resistente e pode ajudar os desenvolvedores em uma futura melhor acurácia para o modelo.

Em relação à bactéria *E. coli* 537B, os dados do Vitek 2 mostraram que a mesma foi resistente apenas a antibióticos beta-lactâmicos (Ampicilina, Ampicilina/Sulbactam, Cefuroxima, Cefuroxima acetil, Cefotaxima, Ceftazidima e Cefoxitina). O DeepARG identificou nesta bactéria, a sequência RG-04, a qual foi indicada com a categoria indefinida.

Entretanto, o Vitek 2 com sua cartela de antibióticos não realiza o teste com todos antibióticos. Além disso, o desenho experimental teve um número limitado de antibióticos a serem testados e não foi possível a realização do teste com as 30 categorias do DeepARG.

Assim, a caracterização funcional teve um número limitado de antibióticos a serem testados, pois para a escolha dos antibióticos foram escolhidos antibióticos de relevância clínica. Como exemplo, o número limitado de antibióticos teve barreiras para a sequência com resistência indefinida para as categorias testadas no modelo e necessita de testes com outras classes de antibióticos para a definição de sua resistência.

Segundo McGarvey, Queitsch e Fields (2012), em comparação com as diferentes metodologias para a caracterização funcional, o ensaio de gotejamento é o mais indicado para a visualização da atribuição de resistência pelo gene inserido no clone bacteriano. O ensaio de gotejamento é visto como teste quantitativo, diferente do teste difusão de disco que é qualitativo. Logo, em comparação com o teste de difusão de disco, a concentração reduzida dos antibióticos poderia indicar um sinal de resistência em comparação ao controle *E. coli* DH10B.

Apesar do teste difusão de disco ser utilizado para espécies selvagens, vale destacar que em relação ao controle positivo (KPC-2) foi possível a visualização da resistência a antibióticos beta-lactâmicos no antibiograma, mesmo em concentrações mais altas comparadas às utilizadas no ensaio de gotejamento. Quanto ao questionamento apresentado sobre se as proteínas estariam sendo expressas, as sequências selecionadas são de bactérias filogeneticamente próximas à *E. coli* e, também, a expressão do gene KPC-2 comprovada na caracterização funcional garante a funcionalidade do constructo genético. Entretanto, para as outras sequências, um ensaio de *western blot* seria necessário para a comprovação de suas expressões. Quanto ao *burden* metabólico, as análises das curvas de crescimento não foram conclusivas e as análises de expressão proteica de cada clone contendo as sequências de interesse são necessárias para a consolidação da hipótese inicial. Ao analisar o antibiograma, a variação de susceptibilidade em relação ao controle negativo (*E. coli* DH10B com pUC19 vazio) pode ser dependente do antibiótico a ser utilizado.

Como apontado pelos desenvolvedores, o DeepARG carece de ARGs suficientes para uma melhoria do seu treinamento e consequente melhor precisão na identificação de sequências resistentes, assim como, uma melhor curadoria de ARGs presentes em bancos de dados é essencial para a diminuição de falsos positivos (ARANGO-ARGOTY *et al.*, 2018).

Em um trabalho desenvolvido por Marini *et al.* (2022), comparações entre diferentes ferramentas de predição de ARGs foram performadas. Entre elas, foram analisadas ferramentas baseadas em alinhamento (AMRPlus), *deep learning* (DeepARG) e modelos ocultos de Markov (Meta-MARC). Com uma coleção 585 isolados bacterianos com resistência a antibióticos, a identificação correta de ARGs pelas diferentes ferramentas foi avaliada.

As ferramentas Resfinder, Karga e AMRPlus obtiveram uma melhor acurácia, entretanto o DeepARG teve uma performance baixa comparada às outras ferramentas. Para genes de resistência a fluoroquinolonas teve uma acurácia balanceada ruim de 0.50 (os valores vão de 0 a 1). Este resultado vai de encontro com os resultados obtidos para a

sequência RG-02, na qual a predição para resistência a fluoroquinolonas não estava correta. Em contrapartida, para beta-lactamases, a ferramenta teve uma ótima predição com uma acurácia balanceada de 0.84 (MARINI *et al.*, 2022).

Como mostrado, o DeepARG pode ter o potencial para identificação de sequências resistentes a beta-lactâmicos, as quais representam a maior diversidade no banco usado para o treinamento e validação do modelo. Como mostrado em seu próprio artigo, Arango-Argoty e colaboradores (2018) mostraram a potencialidade do DeepARG ao identificar corretamente, uma coleção de potenciais 76 novos genes beta-lactamases. Assim, o DeepARG identificou 65 destes genes, tendo uma acurácia de 85%.

De maneira geral, trabalhos acadêmicos têm utilizado o DeepARG para a identificação *in silico* de genes resistentes a antibióticos, mas a verificação do seu potencial na identificação de potenciais ARGs tem sido negligenciada empiricamente. Este projeto é um estímulo para futuros projetos, portanto visando que novas integrações sejam implantadas para identificação de ARGs desconhecidas. Entretanto, como exposto neste tópico, a ferramenta carece de informações concisas e o emprego de dados clínicos são essenciais para contradizer ou confirmar os dados apresentados pelo DeepARG, assim como destacado em nossos resultados.

6 CONCLUSÕES

Com o desenvolvimento deste trabalho, as partes integrantes do projeto tiveram três conclusões referentes aos resultados:

- As sequências selecionadas identificadas pelo DeepARG não foram funcionais nas condições experimentais utilizadas neste estudo;
- Como concluído pelos desenvolvedores do DeepARG, o algoritmo necessita de mais dados para aumentar sua performance; e
- Atualmente, este estudo está limitado quanto ao número de potenciais ARGs e antibióticos testados.

7 BIBLIOGRAFIA

- ABE, K.; NOMURA, N.; SUZUKI, S. **Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism.** FEMS Microbiology Ecology, v. 96, n. 5, p. 31, 1 maio 2020.
- ABUSHAHEEN, M. A. *et al.* **Antimicrobial resistance, mechanisms and its clinical significance.** Disease-a-Month, v. 66, n. 6, 1 jun. 2020.
- ALCOCK, B. P. *et al.* **CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database.** Nucleic acids research, v. 51, n. D1, 6 jan. 2023.
- ALGHAMDI, S. **The role of vaccines in combating antimicrobial resistance (AMR) bacteria.** Saudi Journal of Biological Sciences, v. 28, n. 12, p. 7505–7510, 1 dez. 2021.
- ALMAGRO ARMENTEROS, J. J. *et al.* **SignalP 5.0 improves signal peptide predictions using deep neural networks.** Nature biotechnology, v. 37, n. 4, p. 420–423, 1 abr. 2019.
- ALVAREZ-ORTEGA, C.; OLIVARES, J.; MARTÍNEZ, J. L. **RND multidrug efflux pumps: What are they good for?** Frontiers in Microbiology Frontiers Research Foundation, , 2013.
- AMIRIE, M. *et al.* **Attempts to Construct an Enlarged pUC19 via Insertion of HindIII-digested Coliphage λ DNA.** Journal of Experimental Microbiology and Immunology (JEMI), v. 18, p. 96–101, 2014.
- ARANGO-ARGOTY, G. *et al.* **DeepARG: A deep learning approach for predicting antibiotic resistance genes from metagenomic data.** Microbiome, v. 6, n. 1, p. 1–15, 1 fev. 2018.
- ARNOLD, B. J.; HUANG, I. T.; HANAGE, W. P. **Horizontal gene transfer and adaptive evolution in bacteria.** Nature Reviews Microbiology 2021 20:4, v. 20, n. 4, p. 206–218, 12 nov. 2021.
- ASHKENAZI, S. **Beginning and possibly the end of the antibiotic era.** Journal of Paediatrics and Child Health, v. 49, n. 3, p. E179–E182, 1 mar. 2013.
- BANERJEE, R.; HUMPHRIES, R. **Rapid Antimicrobial Susceptibility Testing Methods for Blood Cultures and Their Clinical Impact.** Frontiers in Medicine, v. 8, p. 167, 10 mar. 2021.
- BARNES, M. D. *et al.* **Klebsiella pneumoniae Carbapenemase-2 (KPC-2), Substitutions at Ambler Position Asp179, and Resistance to Ceftazidime-Avibactam: Unique Antibiotic-Resistant Phenotypes Emerge from β -Lactamase Protein Engineering.** mBio, v. 8, n. 5, 1 set. 2017.
- BATEMAN, A. **UniProt: a worldwide hub of protein knowledge.** Nucleic acids research, v. 47, n. D1, p. D506–D515, 8 jan. 2019.
- BLAIR, J. M. A. *et al.* **Molecular mechanisms of antibiotic resistance.** Nature Reviews Microbiology 2014 13:1, v. 13, n. 1, p. 42–51, 1 dez. 2014.
- BONECA, I. G.; CHIOSIS, G. **Vancomycin resistance: occurrence, mechanisms and strategies to combat it.** <http://dx.doi.org/10.1517/14728222.7.3.311>, v. 7, n. 3, p. 311–328, jun. 2005.

BORELLI, T. C. *et al.* **Combining Functional Genomics and Whole-Genome Sequencing to Detect Antibiotic Resistance Genes in Bacterial Strains Co-Occurring Simultaneously in a Brazilian Hospital.** *Antibiotics* (Basel, Switzerland), v. 10, n. 4, 2021.

(BRASIL). **PAN-BR Plano de Ação Nacional de Prevenção e Controle da Resistência aos Antimicrobianos no Âmbito da Saúde Única.** 2018.

(BRCAST). **Teste de sensibilidade aos antimicrobianos: Método de Disco-Difusão-BrCAST.** [s.l: s.n.]. 2021.

BUCHAN, D. W. A.; JONES, D. T. **The PSIPRED Protein Analysis Workbench: 20 years on.** *Nucleic acids research*, v. 47, n. W1, p. W402–W407, 1 jul. 2019.

BUCHFINK, B.; XIE, C.; HUSON, D. H. **Fast and sensitive protein alignment using DIAMOND.** *Nature Methods* 2014 12:1, v. 12, n. 1, p. 59–60, 17 nov. 2014.

BUSCH, G. *et al.* **Perceptions of antibiotic use in livestock farming in Germany, Italy and the United States.** *Livestock Science*, v. 241, p. 104251, 1 nov. 2020.

CASSINI, A. *et al.* **Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis.** *The Lancet Infectious Diseases*, v. 19, n. 1, p. 56–66, 1 jan. 2019.

CDC. **COVID-19: U.S. Impact on Antimicrobial Resistance, Special Report 2022.** Atlanta, Georgia: [s.n.]. Disponível em: <<https://stacks.cdc.gov/view/cdc/117915>>.

C REYGAERT, W. **An overview of the antimicrobial resistance mechanisms of bacteria.** *AIMS Microbiology*, v. 4, n. 3, p. 482–501, 2018.

CHOKSHI, A. *et al.* **Global Contributors to Antibiotic Resistance.** *Journal of Global Infectious Diseases*, v. 11, n. 1, p. 36, 1 jan. 2019.

CUADRAT, R. R. C. *et al.* **Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples.** *GigaScience*, v. 9, n. 5, p. 1–12, 1 maio 2020.

DA CUNHA, B. R.; FONSECA, L. P.; CALADO, C. R. C. **Antibiotic Discovery: Where Have We Come from, Where Do We Go?** *Antibiotics* 2019, Vol. 8, Page 45, v. 8, n. 2, p. 45, 24 abr. 2019.

ESPINEL-INGROFF, A.; TURNIDGE, J. **The role of epidemiological cutoff values (ECVs/ECOFFs) in antifungal susceptibility testing and interpretation for uncommon yeasts and moulds.** *Revista iberoamericana de micologia*, v. 33, n. 2, p. 63–75, 1 abr. 2016.

(EUCAST). **EUCAST Disk Diffusion Method for Antimicrobial Susceptibility Testing Antimicrobial susceptibility testing EUCAST disk diffusion method.** [s.l: s.n.]. Disponível em: <www.eucast.org>. Acesso em: 3 fev. 2023.

FATHIMA, B.; ARCHER, A. C. **Bacteriophage therapy: recent developments and applications of a renaissance weapon.** *Research in Microbiology*, v. 172, n. 6, p. 103863, 1 set. 2021.

FINN, R. D. *et al.* **Pfam: the protein families database.** *Nucleic acids research*, v. 42, n. Database issue, 1 jan. 2014.

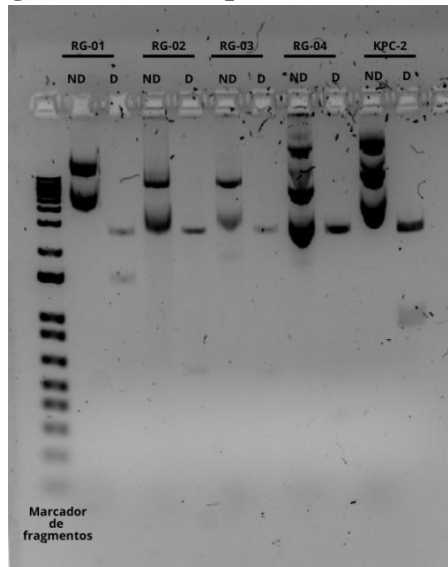
- FROHMAN, M. **Physiological and Pathophysiological roles for Phospholipase D Study on Aquaporin-3-Phospholipase D2 module in skin diseases** View project. Article in Journal of Lipid Research, 2015.
- FU, L. *et al.* **CD-HIT**: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, v. 28, n. 23, p. 3150–3152, 1 dez. 2012.
- GASPAR, G. G. *et al.* **Nosocomial Outbreak of Extensively Drug-Resistant (Polymyxin B and Carbapenem) *Klebsiella pneumoniae* in a Collapsed University Hospital Due to COVID-19 Pandemic**. *Antibiotics*, v. 11, n. 6, 1 jun. 2022.
- GOH, S.; GOOD, L. **Plasmid selection in *Escherichia coli* using an endogenous essential gene marker**. *BMC Biotechnology*, v. 8, n. 1, p. 1–9, 11 ago. 2008.
- HAO, X.; ZHANG, G.; MA, S. **Deep Learning**. *International Journal of Semantic Computing*, v. 10, n. 3, p. 417–439, 1 set. 2016.
- HARBARTH, S. *et al.* **Antimicrobial resistance: One world, one fight!** *Antimicrobial Resistance and Infection Control*, v. 4, n. 1, p. 1–15, 18 nov. 2015.
- HENDRIKSEN, R. S. *et al.* **Using Genomics to Track Global Antimicrobial Resistance**. *Frontiers in Public Health*, v. 7, p. 242, 4 set. 2019.
- JOHNSTON, C. *et al.* **Bacterial transformation: distribution, shared mechanisms and divergent control**. *Nature Reviews Microbiology* 2014 12:3, v. 12, n. 3, p. 181–196, 10 fev. 2014.
- JUMPER, J. *et al.* **Highly accurate protein structure prediction with AlphaFold**. *Nature* 2021 596:7873, v. 596, n. 7873, p. 583–589, 15 jul. 2021.
- KIM, H. H. S.; JUNG, J. H. **Social Isolation and Psychological Distress During the COVID-19 Pandemic: A Cross-National Analysis**. *The Gerontologist*, v. 61, n. 1, p. 103–113, 21 jan. 2021.
- KNIGHT, G. M. *et al.* **Antimicrobial resistance and covid-19: Intersections and implications**. *eLife*, v. 10, p. 1–27, 1 fev. 2021.
- KOTWANI, A.; JOSHI, J.; KALONI, D. **Pharmaceutical effluent: a critical link in the interconnected ecosystem promoting antimicrobial resistance**. *Environmental Science and Pollution Research*, v. 28, n. 25, p. 32111–32124, 1 jul. 2021.
- KOWALSKA-KROCHMAL, B.; DUDEK-WICHER, R. **The Minimum Inhibitory Concentration of Antibiotics: Methods, Interpretation, Clinical Relevance**. *Pathogens (Basel, Switzerland)*, v. 10, n. 2, p. 1–21, 1 fev. 2021.
- KRAWCZYK, P. S.; LIPINSKI, L.; DZIEMBOWSKI, A. **PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures**. *Nucleic Acids Research*, v. 46, n. 6, p. e35–e35, 6 abr. 2018.
- LAI, C. C. *et al.* **Increased antimicrobial resistance during the COVID-19 pandemic**. *International Journal of Antimicrobial Agents* Elsevier B.V., , 1 abr. 2021.
- LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning**. *Nature*, v. 521, n. 7553, p. 436–444, 27 maio 2015.

- LERMINIAUX, N. A.; CAMERON, A. D. S. **Horizontal transfer of antibiotic resistance genes in clinical environments**. Canadian journal of microbiology, v. 65, n. 1, p. 34–44, 2019.
- LIN, L. *et al.* **Factors influencing inappropriate use of antibiotics in outpatient and community settings in China: a mixed-methods systematic review**. BMJ Global Health, v. 5, n. 11, p. e003599, 1 nov. 2020.
- LIU, B.; POP, M. **ARDB--Antibiotic Resistance Genes Database**. Nucleic acids research, v. 37, n. Database issue, 2009.
- LU, S. *et al.* **CDD/SPARCLE: the conserved domain database in 2020**. Nucleic acids research, v. 48, n. D1, p. D265–D268, 1 jan. 2020.
- MACGOWAN, A. P.; WISE, R. **Establishing MIC breakpoints and the interpretation of in vitro susceptibility tests**. p. 1–24, jan. 2005.
- MAHMUD, M. *et al.* **Applications of Deep Learning and Reinforcement Learning to Biological Data**. IEEE Transactions on Neural Networks and Learning Systems, v. 29, n. 6, p. 2063–2079, 1 jun. 2018.
- MARINI, S. *et al.* **Towards routine employment of computational tools for antimicrobial resistance determination via high-throughput sequencing**. Briefings in bioinformatics, v. 23, n. 2, 1 mar. 2022.
- MARTÍNEZ, José L.; COQUE, Teresa M.; BAQUERO, Fernando. **What is a resistance gene? Ranking risk in resistomes**. Nature Reviews Microbiology, [S.L.], v. 13, n. 2, p. 116–123, 15 dez. 2014. Springer Science and Business Media LLC.
<http://dx.doi.org/10.1038/nrmicro3399>.
- MCGARVEY, K. M.; QUEITSCH, K.; FIELDS, S. **Wide variation in antibiotic resistance proteins identified by functional metagenomic screening of a soil DNA library**. Applied and Environmental Microbiology, v. 78, n. 6, p. 1708–1714, mar. 2012.
- MCGINNIS, S.; MADDEN, T. L. **BLAST: at the core of a powerful and diverse set of sequence analysis tools**. Nucleic Acids Research, v. 32, n. suppl_2, p. W20–W25, 1 jul. 2004.
- MURRAY, C. J. *et al.* **Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis**. The Lancet, v. 399, n. 10325, p. 629–655, 12 fev. 2022.
- NEIL, K.; ALLARD, N.; RODRIGUE, S. **Molecular Mechanisms Influencing Bacterial Conjugation in the Intestinal Microbiota**. Frontiers in microbiology, v. 12, 4 jun. 2021.
- NIELSEN, M. A. **Neural Networks and Deep Learning**. [s.l.] Determination Press, 2015.
- NIKAIDO, H. **STRUCTURE AND MECHANISM OF RND-TYPE MULTIDRUG EFFLUX PUMPS**. Advances in enzymology and related areas of molecular biology, v. 77, p. 1, 2011.
- O'NEILL, J. **Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations**. Geneva: [s.n.]. . Acesso em: 2 fev. 2023. 2014.
- ONICIUC, E. A. *et al.* **The Present and Future of Whole Genome Sequencing (WGS) and Whole Metagenome Sequencing (WMS) for Surveillance of Antimicrobial Resistant Microorganisms and Antimicrobial Resistance Genes across the Food Chain**. Genes, v. 9, n. 5, 22 maio 2018.

- (OMS). **Global Action Plan on Antimicrobial Resistance**. Geneva: [s.n.]. 2015.
- (BRASIL). **PAN-BR Plano de Ação Nacional de Prevenção e Controle da Resistência aos Antimicrobianos no Âmbito da Saúde Única**. 2018.
- PEREIRA, J. *et al.* **High-accuracy protein structure prediction in CASP14**. *Proteins*, v. 89, n. 12, p. 1687–1699, 1 dez. 2021.
- PEREIRA, R. J. *et al.* **The risk of COVID-19 transmission in favelas and slums in Brazil**. *Public Health*, v. 183, p. 42, 1 jun. 2020.
- PETERSEN, L. M. *et al.* **Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing**. *Journal of Clinical Microbiology*, v. 58, n. 1, 1 jan. 2020.
- PHILLIPS, I. **&bgr;-LACTAMASE-PRODUCING, PENICILLIN-RESISTANT GONOCOCCUS**. *The Lancet*, v. 308, n. 7987, p. 656–657, 25 set. 1976
- RABELLO, R. F. *et al.* **Antimicrobial Resistance in Farm Animals in Brazil: An Update Overview**. *Animals : an Open Access Journal from MDPI*, v. 10, n. 4, p. 552, 1 abr. 2020.
- RANJAN, R. *et al.* **Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing**. *Biochemical and biophysical research communications*, v. 469, n. 4, p. 967–977, 22 jan. 2016.
- ROY, A.; KUCUKURAL, A.; ZHANG, Y. **I-TASSER: a unified platform for automated protein structure and function prediction**. *Nature protocols*, v. 5, n. 4, p. 725–738, 2010.
- RUSK, N. **Deep learning**. Nature Publishing Group, 2015.
- SANTAJIT, S.; INDRAWATTANA, N. **Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens**. BioMed Research International Hindawi Limited, , 2016.
- SELVY, P. E. *et al.* **Phospholipase D: Enzymology, functionality, and chemical modulation**. *Chemical Reviews*, v. 111, n. 10, p. 6064–6119, 12 out. 2011.
- SIGRIST, C. J. A. *et al.* **PROSITE, a protein domain database for functional characterization and annotation**. *Nucleic Acids Research*, v. 38, n. suppl_1, p. D161–D166, 1 jan. 2010.
- SPELLBERG, B. **The future of antibiotics**. *Critical Care*, v. 18, n. 3, p. 1–7, 27 jun. 2014.
- STOGIOS, P. J.; SAVCHENKO, A. **Molecular mechanisms of vancomycin resistance**. *Protein Science : A Publication of the Protein Society*, v. 29, n. 3, p. 654, 1 mar. 2020.
- STOKES, J. M. *et al.* **A Deep Learning Approach to Antibiotic Discovery**. *Cell*, v. 180, n. 4, p. 688–702.e13, 20 fev. 2020.
- TACCONELLI, E. *et al.* **Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics**. Geneva: [s.n.]. Disponível em: <<http://www.cdc.gov/drugresistance/threat-report-2013/>>. Acesso em: 4 fev. 2023.
- TAPPERT, C. C. **Who is the father of deep learning?** Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019. Anais...Institute of Electrical and Electronics Engineers Inc., 1 dez. 2019.
- TENOVER, F. C. **Antimicrobial Susceptibility Testing**. *Encyclopedia of Microbiology*, p. 166–175, 1 jan. 2019.

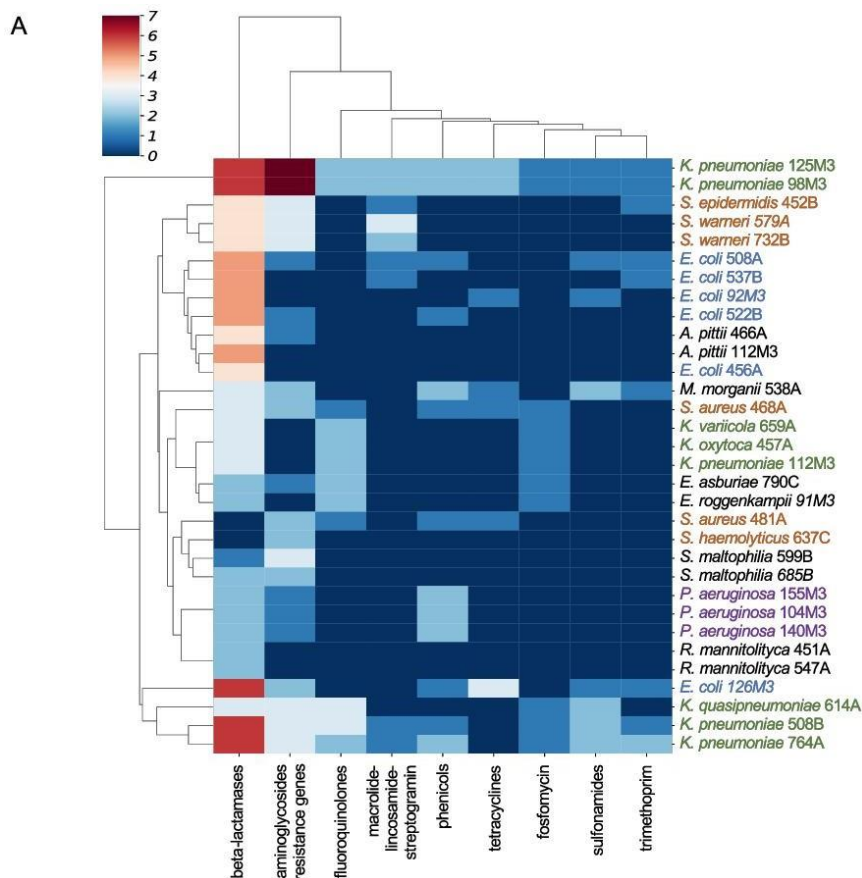
- TOH, S. Y. **The Study of Exclusion Effect of pBR322 using its rop-inactivated Mutant, During Co-transformation with pBR322 and pUC19: Plasmid Copy Number Does Not Relate to the Exclusion of pBR322.** *Journal of Experimental Microbiology and Immunology (JEMI)*, v. 17, p. 109–114, 2013.
- TRIMBLE, M. J. *et al.* **Polymyxin: Alternative Mechanisms of Action and Resistance.** *Cold Spring Harbor Perspectives in Medicine*, v. 6, n. 10, 1 out. 2016.
- VENTOLA, C. L. **The Antibiotic Resistance Crisis: Part 1: Causes and Threats.** *Pharmacy and Therapeutics*, v. 40, n. 4, p. 277, 2015.
- VIROLLE, C. *et al.* **Plasmid transfer by conjugation in gram-negative bacteria: From the cellular to the community level.** *GenesMDPI AG*, , 1 nov. 2020.
- WATERHOUSE, A. *et al.* **SWISS-MODEL: homology modelling of protein structures and complexes.** *Nucleic Acids Research*, v. 46, n. W1, p. W296–W303, 2 jul. 2018.
- WILSON, D. N. **Ribosome-targeting antibiotics and mechanisms of bacterial resistance.** *Nature Reviews Microbiology*, v. 12, n. 1, p. 35–48, jan. 2014.
- WU, G. *et al.* **Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications.** *Trends in Biotechnology*, v. 34, n. 8, p. 652–664, 1 ago. 2016.
- YOUSAF, H. *et al.* **Green synthesis of silver nanoparticles and their applications as an alternative antibacterial and antioxidant agents.** *Materials Science and Engineering: C*, v. 112, p. 110901, 1 jul. 2020.
- ZHOU, J.; TROYANSKAYA, O. G. **Predicting effects of noncoding variants with deep learning-based sequence model.** *Nature Methods* 2015 12:10, v. 12, n. 10, p. 931–934, 24 ago. 2015.

ANEXO A - Resultado da digestão do vetor pUC19 contendo as sequências selecionadas

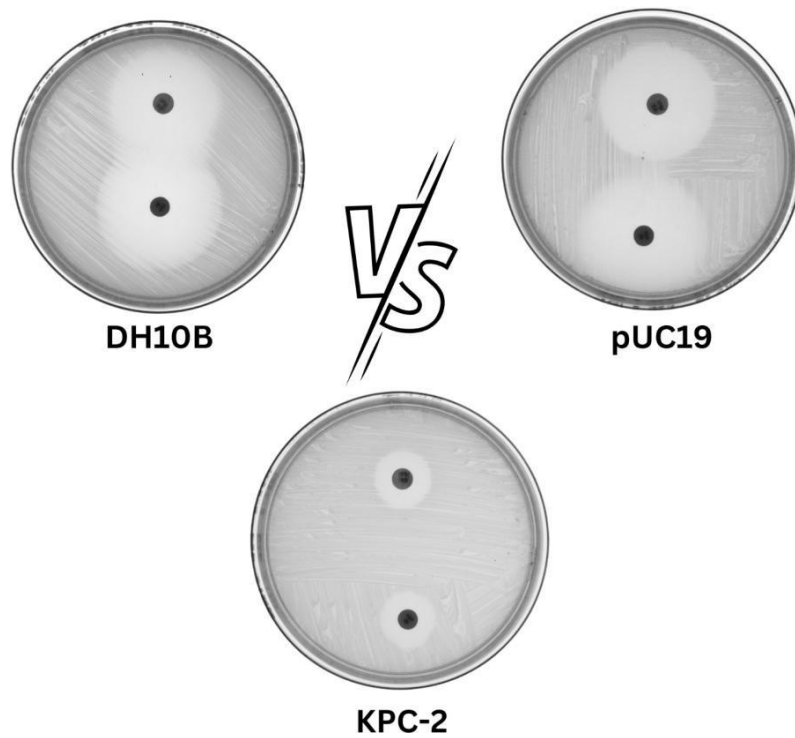


Fonte: Autoral. Gel de agarose da digestão do vetor pUC19. ND - Não digerido, D - Digerido. As sequências possuem os seguintes tamanhos em pares de bases (pb): RG-01(1514 pb), RG-02 (632 pb), RG-03 (605 pb), RG-04 (329 pb) e KPC-2 (989 pb).

ANEXO B - ARGs identificados pelo programa ABRicate



Fonte: Borelli et. al (2021). Mapa de calor com a quantidade de genes de resistência para cada classe de antibióticos identificados pelo ABRicate.

ANEXO C - Antibiograma do clone contendo o gene KPC-2

As imagens do antibiograma foram geradas pelo Sistema de Imagens Chemidoc. A imagem mostra os perfis apresentados pela *E. coli* DH10B (DH10B), *E. coli* DH10B com vetor vazio (pUC19) e com inserto (KPC-2). Na imagem estão os discos de antibióticos: Imipenem (IPM) no topo e cefepime (FEP) no lado inferior.