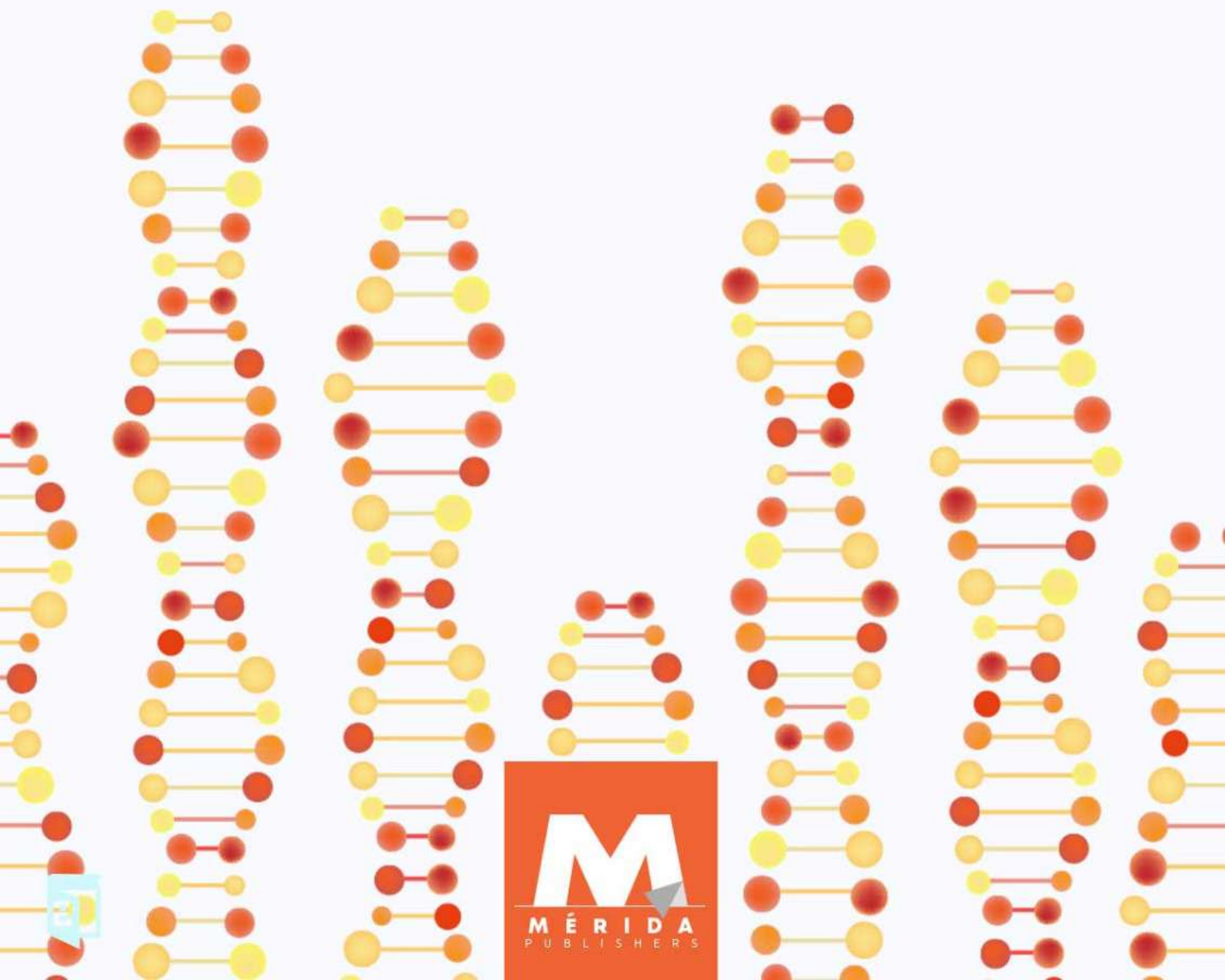


ORGANIZADOR:
Luis Guillermo Ramírez Mérida

Genomas e Biotecnologia:

novas fronteiras em microbiologia



ORGANIZADOR:

Luis Guillermo Ramírez Mérida

Genomas e Biotecnologia:

novas fronteiras em microbiologia

Canoas
2025



Genomas e Biotecnologia: novas fronteiras em microbiologia

© 2025 Mérida Publishers

<https://doi.org/10.69570/mp.978-65-84548-32-9>

Organizador

Luis Guillermo Ramírez Mérida

Adaptação da capa e desenho gráfico

Luis Miguel Guzmán



Canoas - RS - Brasil

contact@meridapublishers.com

www.meridapublishers.com

Todos os direitos autorais pertencem a Mérida Publishers. A reprodução total ou parcial dos trabalhos publicados, é permitida desde que sejam atribuídos créditos aos autores.

Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)

G335 Genomas e Biotecnologia [livro eletrônico] : novas fronteiras em microbiologia / Organizador Luis Guillermo Ramírez Mérida. – Canoas, RS: Mérida Publishers, 2025.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia

Edição bilingue

ISBN 978-65-84548-32-9

1. Biotecnologia. 2. Genoma. 3. Microbiologia. I. Mérida, Luis Guillermo Ramírez.

CDD 570.285

Elaborado por Maurício Amormino Júnior – CRB6/2422

Apresentação

A microbiologia e a biotecnologia têm passado por uma revolução impulsionada por avanços tecnológicos que nos permitem explorar a vida em sua essência molecular e genética. Este livro, " Genomas e Biotecnologia: novas fronteiras em microbiologia", é um convite para explorar as fronteiras do conhecimento nesses campos fascinantes.

Na primeira parte, mergulhamos no mundo da patogenômica, desvendando os genomas dos patógenos e as estratégias moleculares que utilizam para nos infectar. Explorando as técnicas de sequenciamento de alto rendimento, a genômica comparativa e as ferramentas bioinformáticas que nos permitem identificar genes de virulência, entender a relação patógeno-hospedeiro e rastrear a evolução de doenças infecciosas.

Em seguida, investigamos o impacto das técnicas moleculares no diagnóstico, tratamento e prevenção de doenças infecciosas. Se abordaram as últimas inovações em métodos de amplificação de ácidos nucleicos, sequenciamento de DNA, hibridização molecular e diagnóstico baseado em CRISPR, revelando como essas ferramentas estão transformando a microbiologia e a medicina.

A terceira parte do livro é dedicada à probiogenômica, um campo emergente que estuda os genomas dos probióticos e seus benefícios para a saúde humana. Exploram-se as tecnologias ômicas que nos permitem identificar os mecanismos moleculares pelos quais os probióticos exercem seus efeitos benéficos, desde a modulação do sistema imunológico até a produção de compostos bioativos.

Finalmente, embarcamos em uma jornada pela biotecnologia de microalgas, um campo promissor que nos oferece soluções sustentáveis para produção de bioenergia e compostos bioativos. O capítulo discute o potencial das microalgas como fonte de biocombustíveis, alimentos e produtos farmacêuticos, além de explorar os desafios e as perspectivas futuras dessa área.

Este livro é destinado a estudantes, pesquisadores e profissionais das áreas de microbiologia, biotecnologia, medicina e áreas afins. Esperamos que ele sirva como um guia abrangente e inspirador, despertando o interesse pela pesquisa e pelas aplicações práticas desses campos em constante evolução.

Boa leitura!

Autores

Aline Ferreira Maciel de Oliveira

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Ana Lua de Oliveira Vinhal

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Arun Kumar Jaiswal

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Bernardo Buhr Alves Mendonça

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Eduarda Guimarães Sousa

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Fernanda Diniz Prates

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Núcleo Técnico Operacional, Microbiologia, Instituto Hermes Pardini, Vespasiano, Minas Gerais, Brasil.

Gabriel Camargos Gomes

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Janaíne Aparecida de Paula

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Laboratório de Bioinformática, Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Biológicas e Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brasil.

João Matheus Sobral Pena

Laboratório de Biologia Molecular e Fisiologia de Estreptococos, Departamento de Biologia Celular, Universidade do Estado do Rio de Janeiro (UERJ), Rua São Francisco Xavier, 524, Maracanã, 20550013, Rio de Janeiro, Brasil.

Laura Diniz Freitas

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Letícia Isadora Rodrigues

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Lucas Henrique Ramos de Oliveira

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Luis G. Ramírez Mérida

Profesor Titular, Departamento de Biología, Facultad de Ciencias y Tecnología, Universidad de Carabobo, Valencia – Venezuela.

Luiza Pereira Reis

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Marcus Vinicius Canário Viana

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Mariana Letícia Costa Pedrosa

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Pamella Silva Lannes-Costa

Laboratório de Biologia Molecular e Fisiologia de Estreptococos, Departamento de Biologia Celular, Universidade do Estado do Rio de Janeiro (UERJ), Rua São Francisco Xavier, 524, Maracanã, 20550013, Rio de Janeiro, Brasil.

Pedro Henrique Marques

Laboratório de Bioinformática, Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Biológicas e Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brasil.

Prescilla Emy Nagao

Laboratório de Biologia Molecular e Fisiologia de Estreptococos, Departamento de Biologia Celular, Universidade do Estado do Rio de Janeiro (UERJ), Rua São Francisco Xavier, 524, Maracanã, 20550013, Rio de Janeiro, Brasil.

Rafael Júnio de Oliveira

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Sandeep Tiwari

Instituto de Biologia, Universidade Federal da Bahia, Salvador, Bahia, Brasil.

Siomar de Castro Soares

Laboratório de Bioinformática, Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Biológicas e Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brasil.

Talita Pereira Gomes

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Instituto Nacional de Ciência e Tecnologia em Estudos Interdisciplinares e Transdisciplinares em Ecologia e Evolução (INCT IN-TREE), Instituto de Biologia, Universidade Federal da Bahia, Salvador, Bahia, Brasil.

Vasco Ariston de Carvalho Azevedo

Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

Índice

CAPÍTULO 1..... 11

Patogenômica: Explorando os Genomas de Patógenos

Lucas Gabriel Rodrigues Gomes, Eduarda Guimarães Sousa, Fernanda Diniz Prates, Talita Pereira Gomes, Gabriel Camargos Gomes, Janaíne Aparecida de Paula, Pedro Henrique Marques, Ana Lua de Oliveira Vinhal, Bernardo Buhr Alves Mendonça, Aline Ferreira Maciel de Oliveira, Mariana Letícia Costa Pedrosa, Luiza Pereira Reis, Marcus Vinicius Canário Viana, Siomar de Castro Soares, Sandeep Tiwari, Arun Kumar Jaiswal, Vasco Azevedo

CAPÍTULO 2..... 67

Atualizações, impactos e perspectivas das técnicas moleculares: contextualizando a biologia molecular

João Matheus Sobral Pena, Prescilla Emy Nagao, Pamella Silva Lannes-Costa

CAPÍTULO 3..... 96

Probiogenômica: Análise Genômica dos Probióticos e Seus Benefícios à Saúde

Eduarda Guimarães Sousa, Lucas Gabriel Rodrigues Gomes, Arun Kumar Jaiswal, Laura Diniz Freitas, Rafael Júnio de Oliveira, Letícia Isadora Rodrigues, Luiza Pereira Reis, Aline Ferreira Maciel de Oliveira, Siomar de Castro Soares, Vasco Ariston de Carvalho Azevedo

CAPÍTULO 4..... 140

Microalgal biotechnology: a new green economy?

Luis Guillermo Ramírez Mérida

CAPÍTULO 1

Patogenômica: Explorando os Genomas de Patógenos

Lucas Gabriel Rodrigues Gomes, Eduarda Guimarães Sousa, Fernanda Diniz Prates, Talita Pereira Gomes, Gabriel Camargos Gomes, Janaíne Aparecida de Paula, Pedro Henrique Marques, Ana Lua de Oliveira Vinhal, Bernardo Buhr Alves Mendonça, Aline Ferreira Maciel de Oliveira, Mariana Letícia Costa Pedrosa, Luiza Pereira Reis, Marcus Vinicius Canário Viana, Siomar de Castro Soares, Sandeep Tiwari, Arun Kumar Jaiswal, Vasco Azevedo

<https://doi.org/10.69570/mp.978-65-84548-32-9.c1>

Resumo

Com o advento de desenvolvimentos na área da computação, e com a redução de custo de técnicas para o sequenciamento de genomas, a era do *Big Data* resultou na imensa expansão do nosso repertório de genomas de interesse. Hoje, bancos de dados como o GenBank contém grandes volumes de genomas de patógenos, e mais são depositados a cada dia. A disponibilidade desses dados permite o desenvolvimento de técnicas e ferramentas que servem tanto para estudar os genes que conhecemos quanto para usá-los como referência no estudo de novas descobertas. Através da patogenômica, é possível estudar os genomas de patógenos, identificando genes relacionados com a sua patogênese e entendendo a relação patógeno-hospedeiro. Desse modo, o campo busca guiar novos métodos diagnósticos, profiláticos, e de tratamento, usando o genoma como fundamentação. Neste capítulo, propomos explorar as metodologias e tecnologias usadas em estudos de patogenômica, como: sequenciamento de alto rendimento; genômica comparativa; e ferramentas e bancos de dados bioinformática. Estas ferramentas expandiram a capacidade de pesquisadores para caracterizar, compreender e rastrear espécies patogênicas. Também propomos explicar de modo amplo as aplicações práticas da patogenômica, incluindo: diagnóstico de doenças infecciosas; epidemiologia; e o desenvolvimento de novos métodos terapêuticos e vacinas. Também incluídos são exemplos de casos ilustrando como dados em patogenômica podem ser aplicados para desvendar o comportamento complexo de um patógeno, identificar fatores de virulência, e compreender mecanismos de resistência a antibióticos, bem como os desafios e direções futuras de exploração do campo.

Palavras-chave: Big Data, Patogenômica, Sequenciamento, Genomas, Bioinformática.

1. Introdução à Patogenômica

A genômica constitui um ramo especializado da genética dedicado à investigação do genoma, definido como a totalidade do material genético contido no DNA de um organismo. Esse campo de estudo, de caráter abrangente, inclui não apenas o sequenciamento e a análise dos genomas, mas também a elucidação das interações entre os genes e os fatores ambientais, proporcionando uma compreensão aprofundada dos mecanismos que regem a expressão gênica e a sua modulação em diferentes contextos biológicos (HIH, 2022).

O avanço nas técnicas de sequenciamento genético proporcionou à ciência um acesso sem precedentes a uma vasta gama de informações sobre diversas classes de organismos, incluindo microrganismos patogênicos de relevância médica, ambiental e veterinária. A integração da bioinformática nesse contexto permite que programas computacionais e bases de dados especializadas sejam empregados por pesquisadores, otimizando e ampliando o escopo das investigações científicas (PALLEN et al., 2007).

A patogenômica, campo de estudo dedicado à exploração de múltiplos aspectos da biologia dos patógenos, abrange desde a análise de seu conteúdo genético e suas relações evolutivas até a identificação de genes associados à resistência antimicrobiana e à virulência. Além disso, examina os mecanismos pelos quais esses microrganismos são capazes de infectar e se proliferar no hospedeiro. Entre os principais objetivos da patogenômica destacam-se a profilaxia, o diagnóstico e o monitoramento da evolução das doenças, bem como o desenvolvimento de tratamentos terapêuticos e vacinas com maior eficácia (DEMUTH et al., 2008).

A patogenômica está intrinsecamente vinculada a outras disciplinas, como a genômica funcional, estrutural e a metagenômica, que, quando integradas, fornecem uma visão completa da biologia dos patógenos e permitem a realização dos objetivos desta área de estudo. A genômica estrutural se destaca por sua missão de caracterizar tridimensionalmente as proteínas codificadas pelos genes, dada a estreita correlação entre estrutura e função (Goldsmith-Fischman e Honig, 2003). Essa abordagem não apenas facilita novas

descobertas sobre a dinâmica e a interação dessas proteínas com outras moléculas, mas também potencializa suas aplicações na descoberta e aprimoramento de fármacos. Por outro lado, a metagenômica, ao permitir o estudo das estruturas e funções de microrganismos a partir da análise de uma comunidade completa presente em uma amostra ambiental, oferece uma compreensão mais profunda de como esses genes interagem entre si e desempenham suas funções no contexto específico daquele ambiente (NRC, 2007).

Por fim, a genômica funcional oferece uma análise abrangente das funções de um gene, abrangendo seu papel integral em vias bioquímicas, celulares e fisiológicas (GARCIA, 2006). Ao empregar sequenciamentos genômicos completos e ferramentas bioinformáticas, é possível analisar com maior precisão os mecanismos que levam ao surgimento de fenótipos, permitindo a modificação direcionada das funções gênicas. Essa compreensão é importante para a patogenômica, na medida em que facilita a identificação de novos alvos terapêuticos, o desenvolvimento de novos fármacos, e até a personalização de tratamentos médicos. Dessa forma, é possível prever e otimizar a resposta individual de cada organismo a determinados fármacos, incrementando a segurança e a eficácia das intervenções terapêuticas (IOC, 2024).

Os genomas dos patógenos contêm o material genético que codifica os mecanismos necessários para estabelecer infecções e evadir o sistema imune do hospedeiro. Fatores genéticos chamados de fatores de patogenicidade são responsáveis pela capacidade de diferentes microrganismos a invadir o corpo do hospedeiro, a aderir às suas membranas, a evadir o sistema imune e a resistir a intervenções por antibióticos (IRAOLA e BOYCE, 2022). Sendo cruciais para a patogênese dos microrganismos, a compreensão desses fatores passa a ser essencial na compreensão e tratamento de doenças (IRAOLA e BOYCE, 2022).

Desses fatores de virulência, um conjunto é de particular preocupação global, os fatores de resistência. O desenvolvimento de resistência a antimicrobianos (AMR) em patógenos, frequentemente dado por fatores genéticos como mutações e transferência horizontal de genes, pode representar um risco expressivo do ponto de vista de saúde pública (WHO, 2023). Com isso,

um grande foco tem sido colocado na compreensão de mecanismos de resistência a antimicrobianos.

Os estudos em patogenômica feitos com bactérias do gênero *Leptospira* e com a espécie *Escherichia coli* podem exemplificar o quanto esse campo de estudo tem a contribuir com inovações no campo da saúde. As bactérias *E.coli*, por exemplo, presentes em dejetos, ocasionalmente podem colonizar o trato urinário e causar quadros de infecção urinária, doença que atinge muitos pacientes do sexo feminino, que possuem a uretra mais curta que o sexo masculino. Apesar de muito comum, é uma doença que, se não tratada corretamente, pode levar à morte por choque séptico. O crescente número de sequenciamento genético da *E. coli* tem proporcionado avanços no entendimento dos seus mecanismos de patogenicidade. Sendo um organismo que coloniza naturalmente o trato intestinal, a bactéria precisa de mecanismos adicionais para invadir e estabelecer-se na uretra, facilitando a transmissão entre hospedeiros. Esses mecanismos são denominados "fatores de virulência extraintestinais". A patogenômica tem revelado que essa bactéria pode colonizar o trato urinário através de duas principais vias: a via hematogênica, que envolve a infecção inicial do parênquima renal, e a via ascendente, que começa com a migração da bactéria das fezes para o trato urinário, atingindo eventualmente a bexiga (AGARWAL et al., 2012). O sequenciamento de genomas de linhagens multirresistentes como no caso de *Mycobacterium tuberculosis*, portanto, nos permite estudar a maneira em que esses mecanismos genéticos surgem, guiando o desenvolvimento de novos antibióticos e novos regimes de tratamento (EKTEFAIE et al., 2021).

Outro exemplo relevante são os estudos conduzidos com bactérias do gênero *Treponema*, responsáveis por uma variedade de treponematoses, dentre elas a sífilis, uma IST de importância global. Assim como no caso mencionado anteriormente, o aumento progressivo no sequenciamento de linhagens tem proporcionado descobertas significativas. Dentre essas análises, a realização de estudos de pangenoma focados na patogenômica desse gênero permitiu elucidar alguns dos processos biológicos cruciais dessas espiroquetas (KUMAR et al., 2017). É crucial ressaltar que a manipulação genética de bactérias do gênero *Treponema* enfrenta consideráveis desafios técnicos, devido às grandes

limitações para a manutenção de uma cultura ao longo prazo dos microrganismos (EDMONDSON et al., 2018). Devido a estas dificuldades, a aplicação da patogenômica tem possibilitado avanços no entendimento da patogênese e diversidade de *Treponema*, o que pode fomentar o desenvolvimento de meios de cultura aprimorados para o cultivo dessas bactérias, bem como novas abordagens terapêuticas. Adicionalmente, a patogenômica oferece perspectivas no desenvolvimento de uma vacina contra a sífilis, uma vez que, até o momento, não existe uma candidata viável para uso em humanos (GOMES et al., 2022).

Esses exemplos ilustram o valor substancial que os estudos em patogenômica podem agregar à saúde pública. A integração de dados de diferentes vertentes da genômica permite uma compreensão mais profunda de aspectos cruciais relacionados à biologia e transmissão de patógenos, incluindo suas relações evolutivas, metabolismo, fatores de virulência, e as interações patógeno-hospedeiro. Tais dados são fundamentais para o desenvolvimento de diagnósticos precoces, bem como para o aprimoramento de novas abordagens diagnósticas e terapêuticas. Nos Estados Unidos, os Centros de Controle e Prevenção de Doenças (CDC) já implementaram programas de investimento em sequenciamento *next generation* e patogenômica, com o objetivo de fortalecer o controle de doenças infecciosas. Um exemplo notável é o PulseNet, um banco de dados associado ao CDC, utilizado na detecção e controle de pequenos surtos de intoxicação alimentar, que anteriormente só eram identificáveis em larga escala (CDC, 2024). Outro exemplo está na formulação anual das vacinas contra o vírus da influenza. Duas vezes ao ano, a Organização Mundial da Saúde (OMS) organiza encontros com especialistas que, com base em informações sobre as cepas virais circulantes, indicam os melhores candidatos para a composição das vacinas sazonais, destacando o papel fundamental dos estudos genômicos na prevenção de epidemias globais (ARMSTRONG et al., 2019).

2. Histórico da Patogenômica

2.1. Origens e Desenvolvimento Inicial

O termo patogenômica surgiu por volta dos anos 2000 devido aos avanços do Projeto Genoma Humano. Patogenômica é a denominação utilizada

para a análise genômica de microrganismos patogênicos, ou seja, refere-se a análise do DNA, RNA e proteínas de diferentes patógenos, como bactérias, vírus ou fungos para caracterizar estruturas moleculares e funções, definir seus processos evolutivos e mapear os genomas destes microrganismos (PRADEL et al., 2006). Vale ressaltar que o primeiro genoma completamente sequenciado no mundo foi de um microorganismo, o bacteriófago ϕ x174, sequenciado em 1977 por Sanger et al., concretizando o marco inicial da era genômica (SANGER et al., 1977).

Mais tarde, em 1995, o primeiro patógeno tem seu genoma sequenciado por Fleischmann et al., o qual alcançou o genoma completo da bactéria *Haemophilus influenzae Rd* (Fleischmann et al., 1995). Um ano depois, em 1996, é sequenciado o primeiro genoma completo de um eucarioto que atua como patógeno de baixa virulência e que é de extremo interesse industrial, a levedura *Saccharomyces cerevisiae*, por obra de Goffeau et al (GOFFEAU et al., 1996). Em 2003, o primeiro genoma humano foi sequenciado pelo Projeto Genoma Humano iniciado em 1990 e esta prática recebeu ainda mais atenção (IHGSC, 2024).

Por volta de 1970 foi desenvolvido a primeira técnica de sequenciamento, Sanger, a qual permitiu o sequenciamento do primeiro genoma completo de vírus, o bacteriófago ϕ x174. (LAGE e SANTOS, 2022). Nos anos 2000, os sequenciadores de segunda geração (NGS) emergem revolucionando o cenário da genômica ao permitir o sequenciamento simultâneo de milhares de fragmentos de DNA com custo e tempo reduzido (SOUZA e BRUSAMARELLO, 2009).

O sequenciamento de terceira e quarta geração surge pouco depois e inclui novas técnicas de sequenciamento como Single Molecule Real-Time (SMRT) desenvolvida pela Pacific Biosystems e Nanopores proporcionando o sequenciamento de moléculas únicas capaz de realizar leituras mais longas, caracterizando um processo com melhor resolução e precisão (POLLARD et al., 2018). Essas tecnologias de sequenciamento, fornecem insights da expressão e regulação genética por análises de transcriptoma que permite a detecção de variantes splicing e outras modificações pós-traducionais essenciais para

desenvolver tratamentos personalizados a fim de melhorar sua eficácia, efetividade e segurança (STARK et al., 2019; SAMEEK e CHINNAIYAN, 2014).

O desenvolvimento do campo ao longo dos anos foi marcado por uma série de passos que nos permitiram expandir nossa compreensão e repertório de genomas e proteomas conhecidos. A **Figura 1** mostra uma linha do tempo destes avanços.



Figura 1. Linha do tempo com os principais pioneiros e que marcaram o início da patogenômica, marco da integração entre computação e biologia, principais descobertas que impulsionaram esse campo da bioinformática.

2.2. Marcos Históricos

Ao longo dos séculos, o *Mycobacterium tuberculosis* tem constituído uma ameaça significativa para a saúde mundial. Esta doença tem sido associada aos seres humanos há cerca de 70.000 anos (COMAS et al., 2013). A informação

anterior não surgiu ao acaso, e sim como resultado de uma abundância de estudos sobre genomas, bioinformática, filogenética e arqueologia. Nesta seção, o foco será o sequenciamento do genoma do *M. tuberculosis* e do *Plasmodium falciparum* e seu contexto evolutivo (BAÑUS et al., 2015; COLE et al., 1998).

A primeira estirpe de *Mycobacterium tuberculosis*, denominada H37Rv, foi obtida em 1998. Tal como descrito por S. T. Cole et al. sobre a investigação, uma grande parte do genoma foi obtida através dos cosmídeos, aproximadamente 3,2 Mb, enquanto a parte restante foi obtida utilizando 45.000 clones shotgun de genoma completo e clones BACs (Bacterial Artificial Chromosome). Os clones foram criados usando vectores M13, para os cosmídeos e BACs, e pUC18 para o DNA genómico, procurando uma melhor cobertura para a montagem, usando ambas as vertentes, “reverse” e “forward”, as regiões propensas à seleção e deleção foram obtidas por reação em cadeia da polimerase (PCR) e a partir de BACs. No final destes processos foram mostradas trinta diferenças entre as sequências shotgun e cosmídeos, resultantes de erros de sequenciação e mutações de cosmídeos. É também necessário contrastar os recursos computacionais utilizados ao longo do projeto, como o PHRAP, GAP4, um script perl, DIANA, TB-parse, EMBL, TREMBL, SwissProt, PROSITE, a sua base de dados, tRNAscan e tRNAscan-SE (COLE et al., 1998).

A malária é outra doença que tem causado muitos desafios aos seres humanos. Sendo um agente patogênico, o genoma do *Plasmodium falciparum* tem sido estudado há muitos anos (GARRIDO-CARDENAS et al., 2019). O primeiro genoma completo deste microrganismo foi sequenciado em 2002, e este é outro marco que será discutido nesta seção. A principal estratégia utilizada pelos investigadores, Malcolm J. Gardner et al, foi um shotgun de cromossomos inteiros, estes que foram separados em géis de campo pulsado, posteriormente o genoma cromossômico foi extraído para a construção de bibliotecas shotgun, um pormenor importante é o shotgun partilhado dos cromossomas 6, 7 e 8, devido resultados insatisfatórios em gel. Por fim foi necessário trabalhar com o fecho de lacunas, quase totalmente feito com Sequence tagged sites (STSs), marcadores microssatélites e mapeamento HAPPY, e alguns, entre 3 e 37, pequenos gaps ficaram nos limites do artigo (GARDNER et al., 2002).

O Projeto Genoma Humano, forneceu um avanço tecnológico e científico essencial para a era genômica a partir do desenvolvimento de tecnologias avançadas de sequenciamento de DNA, mais rápidas e acessíveis, as quais passaram a ser utilizadas para sequenciar o genoma de patógenos (WONG, 2018). Além disso, com o mapeamento de genes preciso e completo do DNA humano, foi possível identificar genes responsáveis por doenças através da comparação daqueles presentes DNA humano com os presentes em DNA de microrganismos (MSHG, 1988).

Vale salientar que compreender como o metabolismo e os mecanismos de defesa do organismo humano interagem com os agentes patogênicos é um forte contribuinte para a identificação de novos alvos terapêuticos e tratamentos mais efetivos e seguros (COVID-19 HOST GENETICS INITIATIVE, 2021). A descoberta de todos estes caminhos para a compreensão do combate a estes microrganismos permitiu maior rastreo e controle da disseminação de patógenos, bem como, de maiores precauções contra endemias e pandemias, tendo em vista o evento pandêmico mais recente, no início de 2020, a pandemia do SARSCoV-2 (KOMISSAROV et al., 2020; LI et al., 2021).

2.3. Evolução e Expansão do Campo

Por volta da década de 1960, a bioinformática emergiu como uma disciplina essencial para coordenar a enorme quantidade de dados biológicos gerados (PEREIRA et al., 2020). Essa integração entre biologia, informática, computação e estatística representa uma revolução na pesquisa biológica e médica, principalmente, com o surgimento de inúmeras ferramentas avançadas em análise genômica (SILVA e ALVES, 2023).

Dessa forma, a integração da bioinformática com a análise de dados apresenta uma importância técnico-científica e social significativa, pois como descrito, permite análises de grandes volumes de dados, facilita a identificação das variações bioquímicas e genéticas em organismos vivos promovendo um avanço na área das ômicas (COSTA, 2014).

Os diferentes marcos relativos às técnicas de sequenciamento deram origem à forma como diferentes metodologias são agrupadas, servindo como

uma associação autodescritiva, cronologicamente falando. Os métodos, atualmente, são divididos em 3 grandes gerações que revolucionaram o estudo de genomas, aprimorando e promovendo diferentes possibilidades de pesquisa (HU et al., 2021).

Inicialmente faz-se necessária a contextualização da era genômica que antecede a primeira geração de sequenciadores. Quando, na década de 1950, se tornou possível determinar os métodos de análise de proteínas emergentes, campo de estudo predecessor ao estudo de sequências genômicas. Inicialmente, isso foi feito de forma analógica e anos depois, com o trabalho de Margaret Dayhoff e Robert S. Ledley no COMPROTEIN, um programa foi escrito em FORTRAN para determinar sequências primárias de proteínas (PAREEK et al., 2011). Após alguns anos, sucederam grandes conquistas neste campo, e em 1977 foi divulgado o primeiro método de sequenciamento por Maxam e Walter Gilbert, com metodologia associada à degradação do DNA, uso de químicos perigosos e radiação, seguido pelo método de Frederick Sanger, associado à terminação de cadeia análogas à dideoxinucleotídeos, técnica utilizada para o sequenciamento do primeiro genoma, o bacteriófago Φ X174 (5386bp), coroando o surgimento da primeira geração de sequenciamento genômico. Até o início da segunda geração surgiram algumas outras metodologias de sequenciamento, bem como a primeira abordagem automatizada com fluorescência (GAUTHIER et al., 2018; PAREEK et al., 2011).

Marcadas pelo ano de 2004, tecnologias de segunda geração foram instauradas (HU et al., 2021), através do desenvolvimento do método Illumina pela Solexa, associadas à métodos de multiprocessamento de leituras, responsável pelo crescimento significativo da velocidade de sequenciamento, tecnologias baseadas em processamento de leituras curtas, a necessidade de criação de bibliotecas e nível de confiabilidade muito alto, sendo protagonizados pelos métodos Illumina, metodologia de fluorescência, e Ion Torrent, metodologia iônica (HU et al., 2021; RODRIGUEZ e KRISHNAN, 2023), que serão aprofundados posteriormente.

Em sequência, no ano de 2011 foi inaugurada a terceira e última geração de sequenciadores onde os pioneiros e protagonistas são o Pacific Biosciences (PacBio), *single-molecule real-time* (SMRT), e o Oxford Nanopore Technologies

(ONT), nanopores e corrente elétrica (XIAO e ZHOU, 2020). Esses métodos trazem como características comuns a utilização de leituras longas, maior eficiência ao trabalhar com genomas com sequências repetitivas, além da possibilidade de sequenciamento de materiais completos (SMS) e da redução da necessidade de multiplicação genômica (XIAO e ZHOU, 2020; ATHANASOPOULOU et al., 2021).

Por fim, vale ressaltar que nenhum desses métodos são uma abordagem perfeita. Isso se deve a complexidade de montagem que envolve sequenciamentos de segunda geração, ainda mais agravado quando o assunto são genomas mais complexos; a necessidade de uma biblioteca mais robusta, comparado com os métodos de terceira geração; e ao menor nível de confiabilidade nas leituras geradas por métodos de terceira geração (XIAO e ZHOU, 2020; ATHANASOPOULOU et al., 2021; WANG et al., 2021).

A análise de dados provenientes da genômica possui como principais objetivos a caracterização e estudo de determinado genoma para que suas características sejam melhor compreendidas, indicando potenciais favoráveis ou formas de tratamento de doenças causadas, no caso de patógenos (CHIVIAN et al., 2023). Nesse momento iremos abordar vias oportunas que envolvem as ciências de dados como instrumento para a análise de genomas.

O uso de análise computacional de dados, através de bancos de dados robustos, como GenBank, trouxe e trazem diversas transições positivas para a área, possibilitando a origem de algoritmos de análise comparativa, que serão melhor trabalhados posteriormente, análise de contaminação e a criação de tecnologias de Inteligência Artificial (IA) cada vez mais precisas e eficientes, decorrente da maior disponibilidade de dados e desenvolvimento do próprio campo de estudo. Dentre as técnicas envolvendo IA podemos destacar as técnicas de aprendizado profundo, que através de diferentes algoritmos de redes neurais artificiais, Redes Neurais Recorrentes (RNNs), Redes Neurais Convolucionais (CNNs) e algoritmo de atenção, permitem o avanço da área de pesquisa *in silico* nesse sentido, apesar da limitação muitas vezes imposta pela exigência de hardwares poderosos para trabalhar com tais algoritmos (SAYERS et al., 2023; CHOI e LEE, 2023).

Na saúde global, todas essas técnicas de sequenciamento e métodos de análises mais sofisticados originaram o conceito de vigilância genômica, definido como o monitoramento da evolução, disseminação, mecanismos de resistência e até mesmo a predição de cepas circulantes mais prevalentes ou incidentes em determinado período de tempo em um território ou em todo o mundo. Um exemplo positivo disso é o projeto *Global Influenza Surveillance and Response System* (GISRS) da OMS para rastreamento do vírus da influenza (HAY e MCCAULEY, 2018).

A relevância disso é evidente quando observamos os casos de fungos e protozoários que vem demonstrado uma crescente resistência medicamentosa nas últimas décadas (STRUELENS et al., 2024). O fungo *Candida auris*, por exemplo, vem provocando surtos em hospitais por sua resistência a múltiplos medicamentos (FORSBERG et al., 2019) e *Aspergillus* oferece grande risco de infecções graves em pacientes imunocomprometidos (CADENA et al., 2021). Protozoários como *Trichomonas vaginalis* ou *T. foetus* deixados em segundo plano, tanto que o mecanismo que desenvolveram de resistência a antimicrobianos ainda não é totalmente compreendido, além do pouco conhecimento que adquirimos quanto a seus processos metabólicos e algumas de suas estruturas singulares (KIRKCALDY et al., 2012).

Em complemento podemos descrever diversos outros momentos históricos em que a presença dos conhecimentos de bioinformática contribuíram para o entendimento de patógenos não bacterianos em momentos que vários destes configuraram entre novas doenças emergentes, sendo destaque no campo de identificação, Emerging Pathogen Identification (EPI), trazendo contribuições relativas à eficiência de tratamento, vacinas e até na identificação da origem do organismo, podem ser citados como exemplo o SARS-CoV-2, vírus que causou a última pandemia, a partir do ano de 2019, ou a epidemia de Ebola no oeste do continente africano (VASHISHT et al., 2023), através da prática de análise de material genético, a maior fonte de informações de organismos vivos, os pesquisadores identificam genes responsáveis por processos fundamentais do organismo do patógeno, incluindo a os genes de resistência e através destas técnicas computacionais mais aprimoradas e

robustas se tornou possível direcionar estes genes de maneira mais eficaz com novos fármacos (ENGELTHALER, 2024).

O estudo do genoma de patógenos, evidentemente, reinventou a patologia e a farmacologia, diante disto, é indiscutível que dispor da troca de conhecimentos, recursos e tecnologias de diferentes países é sem dúvidas o fator de impulsão do progresso científico e de inovação pela promoção de grandes volumes de informações na íntegra e de diversidade das mesmas (LINS et al, 2021). Um exemplo claro disso é a Iniciativa Global de Resistência Antimicrobiana (GARDP) divulgada em 2016, que envolve a participação da Organização Mundial da Saúde (OMS) e Iniciativa de Medicamentos para Doenças Negligenciadas (DNDi), como também, dos governos da África do Sul, Alemanha, Japão e Reino Unido.

GARDP analisa e classifica os patógenos, através da genômica, quanto a maior necessidade da criação de novas terapêuticas e menor probabilidade de serem alvos de pesquisa de terceiros. E para isso, esse plano de ação que exige parceria entre público e privado, a fim de incentivar o desenvolvimento de novos tratamentos e diagnósticos para combate a resistência antimicrobiana, um dos maiores problemas de saúde pública da atualidade (PIDDOCK, 2019).

Outros projetos de grande importância foram e são desenvolvidos, no entanto, a pandemia do SARS-CoV-2 se tornou o exemplo mais recente e que destacou ainda mais a importância da união de esforços, não somente no campo da patogenômica, mas de toda a ciência, devido magnitude alcançada por este sucesso. Este evento, se converteu em um marco histórico e em uma grande lição de como a conexão das informações impactou e ainda impacta imensamente, a velocidade de resposta do mundo moderno a episódios infortúnios, inesperados e que resultam em consequências destrutivas para a humanidade (SILVA e LIMA, 2021).

3. Áreas de Pesquisa em Patogenômica

3.1. Análise de Genomas de Patógenos

3.1.1. Sequenciamento e Montagem de Genomas

Ao longo do desenvolvimento do campo *in-silico* diversas tecnologias já foram utilizadas para conseguir determinar a sequência de certo fragmento de material genético. É possível dividir as tecnologias de sequenciamento em 2 grupos amplos: leituras pequenas, as quais surgiram na segunda geração, e leituras longas, que surgiram na terceira geração (SATAM et al., 2023). A seguir serão trabalhadas algumas nuances relativas aos métodos de maior destaque. Porém antes de explorar com mais detalhes os processos de sequenciamento é necessário ressaltar e destacar a importância do controle de qualidade na preparação de amostras e bibliotecas de sequenciamento (LARSON et al., 2023; HESS et al., 2020).

As tecnologias de segunda geração (SGS) trouxeram maior dinamicidade, logo maior capacidade de produção, comparado aos seus predecessores, tendo como características principais o uso de leituras pequenas, grande precisão de sequenciamento e sequenciamento paralelo, apesar da complexidade considerável no momento de montagem, devido às leituras, que possuem tamanho reduzido. Os nomes usados para ilustrar a segunda geração são Illumina e Ion Torrent (HU et al., 2021; SATAM et al., 2023).

Já os métodos de terceira geração (TGS), que trabalham com leituras longas, iremos destacar a tecnologia PacBio e Oxford Nanopore. Ambos os métodos oferecem como vantagem a capacidade de montagem mais simples, quando comparado aos SGS, dado o número menor de leituras a serem arranjadas, que, também, por serem mais longas preservam consideráveis sequências de nucleotídeos, apesar da precisão reduzida em comparação aos métodos das outras gerações. Além disso, os TGS possuem um ciclo de sequenciamento mais veloz (HU et al., 2021; SATAM et al., 2023).

Vale ressaltar, também, a possibilidade de abordagens híbridas de segunda e terceira gerações, que apesar de proporcionar um encarecimento do processo, combinam as características positivas de ambas, logo atenuando os pontos fracos da outra. Por exemplo podemos citar o eficiente papel que

sequenciamentos de terceira geração tem em contribuir no fechamento de gaps em sequências de segunda geração montadas, estas que, seguindo outra metodologia, ajudam a fornecer maior confiabilidade e corrigir potenciais erros de sequenciamentos de terceira geração (WANG et al., 2021).

A análise de dados gerados é feita com o uso de softwares apropriados para tratamento e curadoria. Na primeira etapa da análise de *short reads* é feito uma etapa digital de controle de qualidade e chamada de base, onde se infere a sequência de bases das leituras (CACHO et al., 2016), já na segunda é feito o alinhamento das sequências com o uso de um algoritmo de tabela *hash* e arquivo de formato BAM, para armazenamento dos resultados. Por fim, na terceira etapa é realizado um processo de chamada de variantes, buscando identificar mutações baseadas em algum genoma de referência (HU et al., 2021; LARSON et al, 2023; XU, 2018).

Após o processo bem sucedido de sequenciamento chegamos à etapa de montagem do genoma, que pode ser sintetizada como a tentativa de ligação das “peças”, as leituras, para encontrar a ordem correta delas e assim possuir os genomas completos (SEGERMAN, 2020). As metodologias para realização de montagem são *de novo*, baseado em sobreposição de leituras e sistemas de probabilidade, de referência, que utilizam um genoma semelhante como referência para o processo, e via transcriptoma, onde o RNAseq da espécie é a ferramenta de suporte para montagem (ESPINOSA et al., 2024; ZHANG et al., 2022). Alguns dos montadores de destaque ao longo do século são: Celera assembler(2000), *De novo*(Newbler)(2005), Velvet(2008), ALLPATHS(2008), SPAdes(2012), HGAP(2013), Skesa(2018) e Unicycler(2020) (SEGERMAN, 2020).

Quando tratamos da montagem de genomas mais complexos vemos que, por exemplo no caso de plantas, o tamanho do genoma, as diversas sequências com mesmas bases ao longo do material genético e genomas heterozigotos nas espécies são alguns fatores que crescem na complexidade da montagem (68). A grande demanda computacional do processo, além do fechamento de gaps e erros de montagem são desafios (ESPINOSA et al., 2024).

Os sequenciadores de terceira geração, ao passo que suas técnicas de *reads* longos trazem um atenuação para o problema das sequências curtas,

montagem dificultada pelo tamanho do genoma e áreas de sequências idênticas (67,69). Outra estratégia utilizada se faz a partir de uma progressão na tecnologia dos montadores para trabalhar diferentes tipos de genomas complexos e se complementarem no resultado final (ZHANG et al., 2022).

3.1.2. Análise Comparativa de Genomas

A genômica comparativa tem como finalidade mapear, comparar e analisar a estrutura e as regiões genéticas funcionais de genomas individuais de uma única unidade taxonômica, identificando assim as similaridades e diferenças entre elas (ABBY e DAUBIN, 2007). As análises destinadas à comparação de genomas têm evidenciado a existência de uma enorme diversidade genética mesmo entre linhagens de uma mesma espécie, além de uma grande complexidade na estrutura e organização desses genomas, o que tem levado a questionamentos importantes que contribuem para uma melhor compreensão acerca dos mecanismos de evolução (PRASAD et al., 2023).

Grande parte das principais ferramentas utilizadas na comparação de sequências de DNA, RNA ou proteína se baseiam em métodos clássicos de análise computacional, tais como, algoritmos de alinhamento local e global, e algoritmos de alinhamento simples e de múltiplas sequências.

No alinhamento local a comparação entre as sequências ocorre através de pequenas regiões (SMITH e WATERMAN, 1981). Os algoritmos de alinhamento local geralmente são muito utilizados para encontrar semelhanças entre regiões que podem estar relacionadas, mas que são invertidas ou reorganizadas umas em relação às outras (BRAY et al., 2003). Já o global ocorre quando os genomas são comparados levando em consideração toda a sua extensão, sendo recomendado para sequências similares e que apresentam tamanho aproximado (CHOWDHURY e GARAI, 2017).

Já os alinhamentos simples e de múltiplas sequências dizem respeito ao número de sequências que são comparadas durante o alinhamento. Enquanto o alinhamento simples compara apenas duas sequências entre si, o alinhamento múltiplo compara várias sequências ao mesmo tempo. A partir do resultado do alinhamento de múltiplas sequências, pode-se inferir a homologia e as relações

evolutivas entre as sequências analisadas (PAIS et al., 2014). Por outro lado, as ferramentas de alinhamentos simples são usadas para identificar regiões de similaridade, as quais podem indicar relações funcionais, estruturais e/ou evolutivas entre duas sequências biológicas (MULLAN, 2006).

Abaixo algumas ferramentas de bioinformática amplamente utilizadas para comparação genômica que se baseiam nos algoritmos citados acima:

- **BLAST (Basic Local Alignment Search Tool):** é uma das ferramentas de análise de sequências biológicas de domínio público mais utilizadas hoje (ALTSCHUL et al., 1990).
- **Clustal Ômega:** é a ferramenta mais recente dentro da família Clustal e foi desenvolvida com a finalidade de permitir o alinhamento de uma grande quantidade de sequências sem que houvesse um comprometimento nos resultados dos alinhamentos (SIEVERS et al., 2011).
- **MAFFT:** é uma ferramenta muito utilizada para realizar o alinhamento de múltiplas sequências, essencial em estudos de filogenia, estrutura de proteínas e análise funcional de genes, desempenhando um papel essencial na compreensão da biologia molecular (KATOH, 2002).

A análise comparativa de genomas tem sido muito utilizada na identificação de genes e regiões conservadas entre os genomas comparados visando compreender melhor a evolução e a relação filogenética entre os microrganismos, assim como no desenvolvimento de novas drogas e alvos terapêuticos (JUHAS et al. 2012). Os genes e regiões conservadas são aquelas que persistem no genoma de várias linhagens de um grupo e que apresentam mutações mínimas ao longo do tempo. Geralmente codificam proteínas que desempenham papéis significativos em vários processos celulares críticos e que são essenciais para a sobrevivência de um organismo, tais como transcrição e tradução do DNA, e proteínas envolvidas na respiração celular (ROSCONI et al., 2022).

Uma técnica amplamente utilizada e que se baseia na amplificação, sequenciamento e análise de genes conservados entre isolados de uma mesma

espécie é a tipagem molecular MLST, onde isolados são tipados a partir da sequência dos seus genes *housekeeping*. (Multilocus sequence typing) (MAIDEN et al. 1998). Essa metodologia é adequada para analisar a distribuição geográfica e temporal dos STs circulantes, clonalidade, assim como para obter um panorama sobre a epidemiologia e diversidade genética, auxiliando na investigação de casos esporádicos e surtos de várias doenças (IBARZ e MAIDEN., 2009).

A identificação de genes conservados entre as linhagens de uma mesma espécie também é uma das etapas para a análise do pangenoma, que é considerado atualmente o padrão ouro para a comparação de genomas e estudos de evolução e diversidade. O pangenoma corresponde ao repertório completo de genes observados nos genomas de organismos de uma determinada espécie ou grupo taxonômico (ROULI et al, 2015). Esse conjunto de genes é constituído pelo core genome, shared e singleton, onde o core é compartilhado por todos, o compartilhado é presente em múltiplas, mas não todas, linhagens, e o singleton é exclusivo a linhagens específicas (TETTELIN et al., 2005). Identificar quais genes são conservados e compartilhados por todas as linhagens é parte essencial para a escolha de genes adequados ou proteínas candidatas para o desenvolvimento de vacinas contra várias doenças infecciosas através da abordagem da vacinologia reversa (D'MELLO et al., 2019).

A plasticidade da molécula de DNA possibilita a variabilidade genética e a diversidade, que são fundamentais para o processo evolutivo das espécies e adaptação ao ambiente / hospedeiro (SELA et al., 2018). Uma vez que as características da molécula do DNA e seu processo de replicação contribuem para a continuidade da informação, outros mecanismos atuam para gerar variabilidade genética. O primeiro mecanismo engloba as estratégias que provocam alterações locais na sequência do DNA, tais como substituição, deleção ou inserção de alguns nucleotídeos. O segundo mecanismo envolve estratégias que promovem o rearranjo de sequências de DNA, enquanto o terceiro mecanismo baseia-se na abordagem da aquisição de DNA por transferência horizontal (ARBER, 2000). A **Tabela 1** tem exemplos de ferramentas capazes de identificar variantes nos genomas.

Tabela 1. Ferramentas para identificação de regiões conservadas e variantes genéticas

Ferramenta	Aplicação
VariantDetective (VD) (CHARRON e KANG et al., 2024)	Identificação de polimorfismos de nucleotídeo único (SNPs) e variantes estruturais (SVs)
HGTphyloDetect (YUAN et al., 2023)	Identificação e análises filogenéticas de genes obtidos por transferência horizontal.
progressiveMauve (DARLING et al., 2010)	Identificação e alinhamento de regiões conservadas do genoma na presença de eventos de rearranjos e transferência horizontal

3.1.3. Estudo de Variantes Genéticas

As variantes genéticas desempenham um papel fundamental na evolução, especialmente em microrganismos, onde mutações e recombinações promovem a variabilidade genética, mesmo com a reprodução assexuada. Essas variantes podem ser mutações pontuais, como polimorfismos de nucleotídeo único (SNPs), deleções, inserções de nucleotídeos, ou rearranjos de grandes segmentos gênicos e cromossômicos (ARENAS et al., 2018).

Uma mutação pontual refere-se a uma alteração permanente em uma única base do DNA de um gene, ou à deleção ou inserção de uma nova base. Quando uma mutação em um gene resulta em um códon que codifica o mesmo aminoácido, ela é chamada de mutação sinônima. Já quando o códon é alterado para codificar um aminoácido diferente, trata-se de uma mutação não sinônima, que pode modificar a estrutura ou função da proteína, influenciando diretamente a evolução do microrganismo (CHAHAL et al., 2019).

Por outro lado, a inserção ou deleção de bases (indels) pode alterar a sequência de códons e a fase de leitura do gene, o que pode modificar drasticamente a estrutura e função de uma proteína. Em alguns casos, isso pode gerar um códon de parada prematuro, resultando em uma proteína truncada, ou seja, interrompida precocemente durante sua síntese (CHEN et al., 2009).

As mutações geralmente surgem a partir de erros no processo de replicação do DNA, os quais podem ou não ser corrigidos pelos mecanismos de reparo da célula. Dessa forma, as taxas de mutação variam significativamente entre diferentes microrganismos (CHATTERJEE e WALKER, 2017). Embora mutações que afetam a sequência de aminoácidos muitas vezes possam ser prejudiciais, ocasionalmente elas proporcionam vantagens adaptativas, como maior capacidade reprodutiva ou maior sobrevivência em ambientes hostis (YOSHIDA et al., 1990).

Eventos de recombinação, como a transferência horizontal de genes em bactérias (por transdução, conjugação ou transformação), também impactam de forma significativa a variabilidade genética, ao gerar novas combinações ou reorganizar variações existentes que surgiram por mutações. Além disso, há as CNVs (variações no número de cópias), que são variantes estruturais desequilibradas caracterizadas pela perda ou ganho de fragmentos de DNA que variam de cerca de 50 pares de bases até cromossomos inteiros. (WU et al., 2020; TRALAMAZZA et al., 2024).

Um dos principais métodos para identificação de variantes é a chamada de variantes (do inglês, Variant Calling, VC). A chamada de variantes é o processo de comparar duas amostras de genoma para identificar suas diferenças. Geralmente, essas diferenças são restritas a polimorfismos de nucleotídeos únicos (SNPs) e pequenas inserções e deleções (indels). Alterações estruturais mais amplas, como inversões, duplicações e grandes deleções, normalmente não são incluídas na análise de chamada de variantes. Um dos métodos comumente usados é a ferramenta Snippy, disponível publicamente (<https://github.com/tseemann/snippy>), que compara dados de sequenciamento genômico contra um genoma de referência para identificação dos SNPs e Indels. Por esse motivo, a escolha de genomas de referência montados corretamente é essencial, pois todo o resultado da chamada de variantes será em relação ao genoma de referência.

Existem ainda ferramentas mais robustas, que implementam a chamada de variantes com base em diferentes métodos. As ferramentas atuais se dispõem em três categorias principais. A ferramenta bcftools (disponível em <https://github.com/samtools/bcftools>) por exemplo, realiza por meio de uma

estimativa de frequência de alelos com base em leituras mapeadas. A DeepVariant (POPLIN et al., 2018) usa um modelo de rede neural convolucional treinado em leituras mapeadas. E DiscoSNP (disponível em <https://github.com/GATB/DiscoSnp>) utiliza um método baseado em grafos de Bruijn de k-mers de leituras brutas. Demais ferramentas se baseiam em mudanças a partir desses métodos (SEAH et al., 2023). Esses métodos, por exemplo, são eficazes para chamadas de variantes em genomas bacterianos. Para identificar variações no número de cópias (CNV), há ainda para genomas bacterianos a ferramenta CNV-BAC (WU et al., 2020).

Em resposta à pandemia da COVID-19, muitas ferramentas surgiram para a chamada de variantes em genomas virais. A ferramenta ViReflow (MOSHIRI et al. 2022) por exemplo, é um pipeline de reconstrução e análise de sequência de consenso viral. O ViReflow foi criado em resposta à pandemia de COVID-19, mas pode ser aplicado a qualquer patógeno viral (LI, 2011; WILM et al., 2012).

Variantes genéticas, como SNPs, Indels e CNVs, presentes em regiões genômicas associadas à virulência ou resistência ao ambiente, estresse e antibióticos, no caso de bactérias, podem conferir vantagens ou desvantagens na adaptação e interação com o hospedeiro. Um exemplo recente vem da Índia (105), onde três linhagens de diferentes espécies de *Leptospira* spp. foram sequenciadas: *Leptospira interrogans* sorovar Pyrogenes cepa Salinem, *Leptospira santarosai* sorovar Huanuco cepa M-4, e *Leptospira kirschneri* sorovar Ratnapura cepa Wumalasena.

Além do sequenciamento, o estudo também incluiu a montagem de novo, a anotação estrutural e funcional, além da análise de variantes. Dentre as encontradas, foram relatadas variantes em proteínas que desempenham várias funções importantes em bactérias, como invasão celular, ancoragem à membrana e interação com o hospedeiro. A presença de um grande número de SNPs em genes associados à virulência indica que isso pode ser um dos fatores que contribuíram para o surgimento de linhagens altamente patogênicas. No entanto, são necessários estudos complementares para validar essa correlação (LATA et al., 2020).

Tratando-se de resistência a antibióticos, alguns estudos já aplicaram a chamada de variantes a genes de resistência com sucesso (NAIDENOV et al.,

2019). Por exemplo, a *Piscirickettsia salmonis*, patógeno intracelular Gram-negativo causador da Piscirickettsiose, é tratada frequentemente com altas doses de antibióticos. Análises genômicas *in silico* identificaram sete genes relacionados à resistência ao florfenicol e oxitetraciclina. Foram encontradas cepas com diferentes níveis de suscetibilidade aos antibióticos, e análises *in silico* revelaram que mudanças estruturais nas proteínas associadas à resistência, como OmpF, estavam relacionadas à presença de SNPs (FIGUEROA et al., 2019).

3.2. Análise de Interações Patógeno-Hospedeiro

O termo Interação Hospedeiro-Patógeno (IHPs) refere-se a abordagens diretamente relacionadas à forma como um determinado patógeno interage com seu organismo hospedeiro. Existem diferentes tipos de patógenos, como bactérias, fungos, vírus e outros tipos de organismos que, ao conseguirem se aderir ao sistema do hospedeiro, realizam diversas alternativas para a autopromoção da sobrevivência e da infecção, causando doenças em seu hospedeiro devido às suas características patogênicas. Nesta seção, analisaremos os principais aspectos comumente presentes nessas condições.

As IHPs podem ser descritas em níveis: populacional (como quando ocorreu a pandemia de Covid, em que o vírus Sars-Cov-2 causou infecção em uma população (WHO, 2024), organismal (diferentes patógenos infectando o mesmo hospedeiro) e/ou molecular (quando as proteínas do patógeno interagem com as proteínas do hospedeiro, com receptores presentes nas células humanas) (SEM et al., 2016). Além disso, os IHPs têm “estágios” que são respeitados entre os vários patógenos e hospedeiros possíveis, em ordem: invasão do patógeno no hospedeiro ao superar as barreiras primárias; evasão das defesas imunológicas do hospedeiro pelo patógeno; 3. replicação do patógeno e 4. capacidade imunológica do hospedeiro contra o patógeno (SEM et al., 2016). Os meios pelos quais os patógenos realizam esses feitos estão intimamente associados aos seus fatores de virulência.

3.2.1. Identificação de Fatores de Virulência

Fatores de virulência são moléculas específicas, na sua grande maioria representadas por proteínas produzidas e secretadas por bactérias, fungos, protozoários e vírus. Em patógenos bacterianos, os fatores de virulência são codificados por genes específicos que podem estar localizados no cromossomo ou em elementos genéticos móveis, tais como, plasmídeos e transposons (LEITÃO, 2020).

Os fatores de virulência desempenham diversos papéis importantes no estabelecimento de doenças, incluindo a colonização do hospedeiro, imunossupressão, aquisição de nutrientes da célula hospedeira e detecção de mudanças no ambiente (ALSPAUGH, 2015). Várias moléculas estão envolvidas neste cenário de interação patógeno-hospedeiro, englobando adesinas, toxinas, enzimas e cápsulas (CHEN, 2004).

Através da análise da sequência genômica de diversos tipos de patógenos, podemos obter uma melhor compreensão da interação patógeno-hospedeiro. Com isso, a identificação e caracterização precisa de genes e proteínas de virulência tem se tornado cada vez mais essencial na busca por novos alvos terapêuticos, para o desenvolvimento de testes mais específicos e direcionados, no desenvolvimento de vacinas ou para tratamentos que possam ser utilizados como alternativa aos antimicrobianos (OLI et al., 2020).

Com isso, várias ferramentas têm sido desenvolvidas para identificar com melhor precisão fatores de virulências que possam ser cruciais para a patogenicidade. Abaixo algumas dessas ferramentas:

- **Vfanalyzer:** O VFanalyzer constroi grupos ortólogos dentro do genoma de consulta e genomas de referência pré-analisados do banco de dados de fatores de virulência (VFDB) e, em seguida, conduz pesquisas de similaridade de sequência no VFDB para identificar com precisão potenciais fatores de virulência (CHEN, 2004).
- **VirulentPred 2.0:** O modelo de predição utilizado por este servidor web é baseado no treinamento de modelos de aprendizado de máquina com sequências de proteínas virulentas validadas experimentalmente (SHARMA et al., 2023).

Os fatores de virulência aumentam a habilidade de um patógeno em causar doença. De acordo com os mecanismos e suas funções, os fatores de virulência podem ser divididos nas seguintes categorias [Tabela 2]:

- (1) Adesinas: se ligam especificamente a receptores complementares encontrados na superfície de certos tipos celulares, promovendo assim a adesão da bactéria aos tecidos do hospedeiro (KLINE et al., 2009).
- (2) Evasinas: auxiliam as bactérias no processo de escape da fagocitose ou do sistema imune do hospedeiro (CHANDRA et al., 2022).
- (3) Invasinas: rompem as membranas do hospedeiro e estimulam a endocitose. (GARCIAL DEL PORTILLO e FINLAY, 1994).
- (4) Toxinas que provocam danos às células e tecidos do hospedeiro (DONG et al., 2019)
- (5) Polissacarídeos capsulares que protegem os patógenos das defesas do hospedeiro (PATON et al., 2019).
- (6) Sideróforos: secretados por algumas bactérias para a obtenção de ferro, necessário para o seu crescimento (ELLERMAN e ARTHUR, 2017).

Tabela 2. Exemplos de fatores de virulência em patógenos conhecidos

Espécie	Fator de virulência
<i>Escherichia coli</i> enteropatogênica	Proteínas intimina e seu receptor Tir: adesão à superfície do epitélio intestinal (PAKBIN et al., 2021)
<i>Staphylococcus aureus</i>	Toxinas: altamente inflamatórias e que podem causar morte de células leucocitárias por citólise e deleção clonal, respectivamente (TAM e TORRES, 2019).
<i>Neisseria gonorrhoeae</i>	Fimbrias: adesão à mucosa urogenital (TRAMONT, 2008). Quando as fimbrias estão ausentes, devido à mutação por exemplo, a colonização muitas vezes não ocorre, e a doença não evolui.
<i>Bacillus anthracis</i>	Cápsula de ácido D-glutâmico: inibição da fagocitose das células vegetativas pelos macrófagos durante a infecção por antraz (EZZELL e WELKOS, 1999).

Os fatores de virulência têm grande impacto no grau de patogenicidade das bactérias e até mesmo na progressão de algumas doenças [Figura 2]:

A infecção crônica pelo *Helicobacter pylori* é um fator essencial na patogênese do carcinoma gástrico, devido ao gene CagA fator de virulência relacionado a um risco aumentado para o desenvolvimento tumoral (WANG et al., 2014). Outro exemplo de patógeno com um fator de virulência bastante relevante é o *Corynebacterium diphtheriae*, principal agente etiológico da difteria, uma doença potencialmente fatal ao hospedeiro, principalmente devido à ação da toxina diftérica (TD), uma exotoxina produzida pelo microrganismo (SHARMA et al., 2019, SUGIMAN-MARANGOS et al., 2022).

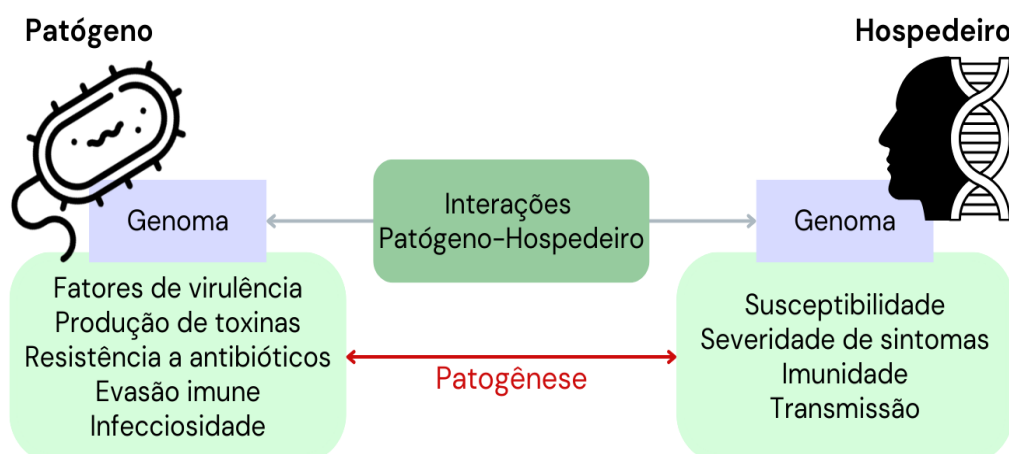


Figura 2. Interações genoma-genoma no contato entre patógeno e hospedeiro

3.2.2. Interações Moleculares

Técnicas para identificação de interações proteína-proteína (PPIs), como a co-imunoprecipitação, ensaios *pull-down*, EMSA (electrophoresis gel mobility shift analysis) e *DNase I footprinting* são bem desenvolvidas e estabelecidas em vários campos de estudo, mas técnicas *in vitro* tendem a ser processos muito caros e trabalhosos, pois exigem um grande número de anticorpos e equipamentos específicos, além da otimização das condições e de outros fatores. Além disso, o grande número de proteínas do hospedeiro e do patógeno representa um problema experimental em tempo real. Nesse sentido, a pesquisa

de bioinformática tende a ser bem-sucedida em vários tipos de análise. Assim, foram desenvolvidas abordagens para realizar a previsão *in silico* de PPIs com base no aprendizado de máquina, inicialmente por Bock et al. (2001) por meio do uso da Máquina de Vetor de Suporte (SVM) para treinar e prever interações com base na estrutura primária e nas propriedades físico-químicas relacionadas.

As técnicas de previsão de PPIs tendem a ser eficientes na triagem primária das interações realizadas entre o patógeno e o hospedeiro, fornecendo dados prováveis que podem ser verificados experimentalmente após o financiamento. Assim, estudos preditivos de PPIs putativos em HPIs são importantes para observar se determinadas vias tendem ou não a ser influenciadas por um determinado patógeno. Há vários tipos de previsão, a saber:

- **Aqueles baseados em homologia:** Quando uma determinada interação entre proteínas tende a ocorrer em sequências similares. (LEE et al., 2008; MATTHEWS et al., 2001).
- **Aquelas baseadas na estrutura:** Quando um determinado par de proteínas tem estruturas semelhantes a um par de proteínas que já interagem (DAVIS et al., 2007, DOOLITTLE e GOMEZ, 2010).
- **Aquelas baseadas na interação entre motivos conservados:** esses tipos de previsões integram PPIs intraespecíficas conhecidas com perfis de domínio/motivo de proteína que interagem (DYER et al., 2007, EVANSet al., 2009).
- **Aqueles baseados em aprendizado de máquina:** Eles extraem dados recuperados de interações entre hospedeiro e patógeno (THIEU et al., 2012, CHAUSSABEL et al., 2003).

3.2.3. Respostas Imunes e Evasão Imunológica

A resposta imunológica do hospedeiro contra um patógeno começa com o macrófago residente próximo ao sítio de entrada do patógeno no corpo humano e, ao observar isto, realiza fagocitose do patógeno visando controlar e “alertar” novas unidades de leucócitos sobre a infecção. Após a fagocitose, o macrófago

libera citocinas e quimiocinas que deixam o local da infecção e se ligam a proteoglicanos presentes na superfície do endotélio do vaso sanguíneo mais próximo que “espera” que um leucócito passe por esse vaso (ABBAS, 2018; KOLACZKOWSKA e KUBES, 2013; MOORE et al., 1995). Esse primeiro estágio é conhecido como o reconhecimento pelo sistema imunológico inato da existência do patógeno no corpo.

Posteriormente, a célula dendrítica (uma das APCs - Antigen Presenting Cells - Células Apresentadoras de Antígeno) leva o epítipo do patógeno, processado e expresso no MHC I ou II (dependendo do tipo de célula T que se deseja ativar), para os órgãos linfóides secundários, como os linfonodos, onde apresentará o epítipo aos linfócitos T ingênuos e desencadeará uma resposta imune celular, como a produção de anticorpos. Após a apresentação, os linfócitos sofrem expansão clonal, aumentando o número de linfócitos ativados específicos para esse agente infeccioso. Essa ação leva alguns dias. As células do sistema imunológico interagem entre si por meio de quimiocinas/citocinas, coordenando as atividades da resposta imune inata e adaptativa, uma vez que todas as células do sistema imunológico são capazes de secretar essas citocinas e receber sinais por meio de receptores para elas (ABBAS, 2018).

Assim, sabemos que, durante o curso de uma infecção, o que determinará o sucesso e a capacidade do patógeno de sobreviver e se replicar em um ambiente hostil, devido à ação do sistema imunológico do hospedeiro, depende de como será a interação entre os dois, principalmente um hospedeiro e um patógeno, que determina a capacidade do micróbio de sobreviver em um ambiente hostil e promove uma resposta imunológica do hospedeiro como medida de proteção. Por exemplo, no caso da *Burkholderia spp*, a ligação inicial aos receptores de superfície celular presentes no hospedeiro desencadeia uma série de liberação de mediadores pró-inflamatórios, iniciando importantes vias de sinalização, como NF κ B, Erk e Akt, bem como a indução de um grande número de citocinas, como as interleucinas 1 β , 6 e 8 (IL-1 β , 6 e 8), o fator de necrose tumoral alfa (TNF- α) e etc. das células epiteliais alveolares e bronquiolares (LU et al., 2012; GILLETTE et al., 2013; SIM et al., 2009). Todas essas vias estão relacionadas à resposta imune inicial contra o patógeno que acabou de entrar no corpo, com organização quimiotática destinada a controlar

e acabar com a infecção (DAVID et al., 2015), no entanto, ele tende a desenvolver mecanismos de evasão imune para sobreviver.

Uma das estratégias de imunização mais eficientes e mais comentadas, é baseada no desenvolvimento de vacinas. Vacinas são preparações biológicas projetadas para prevenir doenças infecciosas, através de treinamento do sistema imune a reconhecer e eliminar futuras infecções de maneira rápida e eficaz. A imunização através de vacinas é uma das ferramentas mais importantes da saúde pública, responsável por erradicar doenças como a varíola e controlar várias outras, como o sarampo e a poliomielite (DOWDLE, 1998; ROUSH, 2007; TALBIRD et al., 2022).

Há vários tipos de formulações vacinais possíveis, pois há inúmeras maneiras de fazer o sistema imunológico reconhecer partes do patógeno e induzir a sua eliminação. Um exemplo clássico e recente, foi quando, em relativamente um curto período de tempo, estavam disponíveis diversos tipos de vacinas contra o Sars-Cov-2, devido a pandemia que se iniciou em meados de 2019. A saber, foram desenvolvidas formulações do tipo “Vírus Inativado”, isto é, o antígeno como versão do vírus SARS-CoV-2 inativado; “Vivo atenuado”, com SARS-CoV-2 enfraquecido; de “Subunidade de proteína”, com a proteína spike do SARS-CoV-2 presente; “Partículas semelhantes a vírus (VLPs)” havendo múltiplas proteínas do SARS-CoV-2 na formulação vacinal; Vetor viral, DNA e RNA, contendo o ácido nucléico do vírus (RAHMAN et al., 2022).

Neste sentido, o desenvolvimento de uma vacina eficaz contra determinado patógeno deve levar em consideração características singulares do agente patológico, incluindo seus mecanismos de evasão imunológica. Ainda sobre vírus, sabe-se que o sistema de produção de interferon tipo I (IFNs- α/β) é ativado pelo reconhecimento de componentes virais por fatores da célula hospedeira (SAMUEL, 2001; SADLER, 2008). Contudo, um dos mecanismos de evasão do sistema imunológico pelos vírus é justamente o desenvolvimento de antagonismo de IFN viral por meio da codificação de proteínas virais capazes de antagonizar a via do IFN (SADLER, 2008; XU et al., 2021).

Ao considerar o desenvolvimento de vacinas para esses grupos de patógenos, é fundamental que elas estimulem tanto a resposta imune inata quanto a adaptativa. No campo da medicina veterinária, as vacinas vivas

atenuadas são geralmente preferidas por serem consideradas mais eficazes do que as inativadas. Essas vacinas vivas induzem tanto a imunidade humoral quanto a celular, enquanto as inativadas promovem principalmente uma resposta de anticorpos. A remoção do antagonismo de interferon (IFN) em vírus tem sido uma estratégia usada para aprimorar a imunogenicidade e aumentar a proteção em vacinas humanas e veterinárias (SU et al., 2023). Um exemplo é a proteína NS5 do vírus da dengue (DENV) no Zika virus (ZIKV), que interage com a proteína STAT2 humana e inibe a via do IFN, destacando-se como um importante fator de virulência e um possível alvo para o desenvolvimento de vacinas atenuadas contra flavivírus (GRANT et al., 2016; ASHOUR et al., 2010).

4. Aplicações da Patogenômica

4.1. Diagnóstico e identificação de patógenos

Organismos patogênicos, tais como bactérias e fungos, eram identificados pelas suas características fisiológicas e morfológicas, à partir da invenção da microscopia, e em métodos como a coloração de Gram, que resultavam em diagnósticos tardios, só possíveis após a manifestação dos primeiros sintomas, causando desconforto para o paciente, dificuldade para o tratamento (podendo levar à morte) e prevenção quanto à transmissão do patógeno. Os avanços na medicina tornaram possível a identificação das infecções de maneira mais precoce, através da identificação de biomarcadores que indicam uma infecção (DE NOOIJER et al., 2023), mas caracterizar e estudar qual patógeno está causando a infecção ainda é difícil, tendo em vista as características do meio ideais de cada microrganismo, como pH, temperatura, salinidade, oxigenação, entre outros (LIU et al., 2022). Os avanços no campo das ômicas (genômica, transcriptômica, proteômica, entre outros) estão tornando possível uma identificação mais precisa dos genes dos patógenos, com o desenvolvimento de programas e algoritmos mais eficientes, além da expansão dos bancos de dados. Pesquisadores estão sendo capazes de identificar marcadores e precursores de metabolismo, sendo possível o tratamento de infecções antes do surgimento de sintomas agudos ou marcadores toxicológicos que só aparecem após a doença, assim aumentando a eficiência dos tratamentos.

Os métodos convencionais envolvem técnicas associadas ou não a meios de cultura (NEHRA et al., 2022). Os mesmos podem ser baseados em imunologia, como por exemplo a detecção por ligação a anticorpos, que permite uma detecção não-invasiva (coleta de saliva), porém a confiabilidade só é alta após um certo tempo depois da infecção inicial, e os resultados perdem a exatidão dependendo da resposta imune dos hospedeiros, que pode não ser tão intensa (LEVA-BUENO et al., 2020). Os métodos baseados em ácidos nucleicos, tais como a reação em cadeia de polimerase (PCR), são atualmente os mais utilizados, especialmente após a pandemia do Sars-Cov (ZENG et al., 2024), e são fundamentais no trabalho com patógenos que não crescem *in vitro*, onde é possível amplificar e identificar genes específicos, porém também dependem do período de infecção para detectar a presença do microrganismo. Além disso, também são usadas técnicas relacionadas com espectrometria de massa, tais como a de MALDI-TOF, que apesar de rápida e relativamente barata, possui baixa discriminação entre espécies e sub-espécies de organismos, podendo levar a erros de identificação (ELBEHIRY et al., 2022). De forma geral, os métodos convencionais possuem uma série de dificuldades para serem executados, tais como o preço dos reagentes e equipamentos, o tempo para obter os resultados e a necessidade de pessoal qualificado, além da chance de resultar em falsos positivos/negativos por falha humana ou de produto.

Com os avanços na genômica, a partir do sequenciamento de nova geração (NGS), foi possível observar diferenças entre os genes de organismos, que, anteriormente eram classificados como taxonomicamente iguais, mas foram observadas grandes diferenças quanto a marcadores ribossomais e/ou metabólicos. Tendo bancos de dados com uma riqueza grande de genes confiáveis, veio à tona métodos de identificação taxonômica como a Identidade Média de Nucleotídeos (ANI), que realiza a identificação taxonômica a partir da semelhança dos genes ortólogos entre dois organismos (JAIN et al., 2018), e a Hibridização Digital de DNA-DNA (dDDH), que avalia a semelhança entre segmentos genômicos homologados para determinar a identidade nucleotídica média entre eles, e assim determinar a proximidade taxonômica com os genomas do banco de dados (MEIER-KOLTHOFF e GÖKER et al., 2019). Uma das aplicações da patogenômica é o monitoramento da presença e ausência de

genes de interesse, e no caso de patógenos, genes de virulência, resistência a antibióticos e elementos genéticos móveis (MGEs). Com essas informações é possível prever o comportamento de microrganismos patogênicos, e planejar o tratamento adequado para as doenças.

Em um estudo de *Daria Bottai et al* (BOTTAI et al., 2020), foi observado que diversas cepas de *Mycobacterium tuberculosis* presentes em epidemias globais de tuberculose teriam a deleção em comum de uma região gênica TbD1. Essa deleção, que aumenta a resistência a estresse oxidativo e hipóxia, é um fator evolutivo que pode ser um ponto de interesse no monitoramento desse patógeno, observando em cepas que ainda possuem o gene. Em outro caso, de *Migle Gabrielaite et al* (GABRIELAITE et al., 2021), foi observado que pacientes com fibrose cística estavam sendo infectados por *Achromobacter spp.*, gerando casos de resistência a antibióticos, por exemplo. Dessa forma foram analisados 101 genomas isolados de 51 pacientes, revelando padrões de mutação que indicam uma adaptação genética convergente, e esse padrão de comportamento pode ser fundamental para prever o resultado dessas mutações e como elas impactam na virulência do patógeno.

4.2. Desenvolvimento de terapias e vacinas

Com a melhoria nos bancos de dados, devido ao aumento do número de genomas e a qualidade dos mesmos, novas análises como a pangenômica estão se tornando viáveis para a identificação de novos meios de tratamento. Nas análises pangenômicas, a partir do estabelecimento de um grupo de organismos (de mesma espécie, por exemplo), se obtém um conjunto de genes essenciais (genes core) pertencentes a todos os indivíduos, fundamentais para a sobrevivência desses microrganismos. A implicação desse fato é, ao desenvolver um modelo vacinal, ou uma droga que atinge um gene essencial, em teoria todos os indivíduos do grupo são atingidos, assim criando um tratamento mais eficiente contra o patógeno. Os modelos de vacinologia estão se atualizando, tendo em vista que os modelos clássicos possuem um desenvolvimento muito custoso em termos de dinheiro e tempo. Com a patogenômica, é possível fazer a predição de quais são as proteínas de

membrana de um grupo de microrganismos, e assim desenvolver uma vacina que abrange todos.

Os microrganismos nosocomiais, ou seja, aqueles que estão relacionados à infecções em ambientes hospitalares, são um foco fundamental de monitoramento, pois estão constantemente expostos a novos hospedeiros, muitas vezes com outras infecções. Dessa forma, esses organismos recebem e transferem genes de forma horizontal, e com a adição do uso constante de antibióticos e drogas, o potencial de surgimento de patógenos com alta virulência e resistência a drogas é alto. Um estudo de *Cyril Lefebvre et al.* (LEFEBVRE et al., 2020) investigou fatores de patogenicidade de *Mycobacterium tuberculosis*, e foi observado que o gene *Rv0504c*, que produz a proteína *HadDMtb*, estaria relacionado com o metabolismo dessa bactéria, participando na formação de biofilmes e resistência a baixas temperaturas. Com essa evidência, a possibilidade de produção de um medicamento que pode inibir essa produção poderia reduzir a taxa de sobrevivência desse microrganismo. Certos tipos de microrganismos, como a espécie *Escherichia coli*, que possui uma grande variedade de cepas patogênicas, precisam ser monitoradas quanto ao seu potencial de se tornarem patógenos. J R Clark e A M Maresso (CLARK et al., 2021) construíram um patótipo a partir de 107 cepas reconhecidas como patógenos, e assim identificaram uma série de genes de interesse prevalentes nesses organismos (evasão do sistema imune, absorção de ferro, adesão e produção de toxinas). Com isso é possível monitorar se alguma cepa está se tornando patogênica, além do desenvolvimento de possíveis tratamentos para esses organismos.

4.3. Monitoramento e Controle de Epidemias

O monitoramento genômico tem transformado a forma pela qual monitoramos e controlamos epidemias infecciosas. Ao monitorar e analisar os genomas de patógenos, pesquisadores podem rastrear a origem de surtos, habilitando o controle de infecções (STRUELENS et al., 2024). Uma metodologia chave para esse monitoramento é o sequenciamento de genoma inteiro (WGS), que permite a pesquisadores o rastreamento de diferentes linhagens e variantes (NAFEA et al., 2024). Durante a pandemia do COVID-19, o WGS foi crucial para

rastrear variantes do vírus SARS-CoV-2, permitindo que cientistas e autoridades fizessem intervenções pontuais com agilidade (PAPA et al., 2023; YALÇIN et al., 2024). Portanto, os dados genômicos obtidos por WGS, em conjunto com modelos epidemiológicos, podem ser usados para prever e controlar epidemias, inclusive, no caso de modelos preditivos, de começar preparativos para controlá-las com antecedência (CARDENAS et al., 2022; WANG et al., 2022). Exemplos disso em tempos recentes foram os surtos de Ebola na África Ocidental e a pandemia da COVID-19, ambos casos em que a patogenômica teve função chave no controle de surtos, habilitando tomadas de decisão através de dados genômicos em tempo real (BARRETT et al., 2022, DI PAOLA et al., 2020).

No futuro, o conceito de patogenômica e seus insights poderá contribuir muito para o desenvolvimento de estratégias preventivas. O monitoramento do conteúdo genético de patógenos informando o desenvolvimento de vacinas e métodos diagnósticos, bem como a previsão de ameaças biológicas (FENG et al., 2024).

4.4. Pesquisa e Desenvolvimento de Novos Alvos Terapêuticos

Em tempos recentes, o desenvolvimento de novas plataformas de WGS permitiu a obtenção de conteúdo genômico de alta qualidade, o que viabiliza o uso da patogenômica para identificar novas proteínas ou genes como alvos de droga no genoma dos patógenos (ALTURKI et al., 2022; HYUN et al., 2023). Essas metodologias dependem do estudo cuidadoso do conteúdo genético dos patógenos em busca de genes chave que são essenciais para a sobrevivência do patógeno, fazendo deles excelentes alvos para intervenção terapêutica (LIANG et al., 2024; ALKATHERI et al., 2023) e também em busca de mecanismos de resistência a antibióticos no genoma (DJORDJEVIC et al., 2024). Algumas técnicas emergentes, como as técnicas de CRISPR-Cas baseadas no genoma e no RNA, são altamente precisas e ajudam na examinação dos genomas de patógenos, podendo indicar tratamentos ainda mais eficientes e precisos (LI et al., 2023; ZHU et al., 2022).

Além disso, com a ajuda do WGS e das descobertas em genômica, o campo de pesquisa farmacêutica visando a descoberta de novos medicamentos foi

revolucionado, levando ao desenvolvimento de alvos terapêuticos mais específicos. Por exemplo, a identificação *in silico* de novos alvos para o tratamento da sífilis por meio da análise das sequências do genoma de *Treponema pallidum*. (KUMAR et al., 2017).

5. Desafios e Perspectivas Futuras

A patogenômica, um campo emergente que combina dados genômicos com informação clínica e imunológica para avançar no diagnóstico e tratamento de doenças infecciosas, enfrenta desafios técnicos, metodológicos e éticos significativos, mas também oferece grandes promessas para a medicina de precisão. Este domínio tem evoluído ao longo dos anos através de constantes avanços tecnológicos, mas continuam a surgir novos desafios à medida que o volume de dados aumenta e cresce a necessidade da sua integração eficiente.

5.1. Desafios técnicos e metodológicos

Apesar dos avanços na tecnologia de sequenciação de nova geração (NGS), subsistem limitações consideráveis. A maioria dos métodos está limitada a dados unimodais, deixando a integração de modalidades, como os dados genômicos e imunológicos, relativamente pouco desenvolvida (FENG et al., 2024). Os desafios incluem a montagem exata de genomas complexos, como os de microrganismos com sequências repetitivas, e a dificuldade de integrar eficazmente informações fenotípicas e genotípicas (FENG et al., 2024). As ferramentas emergentes de inteligência artificial multimodal estão a começar a ultrapassar estas limitações, permitindo correlações entre características genômicas e patológicas, particularmente em oncologia.

Cornish e Guda (2015) sublinham a importância de pipelines bioinformáticas robustas para garantir a precisão na identificação de variantes genéticas, uma área crítica para o diagnóstico genómico preciso em agentes patogênicos (MILLER et al., 2010)

À medida que o volume de dados gerados pelas plataformas de sequenciação aumenta, são necessárias infraestruturas de computação potentes e algoritmos avançados para processar e interpretar estes dados em tempo útil. Um dos principais problemas enfrentados é a baixa proporção de DNA

patogénico nas amostras de infecção, o que exige tecnologia de sequenciação profunda e bioinformática avançada para reconstruir genomas patogénicos históricos ou incompletos (SATAN et al., 2023).

A montagem de genomas complexos, especialmente de agentes patogénicos, é outro grande desafio. As tecnologias atuais são frequentemente incapazes de lidar com genomas que têm sequências repetitivas ou altamente heterogêneas (LYNCH et al., 2016; SEREIKA et al., 2022; HAYASHI e OGURA, 2020). Destacam-se os desafios enfrentados pelos algoritmos de montagem, particularmente em genomas com sequências repetitivas, o que compromete a precisão dos dados de montagem de agentes patogénicos. Além disso, a identificação de variantes genéticas patogénicas é dificultada pela falta de genomas de referência bem anotados para algumas espécies (LYNCH et al., 2016).

O volume crescente de dados genômicos suscitou preocupações éticas sobre a privacidade, a segurança e a utilização desses dados (ODEH et al., 2021; WANGMO et al., 2018; MATHAIYAN et al., 2013). A proteção de dados sensíveis, como a informação genética relacionada com doenças infecciosas, exige o desenvolvimento de quadros éticos e jurídicos sólidos, especialmente no contexto da partilha global de dados durante emergências de saúde pública. A utilização da IA em grandes volumes de dados genômicos, como discutido por Vayena et al. (2018), levanta questões importantes sobre a privacidade e a segurança dos doentes.

Apesar dos avanços na bioinformática, continua a haver necessidade de ferramentas mais eficientes para analisar dados genômicos em grande escala e em tempo real. Embora os atuais pipelines bioinformáticos possam processar grandes volumes de dados, são limitados na sua capacidade de integrar de forma eficiente e normalizada diferentes tipos de dados multimodais (como dados clínicos, imagiológicos e genômicos) (HAYASHI e OGURA, 2020; MANIRUZZAMAN e NOKHODCHI, 2017, HE et al., 2023). A integração de dados de RNA-Seq com outras abordagens genômicas, tal como proposto por Wang, Gerstein e Snyder (2009), é uma estratégia importante para ultrapassar estas limitações e melhorar a análise de dados multimodais. Direcções futuras: Integração multimodal avançada e medicina de precisão

O futuro da patogenômica reside na integração total dos dados multimodais, combinando a genômica, a informação clínica e os dados imagiológicos com a ajuda da inteligência artificial e da aprendizagem automática (FENG et al., 2024; HE et al., 2023). Tal permitirá o desenvolvimento de biomarcadores específicos que combinam métricas genotípicas e fenotípicas, conduzindo a uma medicina de precisão mais eficaz. Além disso, serão essenciais melhores orientações éticas e tecnológicas para garantir a privacidade e a segurança dos dados genômicos a nível mundial, enquanto as ferramentas de análise em tempo real serão fundamentais para lidar com surtos infecciosos em rápida evolução (LEFTEROVA et al., 2015). A utilização da IA, como discutido por Topol (2019), pode acelerar a identificação de agentes patogênicos e a integração de dados multimodais para melhorar o diagnóstico e o tratamento em tempo real.

5.2. Tendências Futuras

O sequenciamento de nova geração (NGS) representa um avanço significativo na área da genômica, permitindo a análise de grandes volumes de dados genéticos de forma rápida e eficiente. Essa tecnologia, que surgiu em 2005, possibilita a leitura simultânea de milhões de bases nitrogenadas, desvendando os complexos mecanismos moleculares que subjazem à vida. As aplicações do NGS são vastas, abrangendo desde a medicina personalizada, com o desenvolvimento de tratamentos direcionados a mutações genéticas específicas, até a agricultura, com a criação de cultivos mais resistentes a pragas e doenças. (CARVALHO e SILVA, 2010)

A genômica e a proteômica revolucionaram a biologia molecular, alterando profundamente nossa compreensão da química dos ácidos nucleicos e das proteínas, e abrindo novas possibilidades para a manipulação de sistemas biológicos. O desenvolvimento de protocolos robustos para o uso de ferramentas moleculares impulsionou avanços em áreas como cultura de células, engenharia de tecidos e reprodução assistida. No Brasil, embora o Projeto Genoma Humano tenha tido menor impacto, os investimentos em genômica de microrganismos e parasitas, e mais recentemente em proteômica, têm sido cruciais. O conhecimento gerado por essas pesquisas têm o potencial de transformar a

prevenção e o tratamento de doenças. A diabetes tipo 1, por exemplo, tem sido alvo de estudos intensivos, e a descoberta do papel do gene Sumo-4 na regulação do sistema imunológico e a capacidade de manipular células-tronco para produzir insulina representam avanços promissores para o tratamento dessa doença (GARCIA, 2006).

O Sequenciamento de Nova Geração (NGS) oferece leituras mais longas e uma capacidade significativamente maior de gerar dados em comparação ao método tradicional de Sanger. Isso possibilita trabalhar com coberturas genômicas mais amplas, facilitando a montagem de genomas desconhecidos. Essa eficiência no sequenciamento é resultado da utilização de clonagem *in vitro* e sistemas de suporte sólido, que eliminam a necessidade de produzir clones bacterianos, montar placas de sequenciamento e separar fragmentos em géis. Essa técnica é aplicada, por exemplo, no desenvolvimento de marcadores moleculares, especialmente marcadores microssatélites de DNA, que ainda não existem para essa espécie (LEÃO, 2007).

Antes de 2000, as tecnologias de sequenciamento, como o método de Sanger, eram caras e demoradas, restringindo seu uso na patogenômica a estudos de pequena escala. Para enfrentar esses desafios, as técnicas de sequenciamento foram automatizadas, o que aprimorou a capacidade de processamento e diminuiu a margem de erro humano. Na década de 2000, com a introdução do Sequenciamento de Nova Geração (NGS), tornou-se possível realizar sequenciamentos em larga escala. No entanto, essa nova tecnologia ainda enfrentava problemas de capacidade e precisão, especialmente em relação a leituras curtas e erros em regiões genômicas complexas. O desenvolvimento de plataformas de NGS, como a Illumina, ajudou a reduzir custos e a facilitar estudos genômicos amplos, embora os desafios de montagem e precisão persistem. Entre 2010 e 2015, o aumento no número de projetos de sequenciamento resultou em grandes volumes de dados, sobrecarregando as ferramentas e a infraestrutura computacional disponíveis. Para resolver essa questão, surgiram pipelines de bioinformática, que incluíam ferramentas específicas para processar e analisar esses dados. No entanto, a padronização dos fluxos de trabalho e a integração de dados multimodais ainda eram limitadas. Entre 2015 e 2020, a integração de dados multimodais, que abrange informações

genômicas, clínicas e de imagem, continuou a ser um desafio. Apesar dos avanços no sequenciamento, a dependência de dados unimodais persistia. A solução surgiu com a introdução de ferramentas de inteligência artificial multimodal, que começaram a integrar esses diferentes tipos de dados. Desde 2020, o aumento do uso de dados patogenômicos levantou questões éticas, especialmente relacionadas à privacidade e consentimento, particularmente em situações de emergência de saúde pública. Para abordar essas preocupações, desenvolveu-se frameworks éticos e ferramentas de bioinformática que permitem análises em tempo real, com o objetivo de proteger a privacidade e promover o compartilhamento responsável de dados (LARSON et al., 2023; LYNCH et al., 2023; ARELLANO et al., 2018).

A combinação de dados genômicos, transcriptômicos e proteômicos é fundamental para uma compreensão mais completa das doenças. Pesquisas futuras devem se concentrar em desenvolver plataformas que integrem esses dados para fornecer insights mais profundos.

A fusão de dados multimodais tem como ideia extrair e integrar informações contextuais complementares de diferentes fontes, proporcionando uma base mais sólida para a tomada de decisões. Essa abordagem é especialmente importante na medicina, onde resultados semelhantes em uma única modalidade podem ser interpretados de maneiras distintas quando considerados em conjunto com outras modalidades (LEÃO, 2007).

6. Conclusão

A patogenômica, como abordado ao longo deste capítulo, representa um avanço essencial no estudo de doenças infecciosas, proporcionando uma compreensão profunda dos mecanismos genéticos que permitem aos patógenos causar infecções e resistir a tratamentos. A disponibilidade de grandes volumes de dados genômicos, combinada com o uso de ferramentas de bioinformática, tem revolucionado o diagnóstico e o tratamento de doenças, permitindo não apenas a identificação precisa de patógenos, mas também o desenvolvimento de estratégias terapêuticas personalizadas e vacinas mais eficazes.

O impacto desse campo para a ciência e a saúde pública é inegável, especialmente no monitoramento e controle de doenças emergentes e patógenos multirresistentes. Ao sequenciar e analisar genomas de patógenos, é possível identificar fatores de virulência e resistência, prever a evolução de surtos e orientar políticas de saúde pública, como ilustrado pela resposta à pandemia de COVID-19. Ademais, as metodologias de sequenciamento permitiram uma abordagem mais abrangente e precisa no diagnóstico molecular, melhorando significativamente as estratégias de detecção precoce.

Além disso, a patogenômica contribui para o desenvolvimento de novos fármacos e antibióticos, auxiliando na identificação de novos alvos terapêuticos e na compreensão das interações patógeno-hospedeiro, que são cruciais para a inovação no combate a patógenos de alta resistência. Em suma, este campo emergente continua a transformar o modo como a ciência compreende e responde às doenças infecciosas, sendo uma ferramenta indispensável na promoção de diagnósticos rápidos, tratamentos eficazes e controle eficiente de surtos.

Diante disso, é evidente a revolução que a patogenômica trouxe para o contexto de terapêutica e prevenção. E à medida que a compreensão da interação patógeno-hospedeiro avançou e a bioinformática foi inserida no desenvolvimento farmacêutico, também foi possível identificar seus mecanismos de resistência a medicamentos antimicrobianos, um dos maiores desafios da era moderna. Dessa forma, a aplicação da genômica na patologia e farmacologia torna a identificação de genes e proteínas essenciais para a sobrevivência destes microrganismos patogênicos uma arma em potencial, reduzindo bastante as chances destes seres contornarem efeitos terapêuticos e conferindo mais eficácia e segurança a nova geração de fármacos.

Neste capítulo, reunimos evidências de como o conhecimento sobre o genoma de patógenos tem revolucionado a medicina nas últimas décadas. O avanço tecnológico proporcionado pela computação e pela internet foi significativo, e a era do Big Data trouxe consigo um volume de dados complexo e desafiador para análise. Isso ressalta a necessidade não apenas de novas ferramentas analíticas, mas também de profissionais qualificados com uma visão

multidisciplinar e interdisciplinar, capazes de integrar diferentes campos do conhecimento.

Espera-se que a acessibilidade às ferramentas de sequenciamento e análise de genomas de patógenos, de forma mais rápida, permita o desenvolvimento de pesquisas e aplicações na profilaxia, tratamento e controle de patógenos na saúde pública global. Na profilaxia, será possível identificar genes candidatos para vacinas, por exemplo. No tratamento, será viável identificar genes candidatos como alvos de fármacos. Quanto ao controle, a vigilância genômica de patógenos permitirá o monitoramento de novas variantes circulantes, identificando aquelas com alta morbidade, o que pode levar a uma pandemia, além de possibilitar o acompanhamento da circulação de genes de resistência a antimicrobianos.

Além disso, espera-se que a patogenômica integre-se a outras “ômicas”, como a transcriptômica e a metabolômica, para melhorar o entendimento sobre a biologia dos patógenos e a interação patógeno-hospedeiro. Combinados esses esforços, espera-se que as 'ômicas' continuem revolucionando a biomedicina e que os profissionais atuantes nesta área estejam atentos aos desafios e preparados para enfrentá-los, promovendo o constante desenvolvimento do conhecimento científico e tecnológico.

7. Referências

- ABBAS, A. K. *Imunologia celular e molecular*. Elsevier, 2018.
- ABBY, S.; DAUBIN, V. Comparative genomics and the evolution of prokaryotes. *Trends Microbiol.**, v. 15, n. 3, p. 135–41
- AGARWAL, J.; SRIVASTAVA, S.; SINGH, M. Pathogenomics of uropathogenic *Escherichia coli*. *Indian J Med Microbiol.**, v. 30, n. 2, p. 141–149, abr. 2012.
- ALKATHERI, A. H. et al. Microbial Genomics: Innovative Targets and Mechanisms. *Antibiotics*, v. 12, n. 2, p. 190, 17 jan. 2023.
- ALSPAUGH, J. A. Virulence mechanisms and *Cryptococcus neoformans* pathogenesis. *Fungal Genet Biol.*, v. 78, p. 55–58, maio 2015.
- ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. *J Mol Biol.*, v. 215, n. 3, p. 403–410, out. 1990.

- ALTURKI, N. A.; MASHRAQI, M. M.; JALAL, K.; KHAN, K.; BASHARAT, Z.; ALZAMAMI, A. Therapeutic Target Identification and Inhibitor Screening against Riboflavin Synthase of Colorectal Cancer Associated **Fusobacterium nucleatum**. **Cancers**, v. 14, n. 24, p. 6260, 19 dez. 2022.
- ARBER, W. Genetic variation: molecular mechanisms and impact on microbial evolution. **FEMS Microbiol Rev.**, v. 24, n. 1, p. 1–7, jan. 2000.
- ARELLANO, A. M.; DAI, W.; WANG, S.; JIANG, X.; OHNO-MACHADO, L. Privacy Policy and Technology in Biomedical Data Science. *Annu Rev Biomed Data Sci.*, v. 1, n. 1, p. 115–129, 20 jul. 2018.
- ARENAS, M.; ARAUJO, N. M.; BRANCO, C.; CASTELHANO, N.; CASTRO-NALLAR, E.; PÉREZ-LOSADA, M. Mutation and recombination in pathogen evolution: Relevance, methods and controversies. **Infect Genet Evol.**, v. 63, p. 295–306, set. 2018.
- ARMSTRONG, G. L. et al. Pathogen Genomics in Public Health. **N Engl J Med.**, v. 381, n. 26, p. 2569–2580, 26 dez. 2019.
- ASHOUR, J. et al. Mouse STAT2 Restricts Early Dengue Virus Replication. **Cell Host Microbe**, v. 8, n. 5, p. 410–421, nov. 2010.
- ATHANASOPOULOU, K.; BOTI, M. A.; ADAMOPOULOS, P. G.; SKOUROU, P. C.; SCORILAS, A. Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. **Life**, v. 12, n. 1, p. 30, 26 dez. 2021.
- BAÑÜLS, A. L.; SANOU, A.; VAN ANH, N. T.; GODREUIL, S. Mycobacterium tuberculosis: ecology and evolution of a human bacterium. *J Med Microbiol.*, v. 64, n. 11, p. 1261–1269, 1 nov. 2015.
- BARRETT, A. D. T. et al. The rapid progress in COVID vaccine development and implementation. **Npj Vaccines**, v. 7, n. 1, p. 20, s41541-022-00442–8, 10 fev. 2022.
- BOCK, J. R.; GOUGH, D. A. Predicting protein–protein interactions from primary structure. **Bioinformatics**, v. 17, n. 5, p. 455–460, 1 maio 2001.
- BOTTAI, D. et al. TbD1 deletion as a driver of the evolutionary success of modern epidemic **Mycobacterium tuberculosis** lineages. **Nat Commun.**, v. 11, n. 1, p. 684, 4 fev. 2020.
- BRAY, N.; DUBCHAK, I.; PACHTER, L. AVID: A Global Alignment Program. *Genome Res.*, v. 13, n. 1, p. 97–102, 1 jan. 2003.
- CACHO, A.; SMIRNOVA, E.; HUZURBAZAR, S.; CUI, X. A Comparison of Base-calling Algorithms for Illumina Sequencing Technology. **Brief Bioinform.**, v. 17, n. 5, p. 786–795, set. 2016.
- CADENA, J.; THOMPSON, G. R.; PATTERSON, T. F. Aspergillosis. **Infect Dis*

- Clin North Am.*, v. 35, n. 2, p. 415–434, jun. 2021.
- CÁRDENAS, P.; CORREDOR, V.; SANTOS-VEGA, M. Genomic epidemiological models describe pathogen evolution across fitness valleys. *Sci Adv.*, v. 8, n. 28, p. eabo0173, 15 jul. 2022.
- CARVALHO, M. C. D. C. G. D.; SILVA, D. C. G. D. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. Ciênc Rural, v. 40, n. 3, p. 735–744, mar. 2010.
- CDC. Outbreak Detection [Internet]. CDC PulseNet, 2024 [citado em 4 out. 2024]. Disponível em: <https://www.cdc.gov/pulsenet/hcp/about/outbreak-detection.html>
- CHAHAL, G.; TYAGI, S.; RAMIALISON, M. Navigating the non-coding genome in heart development and Congenital Heart Disease. *Differentiation*, v. 107, p. 11–23, maio 2019.
- CHANDRA, P.; GRIGSBY, S. J.; PHILIPS, J. A. Immune evasion and provocation by *Mycobacterium tuberculosis*. *Nat Rev Microbiol.*, v. 20, n. 12, p. 750–766, dez. 2022.
- CHARRON, P.; KANG, M. VariantDetective: an accurate all-in-one pipeline for detecting consensus bacterial SNPs and SVs. ALKAN, C. (Ed.). *Bioinformatics*, v. 40, n. 2, p. btae066, 1 fev. 2024.
- CHATTERJEE, N.; WALKER, G. C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen.*, v. 58, n. 5, p. 235–263, jun. 2017.
- CHAUSSABEL, D. et al. Unique gene expression profiles of human macrophages and dendritic cells to phylogenetically distinct parasites. Blood. 2003 Jul 15;102(2):672–81.
- CHEN, J. Q.; WU, Y.; YANG, H.; BERGELSON, J.; KREITMAN, M.; TIAN, D. Variation in the Ratio of Nucleotide Substitution and Indel Rates across Genomes in Mammals and Bacteria. *Mol Biol Evol.*, v. 26, n. 7, p. 1523–1531, 1 jul. 2009.
- CHEN, L. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.*, v. 33, n. Database issue, p. D325–D328, 17 dez. 2004.
- CHIVIAN, D. et al. Metagenome-assembled genome extraction and analysis from microbiomes using KBase. *Nat Protoc.*, v. 18, n. 1, p. 208–238, jan. 2023.
- CHOI, S. R.; LEE, M. Transformer Architecture and Attention Mechanisms in Genome Data Analysis: A Comprehensive Review. *Biology*, v. 12, n. 7, p. 1033, 22 jul. 2023.
- CHOWDHURY, B.; GARAI, G. A review on multiple sequence alignment from the perspective of genetic algorithm. Genomics, v. 109, n. 5–6, p. 419–431, out. 2017.

- CLARK, J. R.; MARESSO, A. M. Comparative Pathogenomics of *Escherichia coli*: Polyvalent Vaccine Target Identification through Virulome Analysis. MONACK, D. (Ed.). *Infect Immun.*, v. 89, n. 8, p. e00115-21, 15 jul. 2021.
- COLE, S. T. et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, v. 393, n. 6685, p. 537–544, jun. 1998.
- COMAS, I. et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat Genet.*, v. 45, n. 10, p. 1176–1182, out. 2013.
- COSTA, F. F. Big data in biomedicine. *Drug Discov Today*, v. 19, n. 4, p. 433–440, abr. 2014.
- COVID-19 Host Genetics Initiative, COVID-19 Host Genetics Initiative Leadership, Niemi, M. E. K., Karjalainen, J., Liao, R. G., Neale, B. M. et al. Mapping the human genetic architecture of COVID-19. *Nature*, v. 600, n. 7889, p. 472–477, 16 dez. 2021.
- D'MELLO, A.; AHEARN, C. P.; MURPHY, T. F.; TETTELIN, H. ReVac: a reverse vaccinology computational pipeline for prioritization of prokaryotic protein vaccine candidates. *BMC Genomics*, v. 20, n. 1, p. 981, dez. 2019.
- DARLING, A. E.; MAU, B.; PERNA, N. T. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. STAJICH, J. E. (Ed.). *PLoS ONE*, v. 5, n. 6, p. e11147, 25 jun. 2010.
- DAVID, J.; BELL, R. E.; CLARK, G. C. Mechanisms of Disease: Host-Pathogen Interactions between Burkholderia Species and Lung Epithelial Cells. *Front Cell Infect Microbiol* [Internet], v. 5, 18 nov. 2015 [citado em 16 out. 2024]. Disponível em: <http://journal.frontiersin.org/Article/10.3389/fcimb.2015.00080/abstract>
- DAVIS, F. P.; BARKAN, D. T.; ESWAR, N.; MCKERROW, J. H.; SALI, A. Host–pathogen protein interactions predicted by comparative modeling. *Protein Sci.*, v. 16, n. 12, p. 2585–2596, dez. 2007.
- DE NOOIJER, A. H.; PICKKERS, P.; NETEA, M. G.; KOX, M. Inflammatory biomarkers to predict the prognosis of acute bacterial and viral infections. *J Crit Care*, v. 78, p. 154360, dez. 2023.
- DEMUTH, A. et al. Pathogenomics: An updated European Research Agenda. *Infect Genet Evol.*, v. 8, n. 3, p. 386–393, maio 2008.
- DI PAOLA, N.; SANCHEZ-LOCKHART, M.; ZENG, X.; KUHN, J. H.; PALACIOS, G. Viral genomics in Ebola virus research. *Nat Rev Microbiol.*, v. 18, n. 7, p. 365–378, jul. 2020.
- DJORDJEVIC, S. P. et al. Genomic surveillance for antimicrobial resistance — a One Health perspective. *Nat Rev Genet.*, v. 25, n. 2, p. 142–157, fev.

2024.

- DONG, M.; MASUYER, G.; STENMARK, P. Botulinum and Tetanus Neurotoxins. **Annu Rev Biochem.**, v. 88, n. 1, p. 811–837, 20 jun. 2019.
- DOOLITTLE, J. M.; GOMEZ, S. M. Structural similarity-based predictions of protein interactions between HIV-1 and **Homo sapiens**. **Virol J.**, v. 7, n. 1, p. 82, dez. 2010.
- DOWDLE, W. R. The principles of disease elimination and eradication. *Bull World Health Organ.*, v. 76, n. Suppl 2, p. 22–25, 1998.
- DYER, M. D.; MURALI, T. M.; SOBRAL, B. W. Computational prediction of host-pathogen protein–protein interactions. **Bioinformatics**, v. 23, n. 13, p. i159–i166, 1 jul. 2007.
- EDMONDSON, D. G.; HU, B.; NORRIS, S. J. Long-Term In Vitro Culture of the Syphilis Spirochete **Treponema pallidum** subsp. **pallidum**. ROSA, P. A. (Ed.). **mBio**, v. 9, n. 3, p. e01153-18, 5 jul. 2018.
- EKTEFAIE, Y. et al. Globally diverse **Mycobacterium tuberculosis** resistance acquisition: a retrospective geographical and temporal analysis of whole genome sequences. **Lancet Microbe**, v. 2, n. 3, p. e96–e104, mar. 2021.
- ELBEHIRY, A. et al. How MALDI-TOF Mass Spectrometry Technology Contributes to Microbial Infection Control in Healthcare Settings. **Vaccines**, v. 10, n. 11, p. 1881, 8 nov. 2022.
- ELLERMANN, M.; ARTHUR, J. C. Siderophore-mediated iron acquisition and modulation of host-bacterial interactions. **Free Radic Biol Med.**, v. 105, p. 68–78, abr. 2017.
- ENGELTHALER, D. M. Genomic surveillance and pathogen intelligence. **Front Sci.**, v. 2, p. 1397048, 25 abr. 2024.
- ESPINOSA, E.; BAUTISTA, R.; LARROSA, R.; PLATA, O. Advancements in long-read genome sequencing technologies and algorithms. **Genomics**, v. 116, n. 3, p. 110842, maio 2024.
- EVANS, P.; DAMPIER, W.; UNGAR, L.; TOZEREN, A. Prediction of HIV-1 virus-host protein interactions using virus and host sequence motifs. **BMC Med Genomics.**, v. 2, n. 1, p. 27, dez. 2009.
- EZZELL, J. W.; WELKOS, S. L. The capsule of **Bacillus anthracis**, a review. **J Appl Microbiol.**, v. 87, n. 2, p. 250–250, ago. 1999.
- FENG, X. et al. Pathogenomics for accurate diagnosis, treatment, prognosis of oncology: a cutting edge overview. **J Transl Med.**, v. 22, n. 1, p. 131, 3 fev. 2024.
- FIGUEROA, J. et al. Analysis of single nucleotide polymorphisms (SNPs) associated with antibiotic resistance genes in Chilean *Piscirickettsia*

- salmonis strains. J Fish Dis., v. 42, n. 12, p. 1645–1655, dez. 2019.
- FLEISCHMANN, R. D. et al. Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd. *Science*, v. 269, n. 5223, p. 496–512, 28 jul. 1995.
- FORSBERG, K. et al. *Candida auris*: The recent emergence of a multidrug-resistant fungal pathogen. *Med Mycol.*, v. 57, n. 1, p. 1–12, 1 jan. 2019.
- GABRIELAITE, M.; NIELSEN, F. C.; JOHANSEN, H. K.; MARVIG, R. L. *Achromobacter* spp. genetic adaptation in cystic fibrosis. *Microb Genomics* [Internet], v. 7, n. 7, 28 jul. 2021 [citado em 16 out. 2024]. Disponível em: <https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000582>
- GARCIA, E. S. *Genes: fatos e fantasias* [Internet]. Editora FIOCRUZ, 2006 [citado em 16 out. 2024]. Disponível em: <http://books.scielo.org/id/m7tn7>
- GARCIA-DEL PORTILLO, F.; FINLAY, B. B. *Salmonella* invasion of nonphagocytic cells induces formation of macropinosomes in the host cell. *Infect Immun.*, v. 62, n. 10, p. 4641–4645, out. 1994.
- GARDNER, M. J. et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, v. 419, n. 6906, p. 498–511, out. 2002.
- GARRIDO-CARDENAS, J. A.; GONZÁLEZ-CERÓN, L.; MANZANO-AGUGLIARO, F.; MESA-VALLE, C. *Plasmodium* genomics: an approach for learning about and ending human malaria. *Parasitol Res.*, v. 118, n. 1, p. 1–27, jan. 2019.
- GAUTHIER, J.; VINCENT, A. T.; CHARETTE, S. J.; DEROME, N. A brief history of bioinformatics. *Brief Bioinform.*, v. 20, n. 6, p. 1981–1996, 27 nov. 2019.
- GILLETTE, D. D. et al. Analysis of Human Bronchial Epithelial Cell Proinflammatory Response to *Burkholderia cenocepacia* Infection. *J Biol Chem.*, v. 288, n. 6, p. 3691–3695, fev. 2013.
- GOFFEAU, A. et al. Life with 6000 Genes. *Science*, v. 274, n. 5287, p. 546–567, 25 out. 1996.
- GOLDSMITH-FISCHMAN, S.; HONIG, B. Structural genomics: Computational methods for structure analysis. *Protein Sci.*, v. 12, n. 9, p. 1813–1821, set. 2003.
- GOMES, L. G. R. et al. In Silico Designed Multi-Epitope Immunogen “Tpme-VAC/LGCM-2022” May Induce Both Cellular and Humoral Immunity against *Treponema pallidum* Infection. *Vaccines*, v. 10, n. 7, p. 1019, 25 jun. 2022.
- GRANT, A. et al. Zika Virus Targets Human STAT2 to Inhibit Type I Interferon

- Signaling. **Cell Host Microbe**, v. 19, n. 6, p. 882–890, jun. 2016.
- HAY, A. J.; MCCAULEY, J. W. The WHO global influenza surveillance and response system (GISRS)—A future perspective. **Influenza Other Respir Viruses**, v. 12, n. 5, p. 551–557, set. 2018.
- HAYASHI, S.; OGURA, Y. ERK signaling dynamics in the morphogenesis and homeostasis of *Drosophila*. *Curr Opin Genet Dev.*, v. 63, p. 9–15, ago. 2020.
- HE, Q. et al. A graph-based genome and pan-genome variation of the model plant *Setaria*. *Nat Genet.*, v. 55, n. 7, p. 1232–1242, jul. 2023.
- HESS, J. F. et al. Library preparation for next generation sequencing: A review of automation strategies. **Biotechnol Adv.**, v. 41, p. 107537, jul. 2020.
- HU, T.; CHITNIS, N.; MONOS, D.; DINH, A. Next-generation sequencing technologies: An overview. *Hum Immunol.*, v. 82, n. 11, p. 801–811, nov. 2021.
- HYUN, J. C.; MONK, J. M.; SZUBIN, R.; HEFNER, Y.; PALSSON, B. O. Global pathogenomic analysis identifies known and candidate genetic antimicrobial resistance determinants in twelve species. **Nat Commun.**, v. 14, n. 1, p. 7690, 24 nov. 2023.
- IBARZ PAVÓN, A. B.; MAIDEN, M. C. J. Multilocus Sequence Typing. In: CAUGANT, D. A. (Ed.). **Molecular Epidemiology of Microorganisms** [Internet]. Totowa, NJ: Humana Press, 2009 [citado em 16 out. 2024]. p. 129–140. (Methods in Molecular Biology; v. 551). Disponível em: http://link.springer.com/10.1007/978-1-60327-999-4_11
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. Finishing the euchromatic sequence of the human genome. **Nature**, v. 431, n. 7011, p. 931–945, out. 2004.
- IOC J. Genômica funcional: desafios [Internet]. IOC, 2010 [citado em 4 out. 2024]. Disponível em: <https://www.ioc.fiocruz.br/noticias/genomica-funcional-desafios>
- IRAOLA, G.; BOYCE, J. D. Understanding Pathogenesis Through Pathogenomics and Bioinformatics. In: PRESCOTT, J. F.; RYCROFT, A. N.; BOYCE, J. D.; MACINNES, J. I.; VAN IMMERSEEL, F.; VÁZQUEZ-BOLAND, J. A. (Eds.). **Pathogenesis of Bacterial Infections in Animals** [Internet]. 1. ed. Wiley, 2022 [citado em 16 out. 2024]. p. 32–56. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/9781119754862.ch3>
- JAIN, C.; RODRIGUEZ-R, L. M.; PHILLIPPY, A. M.; KONSTANTINIDIS, K. T.; ALURU, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. **Nat Commun.**, v. 9, n. 1, p. 5114, 30 nov. 2018.
- JUHAS, M.; EBERL, L.; CHURCH, G. M. Essential genes as antimicrobial

- targets and cornerstones of synthetic biology. **Trends Biotechnol.**, v. 30, n. 11, p. 601–607, nov. 2012.
- KATOH, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, v. 30, n. 14, p. 3059–3066, 15 jul. 2002.
- KIRKCALDY, R. D. et al. **Trichomonas vaginalis** Antimicrobial Drug Resistance in 6 US Cities, STD Surveillance Network, 2009–2010. **Emerg Infect Dis.**, v. 18, n. 6, p. 939–943, jun. 2012.
- KLINE, K. A.; FÄLKER, S.; DAHLBERG, S.; NORMARK, S.; HENRIQUES-NORMARK, B. Bacterial Adhesins in Host-Microbe Interactions. **Cell Host Microbe**, v. 5, n. 6, p. 580–592, jun. 2009.
- KOLACZKOWSKA, E.; KUBES, P. Neutrophil recruitment and function in health and inflammation. *Nat Rev Immunol.*, v. 13, n. 3, p. 159–175, mar. 2013.
- KOMISSAROV, A. B. et al. Genomic epidemiology of the early stages of SARS-CoV-2 outbreak in Russia [Internet]. 2020 [citado em 16 out. 2024]. Disponível em: <http://medrxiv.org/lookup/doi/10.1101/2020.07.14.20150979>
- KUMAR JAISWAL, A.; TIWARI, S.; JAMAL, S.; BARH, D.; AZEVEDO, V.; SOARES, S. An In Silico Identification of Common Putative Vaccine Candidates against **Treponema pallidum**: A Reverse Vaccinology and Subtractive Genomics Based Approach. **Int J Mol Sci.**, v. 18, n. 2, p. 402, 14 fev. 2017.
- LAGE, F. S. D.; SANTOS, F. B. Sequenciamento de primeira geração: método de Sanger. In: MARIANO, D.; DEZORDI, F. Z.; MARTINS, P.; XAVIER, J.; SOUSA, T. D. J.; LIMA, L. et al. (Eds.). **BIOINFO - Revista Brasileira de Bioinformática e Biologia Computacional** [Internet]. 1. ed. Alfahelix, 2021 [citado em 16 out. 2024]. Disponível em: <https://bioinfo.com.br/sequenciamento-de-primeira-geracao-metodo-de-sanger/>
- LARSON, N. B.; OBERG, A. L.; ADJEI, A. A.; WANG, L. A Clinician’s Guide to Bioinformatics for Next-Generation Sequencing. **J Thorac Oncol.**, v. 18, n. 2, p. 143–157, fev. 2023.
- LATA, K. S. et al. Whole genome sequencing and de novo assembly of three virulent Indian isolates of **Leptospira**. **Infect Genet Evol.**, v. 85, p. 104579, nov. 2020.
- LEÃO, U. S. Sequenciamento de Nova Geração (NGS) de *Urochloa mosambicensis*: uma forrageira promissora para o semiárido. *RENORBIO 2017.*, 2017 [citado em 4 out. 2024]. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/handle/doc/1080722>
- LEE, S. A. et al. Ortholog-based protein-protein interaction prediction and its application to inter-species interactions. **BMC Bioinformatics**, v. 9, n. S12,

p. S11, dez. 2008.

- LEFEBVRE, C. et al. Discovery of a novel dehydratase of the fatty acid synthase type II critical for ketomycolic acid biosynthesis and virulence of *Mycobacterium tuberculosis*. **Sci Rep.**, v. 10, n. 1, p. 2112, 7 fev. 2020.
- LEFTEROVA, M. I.; SUAREZ, C. J.; BANAEI, N.; PINSKY, B. A. Next-Generation Sequencing for Infectious Disease Diagnosis and Management. *J Mol Diagn.*, v. 17, n. 6, p. 623–634, nov. 2015.
- LEITÃO, J. H. Microbial Virulence Factors. *Int J Mol Sci.*, v. 21, n. 15, p. 5320, 27 jul. 2020.
- LEVA-BUENO, J.; PEYMAN, S. A.; MILLNER, P. A. A review on impedimetric immunosensors for pathogen and biomarker detection. **Med Microbiol Immunol (Berl)**, v. 209, n. 3, p. 343–362, jun. 2020.
- LI, B. sheng et al. Genomic Evolution and Variation of SARS-CoV-2 in the Early Phase of COVID-19 Pandemic in Guangdong Province, China. *Curr Med Sci.*, v. 41, n. 2, p. 228–235, abr. 2021.
- LI, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. **Bioinformatics**, v. 27, n. 21, p. 2987–2993, 1 nov. 2011.
- LI, T. et al. CRISPR/Cas9 therapeutics: progress and prospects. *Signal Transduct Target Ther.*, v. 8, n. 1, p. 36, 16 jan. 2023.
- LI, W. et al. Plant pan-genomics: recent advances, new challenges, and roads ahead. **J Genet Genomics.**, v. 49, n. 9, p. 833–846, set. 2022.
- LIANG, Y.; LUO, H.; LIN, Y.; GAO, F. Recent advances in the characterization of essential genes and development of a database of essential genes. **iMeta**, v. 3, n. 1, p. e157, fev. 2024.
- LINS, R. A.; ASSEF NETO, R. C. D. S.; SILVA, C. H. D.; GUIMARÃES, M. C. S. Coronavírus: um panorama sobre a colaboração científica internacional do Brasil ao longo da história. **Rev Eletrônica Comun Informação E Inov Em Saúde** [Internet], v. 15, n. 4, 10 nov. 2021 [citado em 16 out. 2024]. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/2398>
- LIU, S.; MOON, C. D.; ZHENG, N.; HUWS, S.; ZHAO, S.; WANG, J. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. **Microbiome**, v. 10, n. 1, p. 76, 12 maio 2022.
- LU, R.; POPOV, V.; PATEL, J.; EAVES-PYLES, T. *Burkholderia mallei* and *Burkholderia pseudomallei* stimulate differential inflammatory responses from human alveolar type II cells (ATII) and macrophages. *Front Cell Infect Microbiol* [Internet], v. 2, 2012 [citado em 16 out. 2024]. Disponível em: <http://journal.frontiersin.org/article/10.3389/fcimb.2012.00165/abstract>

- LYNCH, F. et al. Australian public perspectives on genomic data storage and sharing: Benefits, concerns and access preferences. *Eur J Med Genet.*, v. 66, n. 1, p. 104676, jan. 2023.
- LYNCH, T. et al. A Primer on Infectious Disease Bacterial Genomics. *Clin Microbiol Rev.*, v. 29, n. 4, p. 881–913, out. 2016.
- MAIDEN, M. C. J. et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. **Proc Natl Acad Sci.**, v. 95, n. 6, p. 3140–3145, 17 mar. 1998.
- MANIRUZZAMAN, M.; NOKHODCHI, A. Continuous manufacturing via hot-melt extrusion and scale up: regulatory matters. *Drug Discov Today*, v. 22, n. 2, p. 340–351, fev. 2017.
- Mapping and Sequencing the Human Genome [Internet]. Washington, D.C.: National Academies Press, 1988 [citado em 16 out. 2024]. Disponível em: <http://www.nap.edu/catalog/1097>
- MATHAIYAN, J.; CHANDRASEKARAN, A.; DAVIS, S. Ethics of genomic research. *Perspect Clin Res.*, v. 4, n. 1, p. 100, 2013.
- MATTHEWS, L. R. et al. Identification of Potential Interaction Networks Using Sequence-Based Searches for Conserved Protein-Protein Interactions or “Interologs.” **Genome Res.**, v. 11, n. 12, p. 2120–2126, 1 dez. 2001.
- MEIER-KOLTHOFF, J. P.; GÖKER, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. **Nat Commun.**, v. 10, n. 1, p. 2182, 16 maio 2019.
- MILLER, J. R.; KOREN, S.; SUTTON, G. Assembly algorithms for next-generation sequencing data. *Genomics*, v. 95, n. 6, p. 315–327, jun. 2010.
- MOORE, K. L. et al. P-selectin glycoprotein ligand-1 mediates rolling of human neutrophils on P-selectin. *J Cell Biol.*, v. 128, n. 4, p. 661–671, 15 fev. 1995.
- MOSHIRI, N. et al. The ViReflow pipeline enables user friendly large scale viral consensus genome reconstruction. **Sci Rep.**, v. 12, n. 1, p. 5077, 24 mar. 2022.
- MULLAN, L. Pairwise sequence alignment—it’s all about us! *Brief Bioinform.*, v. 7, n. 1, p. 113–115, mar. 2006.
- NAFEA, A. M. et al. Application of next-generation sequencing to identify different pathogens. **Front Microbiol.**, v. 14, p. 1329330, 29 jan. 2024.
- NAIDENOV, B. et al. Pan-Genomic and Polymorphic Driven Prediction of Antibiotic Resistance in **Elizabethkingia**. **Front Microbiol.**, v. 10, p. 1446, 4 jul. 2019.
- NATIONAL RESEARCH COUNCIL (U.S.). **New science of metagenomics:*

- revealing the secrets of our microbial planet*. Washington, DC: National Academies Press, 2007. 1 p.
- NEEDLEMAN, S. B.; WUNSCH, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol.*, v. 48, n. 3, p. 443–453, mar. 1970.
- NEHRA, M.; KUMAR, V.; KUMAR, R.; DILBAGHI, N.; KUMAR, S. Current Scenario of Pathogen Detection Techniques in Agro-Food Sector. **Biosensors**, v. 12, n. 7, p. 489, 4 jul. 2022.
- NIH. A Brief Guide to Genomics [Internet]. genome.gov. 2022 [citado em 4 out. 2024]. Disponível em: <https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics#:~:text=Genomics%20is%20the%20study%20of,and%20with%20the%20person's%20environment>
- ODEH, M. et al. iOntoBioethics: A Framework for the Agile Development of Bioethics Ontologies in Pandemics, Applied to COVID-19. *Front Med.*, v. 8, p. 619978, 21 maio 2021.
- OLI, A. N. et al. Immunoinformatics and Vaccine Development: An Overview. **ImmunoTargets Ther.**, v. 9, p. 13–30, fev. 2020.
- PAIS, F. S. M.; RUY, P. D. C.; OLIVEIRA, G.; COIMBRA, R. S. Assessing the efficiency of multiple sequence alignment programs. *Algorithms Mol Biol.*, v. 9, n. 1, p. 4, dez. 2014.
- PAKBIN, B.; BRÜCK, W. M.; ROSSEN, J. W. A. Virulence Factors of Enteric Pathogenic *Escherichia coli*: A Review. **Int J Mol Sci.**, v. 22, n. 18, p. 9922, 14 set. 2021.
- PALLEN, M. J.; WREN, B. W. Bacterial pathogenomics. **Nature**, v. 449, n. 7164, p. 835–842, out. 2007.
- PAPA MZE, N. et al. High Throughput SARS-CoV-2 Genome Sequencing from 384 Respiratory Samples Using the Illumina COVIDSeq Protocol. **Genes**, v. 14, n. 3, p. 681, 9 mar. 2023.
- PAREEK, C. S.; SMOCZYNSKI, R.; TRETYN, A. Sequencing technologies and genome sequencing. **J Appl Genet.**, v. 52, n. 4, p. 413–435, nov. 2011.
- PATON, J. C.; TRAPPETTI, C. **Streptococcus pneumoniae** Capsular Polysaccharide. FISCHETTI, V. A.; NOVICK, R. P.; FERRETTI, J. J.; PORTNOY, D. A.; BRAUNSTEIN, M.; ROOD, J. I. (Eds.). **Microbiol Spectr.**, v. 7, n. 2, p. 7.2.33, 12 abr. 2019.
- PEREIRA, R.; OLIVEIRA, J.; SOUSA, M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J Clin Med.*, v. 9, n. 1, p. 132, 3 jan. 2020.
- PIDDOCK, L. J. V. The global antibiotic research and development partnership

- (GARDP): researching and developing new antibiotics to meet global public health needs. **MedChemComm.**, v. 10, n. 8, p. 1227–1230, 2019.
- POLLARD, M. O.; GURDASANI, D.; MENTZER, A. J.; PORTER, T.; SANDHU, M. S. Long reads: their purpose and place. **Hum Mol Genet.**, v. 27, n. R2, p. R234–R241, 1 ago. 2018.
- POPLIN, R. et al. A universal SNP and small-indel variant caller using deep neural networks. **Nat Biotechnol.**, v. 36, n. 10, p. 983–987, nov. 2018.
- PRADEL, G.; TEMPLETON, T. J. Genomics of Pathogenic Parasites. In: HACKER, J.; DOBRINDT, U. (Eds.). **Pathogenomics** [Internet]. 1. ed. Wiley, 2006 [citado em 16 out. 2024]. p. 417–444. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/352760801X.ch19>
- PRASAD, A.; ENE, A.; JABLONSKA, S.; DU, J.; WOLFE, A. J.; PUTONTI, C. Comparative Genomic Study of *Streptococcus anginosus* Reveals Distinct Group of Urinary Strains. RAO, K. (Ed.). *mSphere*, v. 8, n. 2, 20 abr. 2023.
- RAHMAN, M. d. M.; MASUM, M. d. H. U.; WAJED, S.; TALUKDER, A. A comprehensive review on COVID-19 vaccines: development, effectiveness, adverse effects, distribution and challenges. *VirusDisease*, v. 33, n. 1, p. 1–22, mar. 2022.
- RODRIGUEZ, R.; KRISHNAN, Y. The chemistry of next-generation sequencing. **Nat Biotechnol.**, v. 41, n. 12, p. 1709–1715, dez. 2023.
- ROSCONI, F. et al. A bacterial pan-genome makes gene essentiality strain-dependent and evolvable. **Nat Microbiol.**, v. 7, n. 10, p. 1580–1592, 12 set. 2022.
- ROULI, L.; MERHEJ, V.; FOURNIER, P. E.; RAOULT, D. The bacterial pangenome as a new tool for analysing pathogenic bacteria. **New Microbes New Infect.**, v. 7, p. 72–85, set. 2015.
- ROUSH, S. W. Historical Comparisons of Morbidity and Mortality for Vaccine-Preventable Diseases in the United States. *JAMA*, v. 298, n. 18, p. 2155, 14 nov. 2007.
- SADLER, A. J.; WILLIAMS, B. R. G. Interferon-inducible antiviral effectors. *Nat Rev Immunol.*, v. 8, n. 7, p. 559–568, jul. 2008.
- SAMEEK, R.; CHINNAIYAN, A. M. Translating Genomics for Precision Cancer Medicine. *Annu Rev Genomics Hum Genet.*, v. 15, n. 1, p. 395–415, 31 ago. 2014.
- SAMUEL, C. E. Antiviral Actions of Interferons. *Clin Microbiol Rev.*, v. 14, n. 4, p. 778–809, out. 2001.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proc Natl Acad Sci.**, v. 74, n. 12, p. 5463–5467, dez. 1977.

- SATAM, H. et al. Next-Generation Sequencing Technology: Current Trends and Advancements. **Biology**, v. 12, n. 7, p. 997, 13 jul. 2023.
- SAYERS, E. W. et al. Database resources of the National Center for Biotechnology Information in 2023. **Nucleic Acids Res.**, v. 51, n. D1, p. D29–D38, 6 jan. 2023.
- SEAH, Y. M. et al. In Silico Evaluation of Variant Calling Methods for Bacterial Whole-Genome Sequencing Assays. DEKKER, J. P. (Ed.). **J Clin Microbiol.**, 10 jul. 2023.
- SEGERMAN, B. The Most Frequently Used Sequencing Technologies and Assembly Methods in Different Time Segments of the Bacterial Surveillance and RefSeq Genome Databases. **Front Cell Infect Microbiol.**, v. 10, p. 527102, 19 out. 2020.
- SELA, I.; WOLF, Y. I.; KOONIN, E. V. Genome plasticity, a key factor of evolution in prokaryotes [Internet]. 2018 [citado em 16 out. 2024]. Disponível em: <http://biorxiv.org/lookup/doi/10.1101/357400>
- SEN, R.; NAYAK, L.; DE, R. K. A review on host–pathogen interactions: classification and prediction. *Eur J Clin Microbiol Infect Dis.*, v. 35, n. 10, p. 1581–1599, out. 2016.
- SEREIKA, M. et al. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nat Methods*, v. 19, n. 7, p. 823–826, jul. 2022.
- SHARMA, A.; GARG, A.; RAMANA, J.; GUPTA, D. VirulentPred 2.0: An improved method for prediction of virulent proteins in bacterial pathogens. **Protein Sci.**, v. 32, n. 12, p. e4808, dez. 2023.
- SHARMA, N. C.; EFSTRATIOU, A.; MOKROUSOV, I.; MUTREJA, A.; DAS, B.; RAMAMURTHY, T. Diphtheria. **Nat Rev Dis Primer**, v. 5, n. 1, p. 81, 5 dez. 2019.
- SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.*, v. 7, n. 1, p. 539, jan. 2011.
- SILVA, H. P. D.; LIMA, L. D. D. Política, economia e saúde: lições da COVID-19. **Cad Saúde Pública**, v. 37, n. 9, p. e00200221, 2021.
- SILVA, R. C. C. D.; ALVES, M. C. S. O uso de ferramentas de bioinformática para análise de dados genéticos: uma revisão. *Sci Electron Arch [Internet]*, v. 17, n. 1, 11 dez. 2023 [citado em 16 out. 2024]. Disponível em: <https://sea.ufr.edu.br/SEA/article/view/1872>
- SIM, S. H. et al. Innate Immune Responses of Pulmonary Epithelial Cells to Burkholderia pseudomallei Infection. NG, L. F. P. (Ed.). *PLoS ONE*, v. 4, n. 10, p. e7308, 6 out. 2009.

- SMITH, T. F.; WATERMAN, M. S. Identification of common molecular subsequences. *J Mol Biol.*, v. 147, n. 1, p. 195–197, mar. 1981.
- SOUZA, A. L. F. D.; BRUSAMARELLO, L. C. C. Sequenciamento de DNA: decifrando o manual de instruções dos seres vivos. **Genética Na Esc.**, v. 4, n. 1, p. 45–52, 20 maio 2009.
- STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nat Rev Genet.**, v. 20, n. 11, p. 631–65
- STRUELENS, M. J. et al. Real-time genomic surveillance for enhanced control of infectious diseases and antimicrobial resistance. **Front Sci.**, v. 2, p. 1298248, 25 abr. 2024.
- SU, C. M.; DU, Y.; ROWLAND, R. R. R.; WANG, Q.; YOO, D. Reprogramming viral immune evasion for a rational design of next-generation vaccines for RNA viruses. *Front Immunol.*, v. 14, p. 1172000, 17 abr. 2023.
- SUGIMAN-MARANGOS, S. N.; GILL, S. K.; MANSFIELD, M. J.; ORRELL, K. E.; DOXEY, A. C.; MELNYK, R. A. Structures of distant diphtheria toxin homologs reveal functional determinants of an evolutionarily conserved toxin scaffold. **Commun Biol.**, v. 5, n. 1, p. 375, 19 abr. 2022.
- TALBIRD, S. E. et al. Impact of Routine Childhood Immunization in Reducing Vaccine-Preventable Diseases in the United States. *Pediatrics*, v. 150, n. 3, p. e2021056013, 1 set. 2022.
- TAM, K.; TORRES, V. J. **Staphylococcus aureus* Secreted Toxins and Extracellular Enzymes*. FISCHETTI, V. A.; NOVICK, R. P.; FERRETTI, J. J.; PORTNOY, D. A.; BRAUNSTEIN, M.; ROOD, J. I. (Eds.). **Microbiol Spectr.**, v. 7, n. 2, p. 7.2.16, 12 abr. 2019.
- TETTELIN, H. et al. Genome analysis of multiple pathogenic isolates of **Streptococcus agalactiae**: Implications for the microbial “pan-genome.” **Proc Natl Acad Sci.**, v. 102, n. 39, p. 13950–13955, 27 set. 2005.
- THIEU, T.; JOSHI, S.; WARREN, S.; KORKIN, D. Literature mining of host–pathogen interactions: comparing feature-based supervised learning and language-based approaches. **Bioinformatics**, v. 28, n. 6, p. 867–875, 15 mar. 2012.
- TRALAMAZZA, S. M.; GLUCK-THALER, E.; FEURTEY, A.; CROLL, D. Copy number variation introduced by a massive mobile element facilitates global thermal adaptation in a fungal wheat pathogen. **Nat Commun.**, v. 15, n. 1, p. 5728, 8 jul. 2024.
- TRAMONT, E. C. Adhesion of **Neisseria Gonorrhoeae** and Disease. In: ELLIOTT, K.; O’CONNOR, M.; WHELAN, J. (Eds.). **Novartis Foundation Symposia** [Internet]. Chichester, UK: John Wiley & Sons, Ltd., 2008 [citado em 16 out. 2024]. p. 188–201. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/9780470720639.ch12>

- VASHISHT, V. et al. Genomics for Emerging Pathogen Identification and Monitoring: Prospects and Obstacles. **BioMedInformatics**, v. 3, n. 4, p. 1145–1177, 7 dez. 2023.
- VESTERGAARD, L. K.; OLIVEIRA, D. N. P.; HØGDALL, C. K.; HØGDALL, E. V. Next Generation Sequencing Technology in the Clinic and Its Challenges. **Cancers**, v. 13, n. 8, p. 1751, 7 abr. 2021.
- WANG, F.; MENG, W.; WANG, B.; QIAO, L. **Helicobacter pylori**-induced gastric inflammation and gastric cancer. **Cancer Lett.**, v. 345, n. 2, p. 196–202, abr. 2014.
- WANG, Y. et al. Linking genomic and epidemiologic information to advance the study of COVID-19. **Sci Data.**, v. 9, n. 1, p. 121, 30 mar. 2022.
- WANG, Y.; ZHAO, Y.; BOLLAS, A.; WANG, Y.; AU, K. F. Nanopore sequencing technology, bioinformatics and applications. **Nat Biotechnol.**, v. 39, n. 11, p. 1348–1365, nov. 2021.
- WANGMO, T. et al. An update on the “empirical turn” in bioethics: analysis of empirical research in nine bioethics journals. *BMC Med Ethics*, v. 19, n. 1, p. 6, dez. 2018.
- WHO. Antimicrobial Resistance [Internet]. WHO Fact-sheets, 2023 [citado em 4 out. 2024]. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>
- WHO. Coronavirus disease (COVID-19) pandemic [Internet]. WHO Emergencies, 2024 [citado em 2024]. Disponível em: <https://www.who.int/europe/emergencies/situations/covid-19>
- WILM, A. et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. **Nucleic Acids Res.**, v. 40, n. 22, p. 11189–11201, 1 dez. 2012.
- WONG, D. W. S. Whole Genome and Next Generation Sequencing. In: *The ABCs of Gene Cloning* [Internet]. Cham: Springer International Publishing, 2018 [citado em 16 out. 2024]. p. 219–230. Disponível em: http://link.springer.com/10.1007/978-3-319-77982-9_24
- WU, L.; WANG, H.; XIA, Y.; XI, R. CNV-BAC: Copy number Variation Detection in Bacterial Circular Genome. *BIROL*, I. (Ed.). **Bioinformatics**, v. 36, n. 12, p. 3890–3891, 1 jun. 2020.
- XIAO, T.; ZHOU, W. The third generation sequencing: the advanced approach to genetic diseases. **Transl Pediatr.**, v. 9, n. 2, p. 163–173, abr. 2020.
- XU, C. A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data. **Comput Struct Biotechnol J.**, v. 16, p. 15–24, 2018.

- XU, Q.; TANG, Y.; HUANG, G. Innate immune responses in RNA viral infection. *Front Med.*, v. 15, n. 3, p. 333–346, jun. 2021.
- YALÇIN, S. et al. Genomic surveillance during the first two years of the COVID-19 pandemic – country experience and lessons learned from Türkiye. **Front Public Health.**, v. 12, p. 1332109, 24 maio 2024.
- YOSHIDA, H.; BOGAKI, M.; NAKAMURA, M.; NAKAMURA, S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of **Escherichia coli**. **Antimicrob Agents Chemother.**, v. 34, n. 6, p. 1271–1272, jun. 1990.
- YUAN, L.; LU, H.; LI, F.; NIELSEN, J.; KERKHOVEN, E. J. HGTphyloDetect: facilitating the identification and phylogenetic analysis of horizontal gene transfer. **Brief Bioinform.**, v. 24, n. 2, p. bbad035, 19 mar. 2023.
- ZENG, D.; JIAO, J.; MO, T. Combination of nucleic acid amplification and CRISPR/Cas technology in pathogen detection. **Front Microbiol.**, v. 15, p. 1355234, 6 fev. 2024.
- ZHANG, T. et al. Complex genome assembly based on long-read sequencing. **Brief Bioinform.**, v. 23, n. 5, p. bbac305, 20 set. 2022.
- ZHU, Y. et al. RNA-based therapeutics: an overview and prospectus. *Cell Death Dis.*, v. 13, n. 7, p. 644, 23 jul. 2022.

Autor

Lucas Gabriel Rodrigues Gomes¹, Eduarda Guimarães Sousa¹, Fernanda Diniz Prates^{1,2}, Talita Pereira Gomes^{1,3}, Gabriel Camargos Gomes¹, Janaíne Aparecida de Paula^{1,4}, Pedro Henrique Marques⁴, Ana Lua de Oliveira Vinhal¹, Bernardo Buhr Alves Mendonça¹, Aline Ferreira Maciel de Oliveira¹, Mariana Letícia Costa Pedrosa¹, Luiza Pereira Reis¹, Marcus Vinicius Canário Viana¹, Siomar de Castro Soares⁴, Sandeep Tiwari⁵, Arun Kumar Jaiswal¹, Vasco Azevedo¹

1. Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.

2. Núcleo Técnico Operacional, Microbiologia, Instituto Hermes Pardini, Vespasiano, Minas Gerais, Brasil.
3. Instituto Nacional de Ciência e Tecnologia em Estudos Interdisciplinares e Transdisciplinares em Ecologia e Evolução (INCT IN-TREE), Instituto de Biologia, Universidade Federal da Bahia, Salvador, Bahia, Brasil.
4. Laboratório de Bioinformática, Departamento de Microbiologia, Imunologia e Parasitologia, Instituto de Ciências Biológicas e Naturais, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brasil.
5. Instituto de Biologia, Universidade Federal da Bahia, Salvador, Bahia, Brasil.

CAPÍTULO 2

Atualizações, impactos e perspectivas das técnicas moleculares: contextualizando a biologia molecular

João Matheus Sobral Pena, Prescilla Emy Nagao, Pamella Silva Lannes-Costa

<https://doi.org/10.69570/mp.978-65-84548-32-9.c2>

Resumo

O potencial das investigações microbiológicas aumentou drasticamente após o desenvolvimento de técnicas moleculares, representando um marco importante na medicina. Na atualidade, os métodos convencionais estão sendo substituídos por métodos avançados como a amplificação de ácidos nucleicos, sequenciamento rápido de DNA, polimorfismo de comprimento de fragmentos de restrição (RFLP), técnicas de hibridização molecular fluorescente *in situ* (FISH), ARISA (análise automatizada de espaçador ribossômico intergênico), diagnóstico baseado na metodologia *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR). Além disso, o desenvolvimento de bibliotecas de clones ribossômicos remodelou a compreensão da microbiologia e revolucionou a abordagem das terapias com agentes antimicrobianos. Essas novas tecnologias facilitaram a investigação aprofundada da microbiota anfibiônica humana, identificando um grande número de organismos existentes ou não cultiváveis. Atualmente, os resultados das técnicas moleculares fornecem informações fundamentais que podem ser correlacionadas com a evolução clínica e o tratamento adequado durante infecções invasivas. As abordagens moleculares para análises epidemiológicas complementam os métodos tradicionais, fornecendo novos métodos filogenéticos e filodinâmicos para rastrear a transmissão de doenças causadas por vários microrganismos, contribuindo para a melhor compreensão e gestão de pandemias e epidemias. Sendo assim, os resultados da investigação translacional podem melhorar a prevenção de doenças e levar a tratamentos médicos mais eficazes e direcionados. Desta forma, o enfoque desta revisão será sobre a detecção de patógenos por meio de métodos rápidos, sensíveis, precisos e acessíveis, demonstrando atualizações, impactos e perspectivas das técnicas moleculares para pesquisas avançadas sobre microrganismos responsáveis por doenças em humanos.

Palavras-chave: DNA, RNA, atualizações, microrganismos, técnicas moleculares.

1. Introdução

Durante a última década, houve um avanço significativo das metodologias moleculares para a detecção de diferentes microrganismos. A implementação das tecnologias de diagnóstico rápido em laboratórios de microbiologia clínica e de pesquisa básica é fundamental para a identificação definitiva de diferentes microrganismos e para o tratamento de amplo espectro. As abordagens utilizando a biologia molecular para análises epidemiológicas avançam e complementam os métodos epidemiológicos tradicionais para a compreensão da frequência e da disseminação das doenças. Atualmente, cada uma das ferramentas disponíveis apresenta vantagens e desvantagens distintas que dependem da natureza do patógeno que está sendo investigado, bem como da qualidade e da quantidade do material genético a ser analisado.

Como marco no século XIX, em 1865, o biólogo austríaco Gregor Mendel foi responsável pela descoberta das leis da herdabilidade que sinalizaram a transformação da biologia em uma ciência comprovada, situada no mesmo nível da química e da física. As quatro leis básicas de Mendel desencadearam uma revolução na biologia, pois forneceram uma base racional para quantificar observações e investigar relações de causa-efeito (GAYON, 2016). Assim, a bioquímica que se concentrava na síntese de novas moléculas passou a ser denominada como biologia molecular. Contudo, a pergunta mais importante sobre a herdabilidade das características genéticas ainda se encontrava em aberto: como as informações hereditárias são armazenadas dentro da célula? Naquela época, o conhecimento do DNA como uma sequência linear de aminoácidos ou nucleotídeos ainda não havia sido descoberta e descrita (KALLENBERGER, 2001).

Classificadas como macromoléculas biológicas, as proteínas e os ácidos nucléicos se tornaram o tema central das investigações dos pesquisadores. Os cientistas buscavam expandir os conhecimentos da recente biologia molecular, área inicialmente definida como a bioquímica relacionada ao DNA e a sua expressão em proteínas. Os estudos da biologia molecular começaram a se concentrar em como a informação genética era armazenada, transmitida e traduzida em fenótipos (KALLENBERGER, 2001). Nos primeiros anos da biologia molecular, os pesquisadores perceberam que os organismos vivos

apresentavam múltiplos sistemas de regulação e *feedback*; ainda nessa primeira temporalidade da biologia molecular, os cientistas defendiam a teoria de que o código genético seria diferente em organismos superiores, especialmente, nos seres humanos (LAI et al, 2001; KELLENBERGER, 2004).

No final da década de 1930, o físico alemão Max Delbruck fez sua migração para a área da biologia para estudar as regras básicas de herança nos organismos mais simples como os vírus bacteriófagos e seus hospedeiros bacterianos. Nos anos seguintes, os estudos com bactérias permitiram a descoberta das sequências de DNA ribossômico 16S, conhecidas como *primer* universal, que permitiram avanços nas análises do microbioma bacteriano (MARTIN; ZENILMAN; LAZARUS, 2010). Essas sequências são encontradas em procariotos, normalmente adjacentes ao DNA ribossômico 16S e, por serem únicas em cada espécie, podem ser usadas para identificação de espécies bacterianas individuais. Posteriormente, ferramentas genômicas, como a amplificação de ácidos nucleicos, sequenciamento rápido de DNA e o desenvolvimento de bibliotecas de clones ribossômicos 16S, revolucionaram nossa capacidade de compreender a microbiologia, podendo revolucionar a abordagem no uso de agentes antimicrobianos para a terapia de doenças crônicas (WHITFIELD, 2013). Desta forma, a genômica bacteriana revolucionou o campo das doenças infecciosas. Os agentes bacterianos causadores da doença da arranhadura do gato e da doença de Whipple, por exemplo, só foram identificados de forma conclusiva com a utilização de métodos genômicos (COSTERTON, 2007).

Atualmente, os marcadores moleculares têm sido usados para identificar os agentes etiológicos de doenças, avaliar o prognóstico e estudar os padrões de transmissão de doenças. Um dos primeiros relatos da aplicação da epidemiologia molecular foi na década de 1970, quando polimorfismos polipeptídicos, identificados por eletroforese em gel de acrilamida, foram usados como marcadores para caracterização viral (RILEY et al., 2018).

O diagnóstico molecular, incluindo sorologia e *Polymerase Chain Reaction* (PCR), fornece ferramentas robustas para a determinação precisa da propagação das doenças causadas por microrganismos. O sequenciamento do genoma viral e bacteriano forneceu novos métodos filogenéticos (ex.

sequenciamento de DNA) e filodinâmicos (ex. técnica de microarray) que são altamente complementares aos meios tradicionais de rastreamento de transmissão. Sendo assim, registros médicos eletrônicos tornam-se um importante recurso para associação de dados demográficos e de cuidados clínicos de pacientes, permitindo uma melhor compreensão da patogênese e da transmissão microbiana (LORENZO-REDONDO, 2021).

As limitações diagnósticas e as incertezas relacionadas à baixa sensibilidade (taxa de positividade), tempo prolongado para detecção de patógenos e tempos de resposta dos resultados, são muitas vezes compensadas pela política liberal do uso de terapia antimicrobiana de amplo espectro. Contudo, considerando o aumento das taxas de resistência antimicrobiana e o surgimento de novos tipos de resistência antimicrobiana, é importante exercer uma abordagem criteriosa para o uso de amplo espectro da terapia antimicrobiana (ASLAM et al., 2018).

Portanto, este capítulo descreve uma atualização e avaliação de melhorias recentes para a detecção de patógenos e para epidemiologia microbiana, se concentrando em tecnologias moleculares de diagnóstico que podem ser classificadas em técnicas específicas para diagnósticos de amostras através de DNA (Reação de cadeia de polimerase, Clonagem, Southern Blotting, Sequenciamento, RFLP, ARISA) , RNA (Northern Blot Analysis, Reação de cadeia de polimerase de transcriptase reversa, Microarray, Hibridização *in situ*) e as técnicas comuns para DNA e RNA (Eletroforese em gel, PCR, Hibridização, *Stable Isotope Probing*, PCR e CRISPR).

2. Técnicas comuns para utilização em amostras de DNA e RNA

A eletroforese e o PCR são as técnicas mais comumente utilizadas em biologia molecular (BARTLETT et al., 2012). A eletroforese é um processo que segrega moléculas, como proteínas ou DNA, com base em seu tamanho, sendo frequentemente utilizada para identificar moléculas específicas ou fragmentos de moléculas (GREEN et al., 2019). A PCR, por outro lado, é um processo usado para amplificar quantidades bem pequenas de DNA, garantindo ampla quantidade de DNA para utilização em outras técnicas que serão descritas ao longo deste capítulo (WATERS et al., 2014).

Dentre as aplicações clínicas e laboratoriais das técnicas moleculares comuns para DNA e RNA podemos citar a identificação de padrões nutrigenéticos que avaliam a variação genética e a dieta alimentar. Esses exames são recomendados para a prevenção de doenças como obesidade, diabetes e hipertensão. Outras aplicações são a avaliação genética personalizada para otimizar o treinamento dos atletas e melhorar o desempenho físico, exames de sexagem fetal e identificação de alterações cromossômicas no bebê, incluindo a síndrome de Down. Sendo assim, existem vários tipos de testes possíveis através da biologia molecular que irão beneficiar diretamente a sociedade.

2.1. Eletroforese em gel

A eletroforese pode ser utilizada para moléculas de DNA e de RNA. A presença de cargas negativas permitem a movimentação no gel em direção ao eletrodo positivo quando adicionadas a uma solução específica e na presença de um campo elétrico (CHANDLER et al., 2006). Os tamanhos das moléculas de DNA e de RNA podem ser utilizados para identificá-las quando posicionadas em uma malha em gel (TOWBIN et al., 1993). O gel é uma malha semi sólida constituída de polímeros à base de agarose ou poliacrilamida que apresenta espaços em sua malha, permitindo que as moléculas do material genético analisado (DNA ou RNA) migrem através desses espaços com base no tamanho da molécula (CHANDLER et al., 2006). As moléculas menores migram de forma mais rápida e facilitada, enquanto as moléculas maiores migram de forma mais lenta (GREEN et al., 2019).

O método mais comum de visualização das moléculas de DNA e RNA é através da coloração do gel de agarose com brometo de etídio, responsável por se intercalar ao longo da estrutura do ácido nucléico, apresentando coloração laranja fluorescente quando submetidos à luz ultravioleta (CHANDLER et al., 2006). Além da coloração tradicional com brometo de etídio, a marcação radioativa com fósforo radioativo para a visualização dos ácidos nucléicos é outra técnica utilizada para coloração do gel de eletroforese (VOYTAS et al., 2001). As aplicações clínicas e laboratoriais da técnica de eletroforese em gel permitem a separação dos componentes individuais presentes em soluções complexas, o

estudo dos componentes de forma isolada e a detecção de doenças genéticas através do padrão de bandas do material genético (BUSZEWSKI et al., 2017; GREEN et al., 2019).

2.2. Hibridização

A hibridização é a formação de moléculas de ácido nucleico com sequências complementares de nucleotídeos nas formas de DNA:DNA, DNA:RNA ou RNA:RNA. O princípio de hibridização, também chamado de anelamento, é utilizado para identificação de produtos de DNA e RNA. A técnica de hibridização é realizada através da ligação de uma sonda previamente marcada ao material genético selecionado, seja esse DNA ou RNA. As sondas são DNAs de cadeia simples marcados com agentes radioativos (radioisótopos) e fluorescentes. A função da sonda utilizada no processo de hibridização é realizar a identificação dos ácidos nucleicos tidos como sequência-alvo; caso a sonda não seja 100% compatível com o alvo, essa será classificada como incompatível (CHANDLER et al., 2006).

Durante a hibridização, os nucleotídeos das sequências complementares são ligados à uma membrana que sofrerá alteração devido ao tamanho dos ácidos nucleicos hibridizados, ao número total de moléculas de guanina e citosina, bem como na concentração dos ácidos nucleicos e o quão próximo os alvos estão das sondas. A hibridização também pode ser realizada através da técnica *in situ* para localizar a expressão de diferentes transcritores em seções de tecido (ex: cortes de tecido, biópsias ou em células), onde as sondas de DNA ou RNA podem conter ou não marcadores radioativos. As vantagens apresentadas para o seu uso são o fato de que ela pode ser aplicada em tecidos congelados ou materiais armazenados e pode ser combinada com técnicas de imunohistoquímica para detectar as proteínas, nem como o mRNA.

A técnica de hibridização *in situ* tem sido utilizada na identificação de sítios de expressão gênica, análise da transcrição de mRNA, mapeamento de sequências de genes em cromossomos. O diferencial dos resultados obtidos através desta técnica é que a localização do mRNA pode ser visualizado no contexto de todo o tecido. A hibridização *in situ* permite estimar o nível e a localização da expressão do RNA mensageiro (MCNICOL et al., 1997). Na

maioria das vezes, as sondas de RNA, também denominadas ribossondas oferecem maior sensibilidade e especificidade em comparação com as sondas de DNA (NOURI-ARIA, 2008).

2.3. SIP (*Stable-Isotope Probing*)

A sondagem de isótopos estáveis de DNA (DNA-SIP) é uma técnica relevante para identificar microrganismos ativos que assimilam substratos de carbono e nutrientes específicos na biomassa celular. A SIP é baseada na assimilação de isótopos estáveis (mais comumente ^{13}C , ^{15}N , ^{18}O e ^2H) na amostra a ser estudada e na subsequente detecção de isótopos na biomassa a granel, células individuais ou componentes celulares usando instrumentação especializada, como hibridização fluorescente *in situ* (FISH) para estudo de amostras de DNA, ou abordagens baseadas no sequenciamento de amplicon de genes marcadores (ex: PCR / sequenciamento de DNA). Após a incubação de uma amostra ambiental com compostos marcados com isótopos estáveis, o ácido nucléico extraído é submetido a ultracentrifugação em gradiente de densidade e subsequente fracionamento do gradiente para separar ácidos nucleicos de diferentes densidades. A purificação de DNA a partir do cloreto de cério permite recuperar o DNA marcado e não marcado para posterior caracterização molecular (impressão digital, microarranjos, bibliotecas de clones, metagenômica) (DUNFORD e NEUFELD, 2010). Sendo assim, esta técnica tem sido importante para atribuir função metabólica às diversas comunidades que habitam ambientes terrestres e aquáticos.

A SIP é uma técnica usada para sondar a identidade e a função de microrganismos não cultivados em amostras ambientais. Empregando isótopos estáveis de elementos comuns, como carbono e nitrogênio, o RNA-SIP consegue explorar um aumento na densidade flutuante do RNA causado pelo metabolismo ativo e pela incorporação de isótopos de massa mais pesada no RNA (GHORI et al, 2019). Os RNAs marcados são subsequentemente separados dos RNAs não marcados por centrifugação em gradiente de densidade, seguida pela identificação dos RNAs por sequenciamento. Portanto, a sondagem de isótopos estáveis de RNA é uma técnica que fornece

informações simultâneas sobre a comunidade, composição e função do microbioma.

A SIP é uma estratégia poderosa para elucidar a ecofisiologia de microrganismos em microbiotas complexas associadas ao hospedeiro. A análise metabolômica produziu uma imensa quantidade de conhecimento sobre a composição e o desenvolvimento dos microbiomas humanos e animais, mas a maioria das funções codificadas no metagenoma são desconhecidas. A SIP é uma estratégia para elucidar a ecofisiologia de microrganismos individuais em microbiotas complexas associadas ao hospedeiro. Uma série de abordagens SIP foi desenvolvida para permitir testar hipóteses de conjuntos de dados metabolômicos e descobrir novas funções. A SIP está sendo usada para determinar as estratégias de utilização de nutrientes dos micróbios *in vivo* e *ex vivo*, bem como as relações ecológicas, como competição, cooperação e alimentação cruzada (ABU-ALI, 2018).

No entanto, os métodos SIP são demorados e muitas vezes requerem instrumentação cara e de alta qualidade, o que limita severamente o número de amostras que podem ser analisadas e, portanto, requer um planejamento experimental cuidadoso. A maioria dos estudos SIP concentra-se na utilização microbiana de compostos específicos e totalmente marcados isotopicamente (SIP mediada por substrato). No entanto, também é possível usar compostos como marcadores de atividade geral, como água na forma de $2\text{H}_2\text{O}$, $^{13}\text{CO}_2$ ou marcação total de ^{15}N , pois as células ativas irão incorporar hidrogênio da água, carbono ^{13}C durante a assimilação autotrófica ou heterotrófica de CO_2 , independentemente de sua fisiologia (WALCZYK et al., 2002).

2.4. Repetições Palindrômicas Curtas Agrupadas E Regularmente Inter Espaçadas (CRISPR_*Clustered Regularly Interspaced Short Palindromic Repeats*)

As repetições palindrômicas curtas agrupadas e regularmente inter-espaçadas, conhecidas pelo nome em inglês *Clustered Regularly Interspaced Palindromic Repeat* (CRISPR) remodelou o diagnóstico atual. CRISPR são as sequências de DNA encontradas em bactérias e arqueas que têm sido amplamente utilizadas em experimentos de edição de genes, são como o

sistema imune adaptativo presente na maioria das arqueas e em muitas bactérias, promovendo proteção (MARRAFFINI et al., 2010; BARRANGOU et al., 2014). Dentro do sistema CRISPR, existe um locus que consiste em repetições breves e consecutivas que são intercaladas com diversas sequências de DNA conhecidas como “espaçadores”. Esses espaçadores são cercados por múltiplos genes *cas*. O processo de “imunidade” dentro do sistema CRISPR-Cas abrange três mecanismos distintos: adaptação, expressão e interferência (GARNEAU et al., 2010; SAMPSON et al., 2014).

Na fase inicial de adaptação, a célula integra fragmentos externos de DNA, chamados de “protoespaçadores”, derivados de vírus e plasmídeos invasores, no arranjo CRISPR para criar novos espaços. Esses espaços servem como sequências de memória que permitem a defesa direcionada contra futuras invasões do mesmo vírus ou plasmídeo. Passando para a fase de expressão, o array CRISPR é transcrito como um precursor conhecido como pré-crRNA que passa por processamento e maturação para produzir RNAs CRISPR (crRNAs) (WESTRA et al., 2012; BARRANGOU, 2013). Finalmente, na fase de interferência, os crRNAs, com o auxílio de proteínas Cas, atuam como guias para identificar alvos específicos e clivar o ácido nucleico do vírus ou plasmídeo correspondente. Estudos recentes indicaram que o sistema CRISPR-Cas tem a capacidade não apenas de defesa contra ameaças, mas também de regular o comportamento coletivo e a patogenicidade (WESTRA et al., 2014; LOUWEN et al., 2014).

Originalmente encontrado como um mecanismo de defesa em bactérias, o sistema CRISPR/Cas é uma grande promessa no combate a doenças infecciosas. Além disso, tem um vasto potencial como ferramenta para o tratamento de doenças genéticas como o câncer e doenças monogênicas, editando ou eliminando mutações genéticas, incluindo oncogenes e proto-oncogenes (JINEK et al., 2012; JIANG et al., 2019). Também desempenham um papel importante na defesa antiviral, pois as sequências são derivadas de bacteriófagos que já infectaram bactérias (BARRANGOU, 2015). Muitas técnicas baseadas em CRISPR estão atualmente em uso ou têm potencial para ser uma opção de teste no local de atendimento para o diagnóstico de patógenos (ISLAM e IQBAL, 2020).

3. Técnicas para utilização em amostras de DNA

3.1. Reação em cadeia da polimerase (PCR_ *Polymerase chain reaction*)

De acordo com a literatura científica, a técnica de PCR foi desenvolvida em 1987 e, atualmente, se apresenta como a técnica de biologia molecular mais utilizada na ciência. As aplicações dessa técnica incluem diagnósticos clínicos, análises genéticas e engenharia genética (ROGERS, 2015; KERI e CHANDRASEKAR, 2023).

A técnica é utilizada para amplificar sequências específicas de DNA através da utilização de um modelo de DNA que contém a sequência alvo a ser estudada, como exemplo, a sequência do paciente. Um par de oligoiniciadores, conhecidos como *primers*, também são necessários. A maioria das técnicas de reação em cadeia da polimerase utiliza um oligoiniciador direto (*forward*) que está localizado na extremidade 5' da sequência alvo que será analisada e um oligoiniciador reverso (*reverse*) que está localizado na extremidade 3' na fita complementar. Enzimas resistentes ao calor são utilizadas para conduzir as reações, sendo a enzima *taq* polimerase (*Thermus aquaticus*) utilizada rotineiramente. Dentre as vantagens conferidas pela termoestabilidade da *taq* polimerase está a sua capacidade de suportar o aquecimento e resfriamento repetidos inerentes à PCR e de sintetizar DNA em altas temperaturas que desnaturam os oligoiniciadores incompatíveis. A enzima, contudo, não é infinitamente resistente ao calor e, para maior eficiência, não deve ser submetida a etapas desnecessárias de desnaturação. Adicionar mais enzima às vezes pode aumentar o rendimento de produtos de PCR inespecíficos em detrimento do produto de interesse. A reação de PCR necessita de um grande suprimento de nucleotídeos para confecção das fitas de DNA e tampão apropriado (KERI e CHANDRASEKAR, 2023).

Posteriormente, a reação criada com a adição de todos esses componentes é colocada em um aparelho termociclador programado para mudar rapidamente a temperatura de acordo com a programação de ciclos pré definidos. A mistura é aquecida a uma temperatura entre 90°C e 96°C para que aconteça a desnaturação ou a separação da fita dupla de DNA adicionado à solução como modelo. Após essa fase inicial, a temperatura cai para um intervalo entre 50°C a 65°C permitindo que os *primers* realizem a ligação e o

emparelhamento com a sequência de DNA adicionada como modelo. Por último, o termociclador eleva a temperatura para 72°C durante 1 a 2 minutos para que a enzima taq polimerase alongue as novas cadeias de DNA iniciadas pelos oligoiniciadores. A fase de alongamento das cadeias ocorre no sentido 5' para 3'. Em cada ciclo é realizada a duplicação da quantidade de DNA presente na mistura inserida no termociclador, sendo assim, ao final de 300 ciclos, haverá mais de 1 milhão de cópias de DNA de fita dupla presentes na amostra de trabalho. O produto da PCR será analisado em eletroforese em gel de agarose (YANG et al. 2004; EVANS et al. 2009).

3.2. Hibridização Fluorescente *in situ* (FISH_fluorescence *in situ* hybridization)

O desenvolvimento de técnicas de hibridização molecular, como a hibridização fluorescente *in situ* (FISH_fluorescence *in situ* hybridization), teve um grande impacto na detecção e caracterização das alterações genéticas que dão origem aos tumores humanos. Com sondas projetadas para identificar cromossomos e regiões cromossômicas específicas, a FISH é usada rotineiramente por laboratórios de citogenética e patologia para identificar anormalidades cromossômicas associadas a doenças hematológicas malignas (GOZZETTI e LE BEAU, 2000).

A técnica FISH envolve a preparação de dois componentes principais: a sonda de DNA e o DNA alvo. A sonda de DNA normalmente vem de fontes clonadas, como plasmídeos, cosmídeos, PACs (*P1 derived artificial cromossomes*), YACs (*Yeast artificial cromossomes*) ou BACs (*Bacterial artificial cromossomes*); onde a inserção pode conter um gene específico ou originar-se de um *locus* cromossômico específico. O DNA purificado pode então ser marcado e detectado indiretamente. Inicialmente a descrição dessa técnica preconizava a utilização de sondas marcadas radioativamente. Na atualidade, os protocolos sugerem a utilização de sondas não-radioativas que apresentam vantagens como: a alta resolução e menor tempo de processamento. A marcação isotópica adiciona enzimas, haptenos ou fluorocromos. Os últimos protocolos utilizam a detecção por flurocromos (moléculas que fluorescem quando excitadas por um comprimento de luz ultravioleta específico).

O DNA alvo pode assumir a forma de cromossomos espalhados ou núcleos interfásicos. As fontes de alvos interfásicos podem vir de preparações citogenéticas ou de tecidos embebidos em parafina. A sonda de DNA marcada e o DNA alvo são desnaturados para um estado de cadeia simples e podem hibridizar entre si. Lavagens pós-hibridização e incubações de anticorpos marcados com fluorescência durante 24 horas possibilitam a visualização por microscopia fluorescente. A interpretação bem-sucedida dos experimentos FISH depende da qualidade dos materiais de partida, da eficiência da hibridização, do rigor das lavagens pós-hibridização e das detecções de anticorpos (TSUCHIYA, 2011).

3.3. Clonagem

A técnica de clonagem pode ser dividida em algumas fases, na primeira fase isola-se o fragmento do DNA de interesse, utilizando-se enzimas de restrição, estas enzimas são responsáveis por reconhecer a sequência alvo específica e cortar de forma seletiva o fragmento que será utilizado. Na segunda fase, o fragmento de DNA é inserido em um vetor; os plasmídeos são moléculas comumente utilizadas como vetores para clonar fragmentos de DNA. Esse processo permite que o fragmento do gene alvo se ligue ao vetor, através da DNA ligase, formando uma molécula de plasmídeo recombinante contendo o gene de interesse. Ao final dessa etapa, o DNA recombinante é transportado para dentro de uma célula hospedeira, na qual ocorre o processo de transformação bacteriana. Assim, a célula que recebeu o DNA recombinante passa a ser denominada de célula transformada. A célula transformada passará por múltiplos ciclos de divisão produzindo várias cópias de DNA recombinante (BROWN et al., 2003).

O DNA obtido como produto do processo de clonagem pode, então, ser utilizado em outros experimentos como a PCR (MIKIĆ, ALOMARI e GOWERS, 2023). Situações onde o gene de interesse não é conhecidamente descrito, a abordagem realizada será alinhar as sequências deste gene como duas ou até três espécies próximas e realizar a comparação da sua homologia. Para realização desta técnica, o DNA da espécie de interesse é usado como modelo. Posteriormente, serão adicionados *primers*, nucleotídeos responsáveis por

sintetizar a nova fita de DNA, bem como a enzima taq polimerase para conduzir a reação (ALBERTS et al., 2002; MIKIĆ, ALOMARI e GOWERS, 2023).

A maioria das técnicas de clonagem utiliza uma temperatura de anelamento mais baixa e um número maior de ciclos, o que permite uma ligação mais íntima entre *primers* e as áreas que não estão exatamente corretas, resultando em múltiplas bandas no gel de agarose. Para descobrir qual banda de DNA foi o produto dessa técnica, outras ferramentas de biologia molecular deverão ser aplicadas em sequência, como exemplo as técnicas de *Southern Blot* ou a técnica de sequenciamento de DNA (ALBERTS et al., 2002).

3.4. Polimorfismo de comprimento de fragmento de restrição (RFLP_Restriction Fragment Length Polymorphism)

Polimorfismo de comprimento de fragmento de restrição (RFLP) é uma diferença nas sequências homólogas de DNA que pode ser detectada pela presença de fragmentos de diferentes comprimentos após digestão das amostras de DNA com endonucleases de restrição específicas. O RFLP, como marcador molecular, é específico para uma única combinação clone/enzima de restrição. O isolamento de DNA suficiente para análise de RFLP é demorado e trabalhoso. Contudo, a PCR pode ser utilizada para amplificar quantidades muito pequenas de DNA, geralmente em 2-3 horas, até aos níveis necessários para a análise RFLP. Portanto, mais amostras podem ser analisadas em menos tempo. Um nome alternativo para a técnica é o ensaio de Sequência Polimórfica Amplificada Clivada (CAPS) (NARAYANAN, 1991).

O uso principal de RFLPs é como parte de uma abordagem de gene suspeito, na qual é analisado um polimorfismo específico conhecido por estar dentro ou muito próximo de um gene candidato. Se o gene candidato não tiver nenhum polimorfismo associado conhecido, uma busca por polimorfismos é usada para encontrar um RFLP. O uso de RFLPs como parte de pesquisas genômicas, embora bastante bem-sucedido no passado, tornou-se menos comum à medida que marcadores mais novos e mais poderosos, como repetições de sequência simples (SSRs), foram desenvolvidos (JARCHO, 2001).

A análise RFLP tem sido usada na localização de genes de doenças para distúrbios mendelianos simples, incluindo condições autossômicas dominantes, autossômicas recessivas e ligadas ao cromossomo X. Além disso, os RFLPs têm sido usados para investigar mutações de deleção, como perda de heterozigosidade em doenças malignas, para analisar distúrbios poligênicos (diabetes e hipertensão) e para distinguir indivíduos uns dos outros em estudos forenses. A maioria dos RFLPs conhecidos ocorre como resultado de mutações pontuais ou variações de nucleotídeo único na sequência de DNA, embora inserções e deleções também sejam responsáveis por alguns sítios polimórficos.

Nem todas as mutações naturais podem ser detectadas pela análise de RFLP, pois muitas não resultam numa alteração no comprimento dos fragmentos de restrição gerados com enzimas conhecidas. Desta forma, aproximadamente 40% das substituições de nucleotídeo único podem ser detectadas usando um conjunto de 30 endonucleases de restrição (WHITE e LALOUEL, 1987). Algumas enzimas de restrição expõem mais variantes de sequência do que outras, devido ao número de sítios de restrição reconhecidos por cada enzima. Em geral, enzimas que possuem mais locais de reconhecimento detectam mais variantes de sequência (WIJSMAN, 1984; HAINES et al., 2001).

3.5. Análise Automatizada de Espaçador Intergênico Ribossômico (ARISA_Automated Ribosomal Intergenic Spacer Analysis)

ARISA (Análise Automatizada de Espaçador Intergênico Ribossômico) é um método de análise molecular que fornece estimativas de riqueza e diversidade microbiana, sendo baseado na heterogeneidade de comprimento do operon rRNA bacteriano 16S-23S denominado espaçador intergênico (também conhecido como espaçador transcrito interno ou ITS). Desde a sua introdução em 1999, o ARISA tem sido utilizado para o estudo do solo, ambientes aquáticos e intestino humano. ARISA tornou-se uma ferramenta muito útil para comparar a estrutura da comunidade em múltiplas amostras com base em padrões previamente descritos. No entanto, também pode ser usado para estimar a diversidade bacteriana, apresentando a possibilidade de estimar a quantidade, pois microrganismos não relacionados podem expor regiões espaçadoras de comprimento idêntico. Por outro lado, múltiplos operons dentro de um único

genoma podem diferir no comprimento do espaçador, levando assim a uma superestimação da variedade. Dessa forma, são necessárias investigações adicionais para obter uma compreensão mais clara dos dados obtidos pela técnica ARISA (KOVACS, YACOBY e GOPHNA, 2010).

3.6. Sequenciamento de DNA

As fitas de DNA são escritas através da combinação e alternância dos nucleotídeos adenina, timina, citosina e guanina (A, T, C, G). Uma sequência de DNA é a ordem própria na qual os nucleotídeos são organizados para realizar a transcrição de genes específicos (GLENN, 2011). Com isso, a técnica de sequenciamento de DNA basicamente é a ferramenta responsável por elucidar a ordem que os nucleotídeos compõem um trecho da fita de DNA amplificada através da técnica de PCR ou através da técnica de clonagem de DNA. As sequências de DNA que apresentam um montante menor que 700 nucleotídeos podem ser obtidas a partir de uma reação de sequenciamento, trechos maiores de DNA precisam ser seccionados em fragmentos menores para então serem sequenciados pelas técnicas (LIU et al., 2012).

De forma simplificada, dois métodos de sequenciamento de DNA são utilizados comumente na rotina laboratorial, o primeiro é através do método químico (Maxam-Gilbert) e o segundo é através do dideoxi enzimático (Sanger). Na década de 1970, os cientistas desenvolveram o método Maxam-Gilbert, uma técnica que envolve a fragmentação química do DNA em segmentos precisos. Esses segmentos são posteriormente isolados e analisados para determinar a sequência de bases nitrogenadas no DNA. Este método inovador desempenhou um papel fundamental no sequenciamento do genoma humano e serviu como pedra angular no Projeto Genoma Humano. O método Maxam-Gilbert possui a capacidade de sequenciar com precisão seções específicas de DNA, o que é vantajoso. Esse recurso permite a identificação de mutações genéticas e o exame da estrutura genética. No entanto, o método Maxam-Gilbert é dificultado pela sua natureza trabalhosa e demorada, o que restringe a sua implementação generalizada (STAMM e LONGO, 1990; VERMA e KULSHRESTHA, 2017).

Outra aplicação envolve produtos de PCR feitos com *primers* degenerados. Frequentemente, resultados ruins são obtidos se *primers*

degenerados forem usados no sequenciamento didesoxi, devido a uma preparação incorreta. Portanto, um método simples que não requer uma etapa enzimática adicional (exonuclease lambda), re-amplificação ou uma correspondência exata de *primer* para gerar fragmentos de DNA para sequenciamento de Maxam-Gilbert (Eren, Taktakoğlu e Pirim, 2022). Embora tenha suas limitações, o método Maxam-Gilbert continua a ser utilizado em laboratórios de pesquisa para sequenciar pequenas seções do genoma. À medida que a tecnologia progrediu, surgiram novos métodos de sequenciamento de DNA. Assim, o método Maxam-Gilbert se destaca como uma técnica pioneira para sequenciamento de DNA e mantém sua relevância em determinadas circunstâncias. No entanto, à medida que a tecnologia avança, surgem métodos novos e rápidos para sequenciar os genomas dos seres humanos e de outros organismos com maior eficiência (VERMA e KULSHRESTHA, 2017).

Além disso, ainda existem duas metodologias mais prevalentes: o método Sanger, também conhecido como método de terminação de cadeia, e o sequenciamento de próxima geração (NGS). O método Sanger, iniciado por Frederick Sanger e Alan Coulson em 1977, é uma das primeiras técnicas empregadas para sequenciamento de DNA. Para ler a sequência de bases, o método Sanger emprega ddNTPs marcados com cores distintas para interromper a replicação do DNA em locais específicos. Amplamente utilizado em laboratórios de pesquisa e diagnóstico, este método é altamente considerado por sua confiabilidade e precisão (KRAMER e COEN, 2001). A análise de múltiplas amostras de DNA simultaneamente e em ritmo acelerado é possível graças ao sequenciamento de próxima geração (NGS), uma técnica mais moderna. Este método depende de sequenciamento paralelo, em que milhões de fragmentos de DNA são sequenciados simultaneamente (EREN, TAKTAKOĞLU e PIRIM, 2022).

Além do método Sanger e do sequenciamento de próxima geração, existem várias outras técnicas amplamente utilizadas na área de sequenciamento. Isso inclui sequenciamento por síntese (Illumina), sequenciamento de torrente de íons (Ion Proton) e sequenciamento de nanoporos (MinION). Cada um destes métodos tem o seu próprio conjunto de benefícios e restrições, onde a escolha da técnica depende do objetivo da

investigação, bem como dos recursos financeiros disponíveis (MARDIS, 2008; MOHAMED e SYED, 2013). Portanto, o método Sanger e o sequenciamento de próxima geração são as principais técnicas empregadas para sequenciamento de DNA nos tempos contemporâneos, complementados por métodos alternativos como sequenciamento sintético, sequenciamento de torrente de íons e sequenciamento de nanoporos. Essas técnicas têm revolucionado a genética e a biologia molecular, possibilitando avanços significativos no estudo do DNA e sua aplicação em diversas áreas da ciência (VERMA e KULSHRESTHA, 2017). As possibilidades de aplicações de sequenciamento de DNA se expandiram nas últimas décadas, moldado em parte pelas crescentes atualizações dessas tecnologias de sequenciamento. Incluindo as aplicações clínicas do sequenciamento, ressequenciamento do genoma individual, sequenciamento clínico e a transformação de sequenciadores em dispositivos de contagem molecular (SHENDURE et al., 2017).

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

Os dados inseridos nas diferentes plataformas de sequenciamento, utilizam diferentes metodologias experimentais, gerando sequências com tamanhos desiguais e em formatos distintos. Além disso, cada plataforma NGS possui tendências associadas à metodologia, tal como, erros de nomeamento de bases (LIU et al., 2012). Para lidar com esses desafios, houve um grande desenvolvimento de programas e ferramentas para análise de dados, desde a avaliação da qualidade das leituras até a anotação e montagem dos genomas. As plataformas NGS utilizam diferentes metodologias de sequenciamento, mas

em comum, elas não precisam passar pela etapa de clonagem dos fragmentos de interesse. Adicionalmente, as tecnologias NGS são capazes de gerar milhões de sequências em uma única corrida (THUDI *et al.*, 2012). Dentre essas plataformas podemos salientar as da Illumina (*Genome Analyser*), pois se destacam pelo grande volume de sequências obtidas, grande espectro de utilidade, além de ter o menor custo por base sequenciada do mercado, fazendo com que a empresa venha dominando a indústria de sequenciamento em larga escala (QUAIL *et al.*, 2012). O sistema de sequenciamento HiSeq da Illumina tem alto desempenho e é amplamente adotado. Em termos de dispêndio monetário, o HiSeq reduziu o custo do sequenciamento a níveis incomparáveis (Illumina). O método de sequenciamento adquirido pela Illumina foi uma grande inovação, e atualmente, essa metodologia viabiliza o sequenciamento de fragmentos de 2 x 150 pb na plataforma HiSeq e 2 x 300 pb na MiSeq (GLENN, 2011).

O surgimento de novas tecnologias de sequenciamento possibilitou o sequenciamento de genomas completos em poucas horas (HENSON; TISCHLER; YI *et al.*, 2014), contribuindo para o avanço no conhecimento dos três domínios da vida. Um grande exemplo ocorreu, nos últimos 15 anos, com domínio *Bacteria* apresentando 83.246 projetos relacionados e também favoreceu o crescimento do número de *drafts*, até o final do ano de 2023, com 4.744 drafts e de 4230 projetos de genomas completos depositados nos bancos de dados públicos. Portanto, se pode observar o avanço e o crescimento impressionante no número de pesquisas envolvendo esta área, se tornando de extrema importância para o desdobramento de análises mais profundas de organismos até então pouco conhecidos.

4. Técnicas para utilização em amostras de RNA

4.1. Reação em cadeia da polimerase de transcriptase reversa (RT-PCR)

A PCR em tempo real só pode ser aplicada para amplificação de moldes na forma de moléculas de DNA. Para detectar e quantificar uma amostra de RNA, a molécula de RNA é primeiramente transcrita reversamente em um DNA complementar (cDNA) usando transcriptase reversa, seguida pela conversão do cDNA de fita simples em DNA de fita dupla. Este DNA de fita dupla é então

amplificado usando PCR padrão. Esse procedimento é conhecido como reação em cadeia da polimerase com transcrição reversa em tempo real (RT-PCR em tempo real) (HILLMANN et al., 2009; BACHMAN et al. 2013; KARUNANATHIE et al., 2022).

A PCR em tempo real, também conhecida como PCR quantitativa, é uma modificação da PCR básica que permite monitorar o progresso da PCR. Além de melhorar a acurácia, sensibilidade e rapidez, uma das principais vantagens da PCR em tempo real sobre a PCR básica é que esta técnica fornece uma relação de quantificação confiável entre o número de sequências alvo iniciais (antes da amplificação por PCR) e a quantidade de amplicon acumulada em um determinado ciclo de PCR. A quantificação precisa dos ácidos nucleicos alvo é uma etapa crítica para a quantificação do RNAm durante a análise da expressão gênica e da determinação da carga viral de um espécime clínico, não havendo a necessidade de processos pós-PCR, minimizando a chance de contaminação cruzada devido a *amplicons* prévios (OLIVEIRA et al., 2005). A técnica de PCR em tempo real revolucionou a detecção e quantificação de ácidos nucleicos alvo, possuindo atualmente, inúmeras aplicações. As principais vantagens deste método são a rapidez de *set-up*, o baixo custo e o menor manuseio das amostras para reduzir erros de pipetagem e contaminação. Entretanto, como este método emprega oligoiniciadores gene-específicos tanto para a RT-PCR quanto para a PCR convencional que ocorrem em um tubo de reação, outros genes de interesse não podem ser amplificados para análise posterior (ARTIKA et al., 2022).

No final de 2019, a pandemia de COVID-19 causada pelo coronavírus 2 da síndrome respiratória aguda grave (SARS-CoV-2), atingiu o mundo e deu origem a um grande desafio para os laboratórios de saúde pública. O diagnóstico padrão-ouro para a infecção SARS-CoV-2 é um “teste de amplificação de ácido nucleico” (nucleic acid amplification test - NAAT), e o ensaio de PCR em tempo real foi a principal técnica aplicada, juntamente com o método de sequenciamento de DNA, a técnica de RT-PCR em tempo real foi aplicada com sucesso para detectar e identificar pela primeira vez o novo coronavírus 2019 recém-emergido em Wuhan, China, em dezembro de 2019. Assim, o procedimento RT-PCR em tempo real foi adotado pela Organização Mundial da

Saúde como um método padrão para a confirmação de infecções agudas por SARS-CoV-2 devido à sua sensibilidade e especificidade (OMS, 2021).

Além de oferecer grande sensibilidade e especificidade, a PCR em tempo real pode ser aplicada tanto para análises qualitativas quanto quantitativas. Portanto, tornou-se o método de escolha para a detecção e quantificação rápida e sensível de ácido nucleico em amostras biológicas para as mais diversas aplicações, tais como análise de expressão gênica, detecção de mutação, determinação do status de câncer, análise de microRNA, detecção de organismos geneticamente modificados, detecção bacteriana, quantificação bacteriana, detecção viral e medição de carga viral. Devido à sua versatilidade e utilidade, a técnica de PCR em tempo real tem sido empregada em muitas áreas de pesquisa, incluindo biomedicina, microbiologia, ciências veterinárias, agricultura, farmacologia, biotecnologia e toxicologia (ARTIKA et al., 2022).

4.2. Análise Northern Blot

A análise Northern Blot é descrita na literatura científica como a técnica mais sensível para a detecção da expressão dos níveis de RNA mensageiro (CHANDLER et al. 2006). O RNA é separado por tamanho através de um gel eletroforético de poliacrilamida e agarose (WIEGARD et al. 2021). De forma similar com as outras formas de eletroforese, a movimentação da molécula no gel se apresenta de forma inversamente proporcional ao seu tamanho, os maiores fragmentos não realizam uma migração tão rápida quanto os fragmentos menores (SCHNEIDER et al., 2018). Como o RNA apresenta como característica sua carga negativa, sua migração no gel de agarose ocorre em direção ao cátodo (CHANDLER et al., 2006).

Após a separação através da eletroforese, o RNA da amostra é transferido para uma membrana de nitrocelulose e o RNA é ligado à membrana através da utilização de irradiação ultravioleta (WANG et al., 2013). Sequências específicas de RNA podem ser detectadas pela membrana de transferência através de técnicas de hibridização (CHANDLER et al., 2006). A sonda pode ser marcada através da ligação de uma proteína como a digoxigenina que, por sua vez, pode ser detectada com o auxílio de um anticorpo anti-digoxigenina conjugado com fosfatase alcalina que converte um substrato corado permitindo a análise; além

disso, a sonda também pode ser radiomarcada e detectada através de autorradiografia (Lin et al., 1993). Os resultados dos experimentos com a técnica de hibridização são responsáveis por entregar a quantidade, o tamanho e a abundância de transcritos, sendo a quantidade do produto proporcional à quantidade do RNA específico presente na amostra analisada (WIEGARD et al., 2021). Através da utilização desta técnica laboratorial, a concentração de um gene de manutenção (*housekeeping*) pode ser medido no mesmo experimento do *Northern blot* para avaliar a intensidade relativa do sinal (HERRERA et al., 2005).

Apesar da técnica apresentar a desvantagem de ser demorada e exigir quantidades significativas de RNA para a análise do material genético da amostra, a literatura científica a descreve como uma técnica rotineiramente utilizada, pois os dados fornecem informações quantitativas permitindo uma avaliação precisa do RNA mensageiro (CHANDLER et al., 2006). Na medicina humana, a análise *Northern Blot* pode ser empregada para o estudo dos níveis de expressão dos genes causadores de câncer, quantificando os genes superexpressos e subexpressos nas células cancerígenas em comparação com as células normais presentes no tecido (WIEGARD et al., 2021).

4.3. Microarranjo (*Microarray*)

O microarray é uma ferramenta da biologia molecular que pode ser empregada para analisar a expressão relativa de um gene através da determinação da quantificação de RNA mensageiro presente na amostra. A técnica de microarray pode ser considerada a técnica reversa a performada nos ensaios de *Northern Blot*, enquanto nos ensaios de *Northern Blot* as sondas analisadas procuram ligar-se ao material genético que está ligada a uma membrana, nos ensaios de microarray, sondas para milhares de transcrições estão imobilizadas na superfície de um chip ou de uma lâmina, onde uma solução contendo RNA marcado é adicionada. Essa tecnologia permite que o material genético seja localizado ou sintetizado diretamente em uma superfície sólida de maneira ordenada, facilitando a análise de forma rápida e eficiente da expressão de milhares de genes em um único experimento (MARZANCOLA e SEDIGHI, 2016).

Após a etapa inicial de hibridização da amostra ser concluída, o material genético é ligado ao microarranjo que é colocado em um scanner composto por lasers exclusivamente projetados, um microscópio e um detector. As tags fluorescentes ficam excitadas pelo laser e a intensidade da coloração refletida entre as cores vermelho e verde corresponderá à quantidade de material genético presente em cada amostra. O microscópio e o detector são responsáveis por gerar uma imagem digital do arranjo e a intensidade do espectro vermelho-amarelo-verde é analisada para cada ponto do microarranjo. Os pesquisadores evidenciam que, embora o microarray seja uma boa técnica para estudar mudanças globais entre duas amostras, essa técnica deve ser validada por outras técnicas de biologia molecular (ANDERSON, 2016).

Várias análises podem ser abordadas usando os microarranjos como ponto inicial de investigação, tais como as mutações em um determinado gene da doença, diferenciação de subcategorias moleculares de doenças, escaneamento do número de cópias do material genético que podem significar doenças adquiridas e a identificação de marcadores associados à doenças (MARZANCOLA e SEDIGHI, 2016). De acordo com a literatura médica, a aplicação mais frequente dos microarranjos na medicina é para a detecção e genotipagem rápida de patógenos incluindo bactérias, fungos e vírus (GOVINDARAJAN et al. 2012).

5. Perspectivas e Considerações Finais

As perspectivas das abordagens moleculares aplicadas aos microrganismos são promissoras, com a expectativa de avanços contínuos em biotecnologia, medicina de precisão e bioinformática. Atualmente, as ferramentas de sequenciamento gênico em alta escala permitem análises em mais de cem genes relacionados à hereditariedade. A demanda por avanços na biologia molecular e na genética é generalizada em escala global. Especificamente, o campo da saúde humana necessita de estratégias inovadoras que simplifiquem a análise molecular diretamente no local de atendimento, incluindo o desenvolvimento de plataformas compactas e portáteis que atendem a doenças infecciosas e não transmissíveis. Além disso, existe uma necessidade premente de redução de custos nos métodos de sequenciamento

de DNA, permitindo análises genômicas para identificação de doenças em um número maior de pacientes.

O estudo da genética e da biologia molecular tem se beneficiado muito com a utilização de diversas técnicas destacadas neste capítulo. No entanto, estas abordagens têm as suas limitações, o que pode levar a lacunas na nossa compreensão do genoma. Uma área importante que requer melhorias é a precisão do sequenciamento e a capacidade de cobrir todo o genoma. Isto é especialmente desafiador para regiões do genoma que são ricas em conteúdo GC ou consistem em longos trechos de DNA homopolímero. Além disso, as plataformas atuais usadas para sequenciamento normalmente produzem comprimentos de leitura curtos, o que dificulta nossa capacidade de analisar com precisão grandes regiões de repetição, *indels* e variações estruturais. Como resultado, porções significativas do genoma permanecem inexploradas ou são representadas de forma imprecisa.

Há um otimismo clínico para o tratamento de doenças potencialmente fatais, como o câncer, graças aos mais recentes avanços na tecnologia CRISPR. Embora tenham sido feitos progressos notáveis, ainda há uma série de desafios técnicos que precisam ser enfrentados como a eficiência limitada no reparo indel ou dirigido por homologia (HDR), entrega de componentes do sistema Cas *in vivo* e potenciais reações imunológicas. A superação destes obstáculos exige avanços tecnológicos significativos ou o desenvolvimento de abordagens inovadoras que possam combater eficazmente doenças graves e, ao mesmo tempo, minimizar quaisquer efeitos adversos. Portanto, este capítulo introduz atualizações e os impactos que as técnicas moleculares têm fornecido para o diagnóstico e tratamento de doenças, especialmente em humanos.

6. Referências

- ABU-ALI. Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat. Microbiol.* 3, 356–366. 2018.
- ANDERSON, Sharon. Chromosome Microarray. *MCN Am J Matern Child Nurs.* Sep-Oct;41(5):272-9. doi: 10.1097/NMC.000000000000260. PMID: 27276104. 2016.
- ALBERTS, B; JOHNSON, A; LEWIS J. *Molecular Biology of the Cell.* 4th edition. New York: Garland Science; Isolating, Cloning, and Sequencing

DNA. 2002.

ASLAM, Bilal; et al. Antibiotic resistance: a rundown of a global crisis. *Infect Drug Resist.* Oct 10;11:1645-1658. doi: 10.2147/IDR.S173867. PMID: 30349322; PMCID: PMC6188119. 2018.

ARTIKA, IM et al. Real-Time Polymerase Chain Reaction: Current Techniques, Applications, and Role in COVID-19 Diagnosis. *Genes (Basel)*. Dec 16;13(12):2387. doi: 10.3390/genes13122387. PMID: 36553654; PMCID: PMC9778061. 2022.

BACHMAN, Julia; Reverse-transcription PCR (RT-PCR). *Methods Enzymol.* 530:67-74. doi: 10.1016/B978-0-12-420037-1.00002-6. PMID: 24034314. 2013.

BARTLETT, Rebecca et al. Fundamental approaches in molecular biology for communication sciences and disorders. *J Speech Lang Hear Res.* doi: 10.1044/1092-4388(2011/11-0152). Epub 2012 Jan 9. PMID: 22232415; PMCID: PMC3418393. 2012.

BARRANGOU, Rodolphe; MARRAFFINI, Luciano. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol Cell* 2014; 54:234–244. PMID: 24766887. 2014.

BROWN, T. A. Clonagem gênica e análise de DNA. Porto Alegre: Artmed, 2003. 376 p.

BUSZEWSKI, Boguslaw et al. Identification of Microorganisms by Modern Analytical Techniques. *J AOAC Int.* 2017 Nov 1;100(6):1607-1623. doi: 10.5740/jaoacint.17-0207. PMID: 28703095. 2017.

CHANDLER, Heather; COLITZ, Carmen. Molecular Biology for the Clinician: Understanding Current Methods. *Journal of the American Animal Hospital Association*, 42(5), 326–335. doi:10.5326/0420326. 2006.

COSTERTON, JW et al. Mechanism of electrical enhancement of efficacy of antibiotics in killing biofilm bacteria. *Antimicrob Agents Chemother.* 2803-9. doi: 10.1128/AAC.38.12.2803. PMID: 7695266; PMCID: PMC188289. 1994.

DUNFORD, Eric; NEUFELD, Josh. DNA stable-isotope probing (DNA-SIP). *J Vis Exp.* (42):2027. doi: 10.3791/2027. PMID: 20729803; PMCID: PMC3156007. 2010.

BERRY, Duncan; LOY, Alexander. Stable-Isotope Probing of Human and Animal Microbiome Function. *Trends Microbiol.* doi: 10.1016/j.tim.2018.06.004. PMID: 30001854; PMCID: PMC6249988. 2018.

EREN, Kubra; TAKTAKOGLU, Nursema, PIRIM, Ibrahim. DNA Sequencing Methods: From Past to Present. *Eurasian J Med.* doi: 10.5152/eurasianjmed.2022.22280. PMID: 36655445; PMCID: PMC11163357. 2022.

- EVANS, Mark. The polymerase chain reaction and pathology practice. *Diagn Histopathol (Oxf)*. doi: 10.1016/j.mpdhp.2009.04.001. Epub 2009 Jul 3. PMID: 32288788; PMCID: PMC7105915. 2009.
- FUHRMAN, Jed et al. A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences*, v. 105, n. 22, p. 7774-7778. 2008.
- GARNEAU, Josiane et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 2010 (11); 468(7320): 67-71.; 2010.
- GAYON, Jean. From Mendel to epigenetics: History of genetics. *Comptes Rendus Biologies*, (), S1631069116300646—. doi:10.1016/j.crvi.2016.05.009. 2016.
- GLENN, Travis; Field guide to next-generation DNA sequencers. *Mol. Ecol. Resour.*, v. 11, n. 5, p. 759-769. 2011.
- GOVINDARAJAN, Rajeshwa et al. Microarray and its applications. *J Pharm Bioallied (Suppl 2)*:S310-2. doi: 10.4103/0975-7406.100283. PMID: 23066278; PMCID: PMC3467903. 2012.
- GOZZETTI, Alessandro; LE BEAU, Michelle. Fluorescence in situ hybridization: uses and limitations. *Semin Hematol*. 320-33. doi: 10.1016/s0037-1963(00)90013-1. PMID: 11071355. 2000.
- GREEN, Michael; SAMBROOK, Joseph. Analysis of DNA by Agarose Gel Electrophoresis. *Cold Spring Harb Protoc*. doi: 10.1101/pdb.top100388. PMID: 30602561. 2019.
- GU, Jun; SMITH, Janice; DOWLING, Patricia. Fluorescence In Situ Hybridization Probe Validation for Clinical Use. *Methods Mol Biol*;1541:101-118. doi: 10.1007/978-1-4939-6703-2_10. PMID: 27910018. 2017.
- HOUSMAN, D; GUSELLA, J. Application of recombinant DNA techniques to neurogenetic disorders. *Res Publ Assoc Res Nerv Ment Dis*. 1983;60:167-72. PMID: 6296951.
- HILLMANN, Andrew; DUNNE, Eimear; KENNY, Dermot. cDNA amplification by SMART-PCR and suppression subtractive hybridization (SSH)-PCR. *Methods Mol Biol*. 496:223-43. doi: 10.1007/978-1-59745-553-4_15. PMID: 18839114. 2009.
- HERRERA, Federico et al. Standard curve for housekeeping and target genes: specific criteria for selection of loading control in Northern blot analysis. *J Biotechnol*. doi: 10.1016/j.jbiotec.2005.03.001. Epub 2005 Apr 25. PMID: 15925716. 2005.
- ISLAM, Khursheed; IQPAL, Jawed. An Update on Molecular Diagnostics for COVID-19. *Front Cell Infect Microbiol*. doi: 10.3389/fcimb.2020.560616.

- PMID: 33244462; PMCID: PMC7683783. 2020.
- JARCHO, John. Restriction fragment length polymorphism analysis. *Curr Protoc Hum Genet*. Chapter 2:Unit 2.7. doi: 10.1002/0471142905.hg0207s01. PMID: 18428271. 2001.
- JINEK, Martin et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. doi: 10.1126/science.1225829. Epub 2012 Jun 28. PMID: 22745249; PMCID: PMC6286148. 2012.
- JIANG, Chunyang; MENG, Lingxiang; YANG, Binjun; LUO, Xin. Application of CRISPR/Cas9 gene editing technique in the study of cancer treatment. *Clin Genet*. doi: 10.1111/cge.13589. PMID: 31231788. 2020.
- KARUNANATHIE, Harsheni et al. PCR enhancers: Types, mechanisms, and applications in long-range PCR. *Biochimie*. 2022 Jun;197:130-143. doi: 10.1016/j.biochi.2022.02.009. PMID: 35231536. 2022.
- KALLENBERGER, Stefan et al. Correlated receptor transport processes buffer single-cell heterogeneity. *PLoS Comput Biol*. 2017 Sep 25;13(9):e1005779. doi: 10.1371/journal.pcbi.1005779. PMID: 28945754; PMCID: PMC5659801. 2018.
- KERI, Vishakh; CHANDRASEKAR, Pranatharthi. Polymerase chain reaction in the diagnosis of invasive aspergillosis: approaches for appropriate use. *Leukemia & Lymphoma*, 64(7), 1330–1334. 2023.
- KIKHNEY, Judith; MOTHER, Annette. Quality Control in Diagnostic Fluorescence In Situ Hybridization (FISH) in Microbiology. *Methods Mol Biol*. doi: 10.1007/978-1-0716-1115-9_20. PMID: 33576998. 2021.
- KRAMER, Martha; Coen, Donald. Enzymatic amplification of DNA by PCR: standard procedures and optimization. *Curr Protoc Immunol* ;Chapter 10:10.20.1-10.20.10. doi: 10.1002/0471142735.im1020s24. PMID: 18432685. 2001.
- KOVACS, Amir; YACOBY, Keren; GOPHNA, Uri. A systematic assessment of automated ribosomal intergenic spacer analysis (ARISA) as a tool for estimating bacterial richness. *Res Microbiol*. doi: 10.1016/j.resmic.2010.01.006. PMID: 20138144. 2010.
- LAI, Chooi-May. Basic molecular biology techniques. *Methods Mol Med*. doi: 10.1385/1-59259-085-3:1. PMID: 21394576. 2000-
- LEE, Julia et al. Validation of reference genes for quantitative RT-PCR studies of gene expression in perennial ryegrass. *BMC Mol Biol*. doi: 10.1186/1471-2199-11-8. PMID: 20089196; PMCID: PMC2827471. 2010.
- LIN, Na-Sheng et al. Post-embedding in situ hybridization for localization of viral nucleic acid in ultra-thin sections. *J Histochem Cytochem*. doi: 10.1177/41.10.8245409. PMID: 8245409. 1993.

- LIU, L. et al. Comparison of next-generation sequencing systems. *BioMed Research International*, v. 12, p. 80-88, 2012.
- LORENZO-REDONDO, Ramon et al. Molecular epidemiology in the HIV and SARS-CoV-2 pandemics. *Curr Opin HIV AIDS*. doi: 10.1097/COH.0000000000000660. PMID: 33186230; PMCID: PMC7723008. 2021.
- LOUWEN, Rogier et al. The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev*. doi: 10.1128/MMBR.00039-13. PMID: 24600041; PMCID: PMC3957734. 2014.
- MARDIS, Elaine. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*. doi: 10.1146/annurev.genom.9.081307.164359. PMID: 18576944. 2008.
- MARTIN, JO et al. Molecular microbiology: new dimensions for cutaneous biology and wound healing. *J Invest Dermatol*. doi: 10.1038/jid.2009.221. PMID: 19626034. 2010.
- MARZANCOLA, Mahsa; SEDIGHI, Abootaleb; LI, Paul. DNA Microarray-Based Diagnostics. *Methods Mol Biol*. doi: 10.1007/978-1-4939-3136-1_12. PMID: 26614075. 2016.
- MARRAFFINI, Luciano, SONTHEIMER, Erik. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*. doi: 10.1038/nrg2749. PMID: 20125085; PMCID: PMC2928866. 2010.
- MCNICOL, AM; FARQUHARSON, MA. In situ hybridization and its diagnostic applications in pathology. *J Pathol*. doi: 10.1002/(SICI)1096-9896(199707)182:3<250::AID-PATH837>3.0.CO;2-S. PMID: 9349226. 1997.
- MIKIC, Ana; ALOMARI, Arqam, GOWERS, Darren. Classical Recombinant DNA Cloning. *Methods Mol Biol*. doi: 10.1007/978-1-0716-3004-4_1. PMID: 36853452. 2023.
- MOHAMED, Sharmarke; SYED, Basharut. Commercial prospects for genomic sequencing technologies. *Nature reviews Drug discovery*, v. 12, n. 5, p. 341-342. 2013.
- NARAYANAN, Sheshadri. Applications of restriction fragment length polymorphism. *Ann Clin Lab Sci*. PMID: 1677556. 1991.
- NOURI-ARIA, Kayhan. In situ Hybridization. *Methods Mol Med*. doi: 10.1007/978-1-59745-366-0_27. PMID: 18612620. 2008.
- Organização Mundial da Saúde (OMS). Recomendações para estratégias nacionais de teste de SARS-CoV-2 e capacidades de diagnóstico: orientação provisória. 2021. Disponível online: <https://www.who.int/publications/i/item/WHO-2019-nCoV-lab-testing-2021.1-eng>.

- RILEY, Lee; BLANTON, Ronald. Advances in Molecular Epidemiology of Infectious Diseases: Definitions, Approaches, and Scope of the Field. *Microbiol Spectr*. doi: 10.1128/microbiolspec.AME-0001-2018. PMID: 30387413; PMCID: PMC6343655. 2018.
- ROGERS, Beverly. The Evolution of the Polymerase Chain Reaction to Diagnose Childhood Infections. *Pediatr Dev Pathol*. doi: 10.2350/15-05-1643-OA.1. Epub 2015 Dec 23. PMID: 26701384. 2015.
- SCHNEIDER, Tim et al. Northern Blot Analysis of Circular RNAs. *Methods Mol Biol*. 2018;1724:119-133. doi: 10.1007/978-1-4939-7562-4_10. PMID: 29322445. 2018.
- SHENDURE, Jay et al. DNA sequencing at 40: past, present and future. doi: 10.1038/s41586-019-1120-8. PMID: 29019985. 2017.
- STAMM, Stefan; LONGO, Frank. Direct sequencing of PCR products using the Maxam-Gilbert method. *Genet Anal Tech*. doi: 10.1016/0735-0651(90)90021-7. PMID: 2091696. 1990.
- TEMPLETO, Nancy. The polymerase chain reaction. History, methods, and applications. *Diagn Mol Pathol*. doi: 10.1097/00019606-199203000-00008. PMID: 1342955. 1992.
- THUDI, Mahendar et al. Current state-of-art of sequencing technologies for plant genomics research. *Brief Funct Genomics*. doi: 10.1093/bfpg/elr045. PMID: 22345601. 2012.
- TOWBIN, J. Introduction to the techniques of molecular biology. *Herz*. PMID: 8375801. 1993.
- TSUCHIYA, Karen. Fluorescence in situ hybridization. *Clin Lab Med*. doi: 10.1016/j.cll.2011.08.011. PMID: 22118735. 2011.
- VERMA, Mansi; KULSHRESTHA, Samarth; PURI, Ayushi. Genome Sequencing. *Methods Mol Biol*. doi: 10.1007/978-1-4939-6622-6_1. PMID: 27896715. 2017.
- VOYTAS, Daniel. Agarose gel electrophoresis. *Curr Protoc Immunol*. Chapter 10:10.4.1-10.4.8. doi: 10.1002/0471142735.im1004s02. PMID: 18432695. 2001.
- WANG, B; SUN, YF; SONG, N et al. Identification of UV-B-induced microRNAs in wheat. *Genet Mol Res*. doi: 10.4238/2013. PMID: 24114216. 2013.
- WATERS, Daniel; SHAPTER, Frances. The polymerase chain reaction (PCR): general methods. *Methods Mol Biol*. doi: 10.1007/978-1-62703-715-0_7. PMID: 24243196. 2014.
- WALCYK, Thomas; COWARD, Andy; SCHOELLER, Dale et al. Stable isotope techniques in human nutrition research: concerted action is needed. *Food Nutr Bull*. PMID: 12362817. 2002.

- WESTRA, Edze; BUCKLING, Angus; FINERAN, Peter. CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Microbiol.* doi: 10.1038/nrmicro3241.PMID: 24704746. 2014.
- WIEGARD, Jana; SCHLUTER, Amri; BURENINNA, Olga et al. Northern Blot Detection of Tiny RNAs. *Methods Mol Biol.* doi: 10.1007/978-1-0716-1386-3_5. PMID: 33792870. 2021.
- WHITFIELD, John. Genetics and molecular biology in laboratory medicine, 1963–2013" *Clinical Chemistry and Laboratory Medicine (CCLM)*, vol. 51, no. 1, 2013, pp. 113-117. doi: 10.1515/cclm-2012-0478. 2013.
- YANG, Samuel; ROTHMAN, Richard. PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. doi: 10.1016/S1473-3099(04)01044-8. PMID: 15172342; PMCID: PMC7106425. 2004.

7. Agradecimentos

Os autores agradecem o suporte financeiro da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) [PDR10 E–26/204.521/2021], Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil), Sub-Reitoria de Pós-Graduação e Pesquisa da Universidade do Estado do Rio de Janeiro (SR–2/UERJ) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código Financeiro 001.

Autores

João Matheus Sobral Pena, Prescilla Emy Nagao, Pamella Silva Lannes-Costa*

Laboratório de Biologia Molecular e Fisiologia de *Streptococcus*, Departamento de Biologia Celular, Universidade do Estado do Rio de Janeiro (UERJ), Rua São Francisco Xavier, 524, Maracanã, 20550013, Rio de Janeiro, Brasil.

*Autor para correspondência: pam_lannes@hotmail.com

CAPÍTULO 3

Probiogenômica: Análise Genômica dos Probióticos e Seus Benefícios à Saúde

Eduarda Guimarães Sousa, Lucas Gabriel Rodrigues Gomes, Arun Kumar Jaiswal, Laura Diniz Freitas, Rafael Júnio de Oliveira, Letícia Isadora Rodrigues, Luiza Pereira Reis, Aline Ferreira Maciel de Oliveira, Siomar de Castro Soares, Vasco Ariston de Carvalho Azevedo

<https://doi.org/10.69570/mp.978-65-84548-32-9.c3>

Resumo

Segundo a Organização para Alimentação e Agricultura (FAO) e Organização Mundial da Saúde (OMS) a definição de probióticos são microrganismos vivos que, quando administrados em quantidades adequadas, conferem benefícios à saúde do hospedeiro. Para cada cepa probiótica a segurança e eficácia deve ser avaliada, e bem como a análise dos efeitos desses microrganismos na garantia de probiose. Nos últimos anos foi observado um avanço nas tecnologias de sequenciamento e bioinformática, os quais impulsionaram a pesquisa em microbiologia. O desenvolvimento das plataformas de análises *in silico*, que utilizam tecnologias baseadas em ômicas, permitiu caracterizar essas linhagens como potenciais probióticos, bem como identificar fatores genéticos, elucidar os complexos mecanismos que promovem a sobrevivência, adaptação ao trato gastrointestinal e promoção de efeitos benéficos desses microrganismos ao hospedeiro. Os dados genômicos possibilitaram uma análise mais aprofundada sobre os fatores genéticos e moleculares relacionados com os efeitos de probiose como atividades antimicrobianas e antifúngicas, produção de metabólitos secundários, genes que codificam compostos bioativos como vitaminas e genes relacionados com a imunomodulação e produção de citocinas anti-inflamatórias. Análises a nível genômico, associadas com estudos *in vitro* e *in vivo*, vêm sendo utilizadas para a caracterização de novos candidatos probióticos, e fornecendo novas perspectivas sobre os principais fatores associados às características funcionais e de segurança destes microrganismos. Através das análises de probiogenômica é possível analisar e revelar a base molecular sobre a diversidade, interação e evolução de bactérias comensais e probióticas na promoção da saúde do hospedeiro. O objetivo deste capítulo é fornecer uma visão global da Probiogenômica, das suas técnicas e das suas aplicações, explorando os seus fundamentos, avanços e perspectivas futuras.

Palavras-chave: Probiogenômica; Probióticos; Bioinformática; Ômicas

1. Introdução

O estudo de probióticos nos últimos anos se tornaram relevantes na indústria farmacêutica e alimentícia, devido sua alta aplicabilidade e benefícios na prevenção de doenças e melhoria do estado nutricional. A probiogenômica é uma evolução nos estudos de probióticos que são definidos como "microrganismos vivos que, quando administrados em quantidades adequadas, conferem benefícios à saúde do hospedeiro" (HILL et al., 2014), a qual é uma área crescente devido o sequenciamento de nova geração e o aumento de técnicas e programas de bioinformática para analisar molecularmente todos os benefícios e segurança de uma cepa probiótica (SIMONE, 2019). Esse fenômeno gerou um crescente interesse tanto na pesquisa quanto no mercado global, impulsionado pela demanda crescente por produtos que contenham probióticos, sejam eles convencionais ou inovadores. Nesse contexto, surge a análise de novas cepas probióticas de última geração, isoladas de ecossistemas convencionais e não convencionais, com características que possam agregar valor ao mercado (LIN et al., 2019).

Em 2009, a probiogenômica surgiu como uma disciplina focada em estudar a diversidade de bactérias probióticas e entender os mecanismos moleculares por trás de seus benefícios à saúde (VENTURA et al., 2008). A disponibilização de sequências genômicas desses microrganismos ampliou significativamente nosso conhecimento sobre sua biologia. Apesar disso, técnicas microbiológicas tradicionais ainda são amplamente utilizadas para identificar, classificar e quantificar probióticos, embora sejam demoradas e nem sempre precisas, especialmente ao identificar bactérias em nível de espécie ou detectar microrganismos não cultiváveis (LUGLI et al., 2022).

O processo de identificação e avaliação de novas cepas com potencial probiótico evoluiu e se tornou mais complexo. As cepas candidatas precisam ser analisadas por meio de uma combinação de testes *in silico* e *in vitro*, levando em conta critérios funcionais, propriedades tecnológicas e aspectos relacionados à segurança (DE MELO PEREIRA et al., 2018). Os resultados obtidos nestes primeiros testes devem ser confirmados por modelos experimentais em animais e, por fim, validados em ensaios clínicos com seres humanos (SHARMA et al., 2020). Os benefícios dos microrganismos probióticos para a saúde humana são

frequentemente agrupados em categorias como o fortalecimento da barreira intestinal, a regulação da resposta imunológica e o combate a patógenos, seja pela produção de substâncias antimicrobianas ou pela competição por locais de adesão na mucosa (MARCO et al., 2006; O'HARA e SHANAHAN, 2007). Embora haja evidências que sustentem essas funções, os mecanismos moleculares subjacentes ainda não são totalmente compreendidos. Nesse contexto, a probiogenômica desempenha um papel crucial ao fornecer ferramentas avançadas para explorar esses mecanismos em nível genético e molecular, possibilitando uma compreensão mais detalhada dos processos que sustentam os benefícios dos probióticos (VENTURA et al., 2008).

Nos últimos anos, a probiogenômica se firmou como uma ferramenta crucial no estudo dos probióticos. Com o sequenciamento dos genomas de *Bifidobacterium longum* NCC2705 e *Lactobacillus plantarum* WCFS1, os pesquisadores passaram a investigar mais profundamente as características genéticas de microrganismos com potencial probiótico (SANCHEZ et al., 2013; CASTRO-LOPEZ et al., 2021). Atualmente, a base de dados National Center for Biotechnology Information (NCBI) reúne cerca de milhares de sequências de bactérias probióticas tradicionais e de nova geração, oferecendo um vasto conteúdo para pesquisa sobre probióticos, seus mecanismos de probiose e segurança (LEE et al., 2023). O progresso das tecnologias de sequenciamento tornou a probiogenômica um campo essencial, permitindo a identificação de genes e propriedades funcionais das cepas probióticas, além de possibilitar a seleção mais precisa de novos probióticos com base em seu perfil genético (FONTANA et al., 2020).

No contexto da microbiota intestinal, a probiogenômica é uma ótima metodologia para a compreensão da interação entre probiótico e hospedeiro. Ao explorar o perfil genético dos microrganismos presentes no trato intestinal e como esses organismos influenciam a saúde do hospedeiro, a probiogenômica permite identificar como diferentes cepas de probióticos interagem com o organismo e a comunidade, modulando processos fisiológicos e imunológicos (VENTURA et al., 2008; CARVALHO et al., 2022). Essa área também se conecta estreitamente à metagenômica, que é uma ferramenta possível de estudar a composição genética da microbiota intestinal em seu conjunto, permitindo uma

análise do ecossistema microbiano, por tecnologias de sequenciamento de genoma completo ou metagenômica parcial, com o gene 16s por exemplo (CARVALHO et al., 2022; AW e FUKUDA, 2015). Ao integrar essas abordagens, é possível revelar como os probióticos podem afetar a função da microbiota intestinal, promovendo a saúde intestinal, prevenindo doenças e melhorando a resposta imunológica, e como os compartilham informações, por meio de estudos de plasticidade genômica, sintenia gênica e genômica comparativa. Dessa forma, a probiogenômica oferece insights valiosos para o desenvolvimento de terapias mais personalizadas e eficazes, bem como uma compreensão mais aprofundada na interação entre microrganismo e hospedeiro e toda a organização do ecossistema microbiano (DA SILVA et al., 2024).

A probiogenômica é uma área interdisciplinar que combina diversas tecnologias ômicas, como genômica, transcriptômica, proteômica e metabolômica, proporcionando uma compreensão abrangente e aprofundada dos mecanismos pelos quais os probióticos exercem seus efeitos benéficos no organismo. A integração dessas abordagens permite não apenas o mapeamento detalhado das propriedades fisiológicas dos microrganismos, mas também a análise das interações complexas entre esses microrganismos e o hospedeiro, especialmente no ambiente intestinal (KWOJI et al., 2023).

Por meio da genômica, é possível identificar os genes presentes em cepas probióticas, enquanto a transcriptômica permite o estudo dos padrões de expressão gênica, revelando como essas cepas respondem a diferentes estímulos e condições ambientais (YETIMAN et al., 2023; LEE et al., 2023). A proteômica, por sua vez, ajuda a identificar as proteínas produzidas por essas bactérias e suas funções no contexto do hospedeiro, enquanto a metabolômica oferece uma visão detalhada dos metabólitos produzidos, ajudando a entender como esses compostos influenciam a saúde do hospedeiro, incluindo a modulação da microbiota intestinal e o metabolismo (ESCOBAR-SANCHEZ et al., 2023; ECHEGARAY et al., 2023).

Essas ferramentas ômicas, em conjunto, possibilitam a análise não apenas das interações entre os probióticos e o ambiente intestinal, mas também o exame da adaptação dos microrganismos ao hospedeiro e de como eles influenciam funções imunológicas e metabólicas. Além disso, elas são

essenciais para a descoberta de genes relacionados a características desejáveis nos probióticos, como resistência a condições adversas do Trato Gastrointestinal (TGI) e produção de compostos bioativos, ao mesmo tempo em que permitem a identificação de traços indesejáveis, como a produção de substâncias potencialmente prejudiciais (JIANG et al., 2023; ZMORA et al., 2018).

A integração de abordagens microbiológicas clássicas com as novas ferramentas genômicas permite uma melhor caracterização de microrganismos benéficos, o que pode levar ao desenvolvimento de terapias mais eficazes baseadas em probióticos. Além disso, os avanços na pesquisa de microbiota e sequenciamento de nova geração têm um potencial significativo para personalizar o uso de probióticos e fornecer tratamentos específicos para diferentes condições de saúde. Dessa forma, este capítulo abordará a relevância crescente do estudo de probióticos na probiogenômica, destacando oportunidades e desafios enfrentados pelos pesquisadores. Será discutido o estudo multiômico na caracterização de linhagens probióticas, a importância de sua identificação e a avaliação da probiose e segurança dos microrganismos como potenciais probióticos, conforme ilustrado na **Figura 1**.

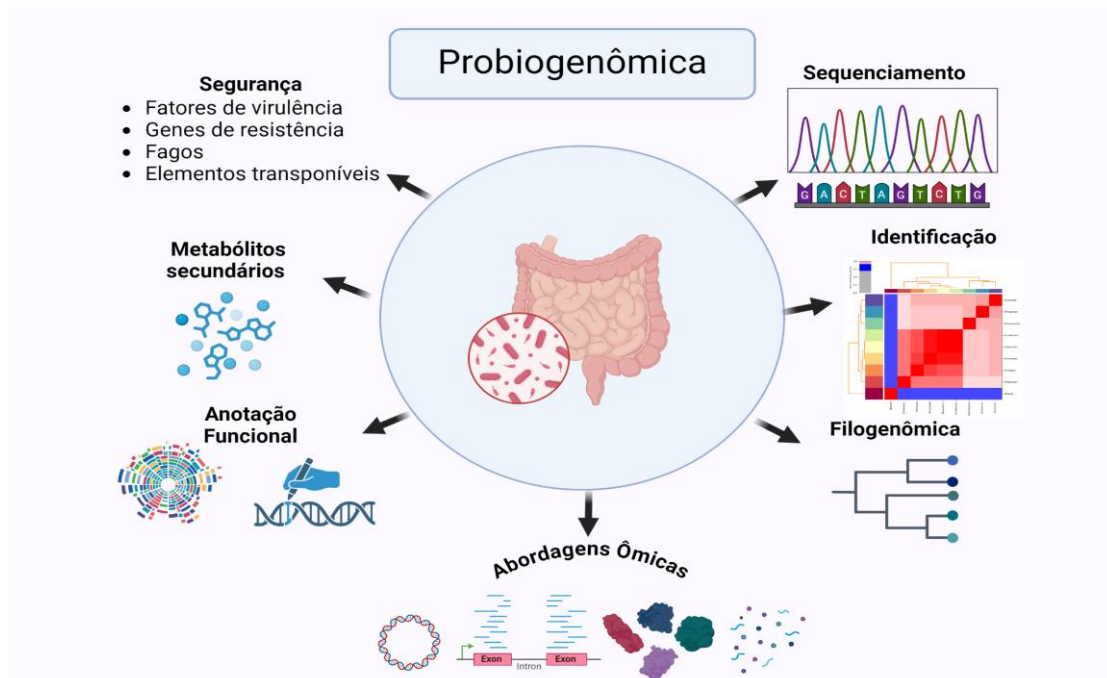


Figura 1. Diagrama abrangente dos principais aspectos relacionados ao estudo de Probiogenômica, destacando as áreas centrais de investigação e as metodologias integradas utilizadas nesse campo interdisciplinar.

2. Abordagens Ômicas na Pesquisa Probióticos

O TGI é um ecossistema complexo com um microbioma consistente de microrganismos em relações que podem variar entre simbiose, comensalidade e patogênese (VENTURA et al., 2012). A interação entre o hospedeiro e sua microbiota intestinal tem efeitos importantes na manutenção da saúde, podendo promover imunidade, garantir a integridade do intestino, melhorar nutrição e defender contra patógenos (YIN et al., 2019). Probióticos, portanto, são definidos como “microrganismos que, quando aplicados em doses adequadas, conferem benefícios à saúde do hospedeiro” através da modulação da microbiota (HILL et al., 2014).

Os mecanismos moleculares dos probióticos são pouco compreendidos e específicos à linhagem e hospedeiro (VENTURA et al., 2012). Sua compreensão requer abordagens que avaliem seus mecanismos em todos os níveis, dos genes à regulação no hospedeiro e microbiota, até a síntese de proteínas e metabólitos (KWOJI et al., 2023). As ciências ômicas, como Genômica, Transcriptômica, Proteômica e Metabolômica (**Figura 2**), são ferramentas importantes para avaliar probióticos nesses níveis.

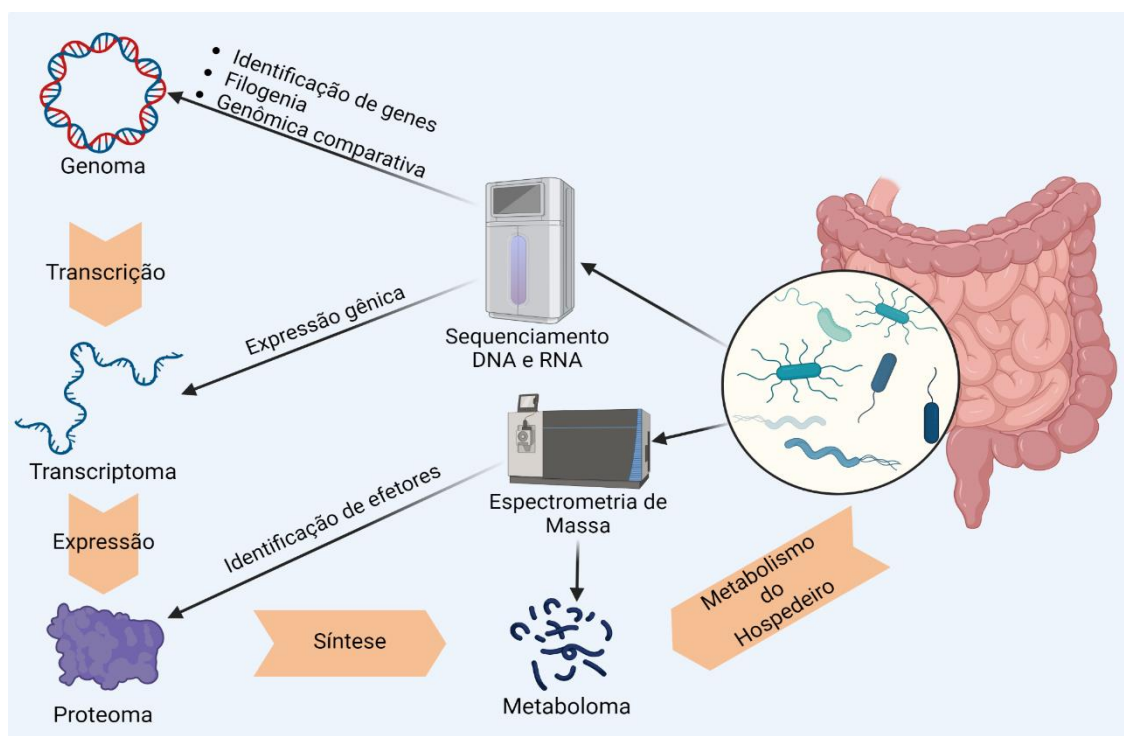


Figura 2. Diagrama das aplicações das ciências ômicas que são chave para a Probiogenômica, destacando as metodologias principais de extração de informação sobre as ômicas dos probióticos e como impactam o campo.

2.1. Genômica

Desenvolvimentos recentes no campo de sequenciamento de genomas têm possibilitado, cada vez mais, o desenvolvimento de robustos bancos de dados como o GenBank, que contém as sequências de uma ampla gama de microrganismos (SAYERS et al., 2023). A análise das sequências codificantes que compõem o genoma de bactérias permite identificar genes relacionados à sua atividade e probiose, prometendo elucidar os fundamentos moleculares por trás de seus benefícios à saúde (VENTURA et al., 2008).

Além da capacidade de sequenciar e caracterizar genes presentes nos genomas de probióticos, a expansão de aplicações da genômica nos permite realizar análises comparativas entre eles. Análises de pangenoma nos permitem caracterizar o repertório gênico de espécies probióticas e avaliar genes compartilhados por grupos de linhagens probióticas, bem como os exclusivos a linhagens de interesse (VENTURA et al., 2008; MAKAROVA et al., 2006; CLAEISSON et al., 2008).

Ainda no campo do sequenciamento de genomas, metodologias amplas para análise de metagenomas também têm tido grande relevância em estudos sobre o TGI. A metagenômica permite avaliar a composição da microbiota e abundância de espécies, bem como avaliar os efeitos da aplicação de probióticos com base na saúde do microbioma MAAKE et al., 2021.

2.2. Transcriptômica

Enquanto a genômica se concentra em sequenciar e analisar os genomas de microrganismos probióticos, a transcriptômica é o estudo do que é transcrito a partir desses genomas. Isto é, a transcriptômica analisa todas as formas de RNA, mensageiro ou não-codificante, bem como suas funções e padrões de transcrição (MILWARD et al., 2016). Isso permite avaliar genes que são transcritos em diferentes condições, como sob estresses ou em contato com outros microrganismos, e as diferenças em expressão entre elas (MILWARD et al., 2016).

Através dessas análises, é possível identificar fatores gênicos envolvidos em sobrevivência a estresses, como por exemplo na passagem dos probióticos

pelo TGI até chegarem na região intestinal (CHEN et al., 2021). A transcriptômica aplicada em comunidades através da metatranscriptômica nos permite avaliar a expressão diferenciada de genes em comunidade com a microbiota e durante processos como a fermentação (EDLUND et al., 2018; SONG et al., 2017).

Indo além da microbiota, a transcriptômica também pode ser usada para medir os efeitos modulatórios dos probióticos nos hospedeiros. Ao avaliar a expressão diferencial de genes em condições de disbiose e com a aplicação de probióticos, é possível avaliar a influência de probióticos em fatores da saúde do hospedeiro (RODRIGUES et al., 2021; KIU et al., 2020).

2.3. Proteômica

A proteômica, por sua vez, é o estudo das proteínas produzidas por uma célula ou tecido através de metodologias capazes de identificar, contabilizar e caracterizar a síntese proteica dos probióticos (RUIZ et al., 2016). Dessa forma, essa disciplina permite a avaliação funcional de efetores de probiose de forma mais específica, uma vez que dados de genoma e transcriptoma não contêm informação precisa de expressão proteica nem de modificações pós-traducionais (KOPONEN et al., 2012).

Em integração com outras análises ômicas, a proteômica permite contabilizar fatores essenciais para a adesão dos probióticos com a mucosa, colonização do TGI, atividade antimicrobiana, dentre outras características funcionais dos probióticos (RUIZ et al., 2016). Adicionalmente, também se podem realizar análises da expressão proteica do hospedeiro e do restante da microbiota através da metaproteômica, permitindo avaliar as interações feitas pelas proteínas dos probióticos (ZHANG et al., 2018). Ao contabilizar e caracterizar proteínas e peptídeos efetores e suas interações com proteínas e tecidos do hospedeiro, é possível analisar como certos probióticos podem modular a saúde humana (ZHANG et al., 2020).

2.4. Metabolômica

Contudo, os efeitos dos probióticos não se dão apenas através de efetores proteicos. A metabolômica é um campo de estudo que se centra na

contabilização e caracterização dos perfis de metabólitos que são fruto de processos biológicos. Esses metabólitos podem vir tanto do hospedeiro quanto da microbiota, sendo eles peptídeos, carboidratos, lipídeos, vitaminas, ácidos orgânicos e outras pequenas moléculas que resultam do metabolismo celular). Probióticos podem, portanto, efetivar mudança no estado de saúde do hospedeiro através da metabolização de compostos no lúmen do TGI e produção de metabólitos secundários bioativos (SALMINEN et al., 2021).

3. Aplicação da probiogenômica para determinação do potencial probiótico

3.1. Identificação molecular dos candidatos à probióticos

A identificação bacteriana é o conjunto de procedimentos utilizados para determinar a espécie ou o grupo ao qual uma determinada bactéria pertence. Esse processo é fundamental tanto para a microbiologia clínica quanto para a pesquisa microbiológica em geral. Ele pode envolver uma variedade de métodos, desde abordagens tradicionais, como a cultura bacteriana em meios específicos e testes bioquímicos, até métodos mais modernos e precisos, como o sequenciamento genético. No contexto do sequenciamento de nova geração (NGS), a identificação bacteriana é baseada principalmente na análise de sequências de DNA extraídas do genoma bacteriano. Essa abordagem molecular permite identificar as características genéticas específicas de uma bactéria, muitas vezes com mais precisão e rapidez do que os métodos tradicionais (WINAND et al., 2019). Demonstrando assim sua eficácia e agilidade nas pesquisas científicas, essas ferramentas baseadas em NGS permitem gerar uma ampla gama de dados que facilitam a classificação dos microrganismos em estudo.

Essas tecnologias têm sido amplamente empregadas em pesquisas sobre potenciais probióticos, possibilitando análises detalhadas de genomas, a identificação de espécies e a avaliação de suas características funcionais. De acordo com o estudo (GARCÍA-HERNÁNDEZ et al., 2016), a determinação das características moleculares de uma bactéria potencialmente probiótica depende, fundamentalmente, do conhecimento da espécie a que ela pertence. Existem diversos métodos de identificação que podem ser aplicados após a montagem

do genoma. Entre eles, destaca-se a análise por Average Nucleotide Identity (ANI) e Hibridização Digital de DNA-DNA, que surgiu da necessidade de um método confiável para a definição de espécies bacterianas, útil para identificar relações evolutivas em cepas com genomas completos ou parciais, sendo um importante programa de bioinformática quando estamos tratando de identificação de cepas bacterianas (LINDSEY et al., 2023; RICHTER e ROSSELLÓ-MÓRA, 2009; AUCH et al., 2010).

O ANI calcula a identidade média de nucleotídeos entre dois genomas bacterianos, comparando as sequências de DNA em regiões conservadas. A alta identidade de ANI (acima de 95%) sugere que as duas amostras pertencem à mesma espécie, enquanto uma baixa identidade (abaixo de 95%) indica que são de espécies diferentes. Por ser uma análise genômica de alta precisão, o ANI oferece uma alternativa eficiente e mais confiável do que métodos tradicionais, como a hibridização do DNA. O pyANI, por exemplo, é uma ferramenta amplamente utilizada para realizar essa análise, sendo útil para identificar relações evolutivas em cepas com genomas completos ou parciais. Essa abordagem tem se consolidado como uma das principais na identificação de cepas bacterianas (PRITCHARD et al., 2015). Outro exemplo é o O GTDB-TK é uma ferramenta de classificação taxonômica, baseada no cálculo de ANI, que utiliza um banco de dados atualizado para atribuir genomas microbianos a uma taxonomia padronizada, melhorando a consistência e precisão na nomenclatura microbiana <https://doi.org/10.1093/bioinformatics/btz848>.

Já o dDDH é uma análise que compara sequências de genomas e mede a similaridade entre dois DNAs pela formação de híbridos e calculando a distância genômica. O dDDH é utilizado para delinear espécies microbianas, baseando-se na quantidade de hibridização entre os fragmentos de DNA. Ferramentas como TYGS são comumente utilizadas para realizar essas análises, para calcular as distâncias genéticas entre pares de genomas. Quando a porcentagem de hibridização é alta (>70%), as cepas são consideradas da mesma espécie, enquanto uma porcentagem baixa (<70%) sugere espécies diferentes (MEIER-KOLTHOFF e GÖKER, 2019).

A anotação de genomas é um processo fundamental na bioinformática que envolve a identificação e a interpretação das funções dos genes presentes

em uma sequência. Esse processo não se limita apenas à localização dos genes, mas também inclui a atribuição de funções, a predição de regiões codificantes e a análise de elementos regulatórios. No contexto dos probióticos, a anotação de genomas desempenha um papel crucial, pois permite a identificação precisa de características genéticas que determinam a viabilidade, a segurança e a eficácia dessas bactérias. Dessa forma, a anotação genômica é uma ferramenta essencial para o desenvolvimento de probióticos mais eficazes e seguros, além de facilitar a compreensão dos mecanismos moleculares envolvidos em seus efeitos terapêuticos (STEIN, 2001; LOBB et al., 2020).

O Prokaryotic Genome Annotation (Prokka), é uma ferramenta de anotação e automática de genomas de procariotos, identificando genes, RNAs e outros elementos genéticos, fornecendo uma visão detalhada das funções e características presentes no genoma (SEEMANN, 2014), a qual integra programas como o BLAST, HMMER, Prodigal e Aragorn, para realizar buscas em bancos de dados de genes conhecidos e prever funções para os novos genes encontrados. No contexto de probióticos, a anotação de genomas com o Prokka tem grande relevância, pois ela permite a identificação de genes-chave que são essenciais para a atividade probiótica das cepas bacterianas.

A anotação funcional é uma etapa importante quando se trata de identificação de possíveis candidatos a probióticos, predizendo seus genes e anotando o seu genoma de forma completa a fim de observar sua funcionalidade, analisar vias metabólicas bacterianas. A qual irá auxiliar na avaliação das propriedades funcionais dos probióticos, como a capacidade de metabolizar substâncias, produzir enzimas benéficas ou até mesmo interagir de forma positiva com a microbiota intestinal, dentre outras especificidades microbianas (RODRIGUES, 2023).

Há diversas ferramentas de bioinformática que impulsionam a pesquisa na probiogenômica a fim de determinar as características, anotações funcionais e moleculares de microrganismos potencialmente probióticos. Ferramentas como o Rapid Annotation using Subsystem Technology (RAST) é um servidor que identifica genes codificadores de proteínas, rRNA e tRNA além de atribuir aos genes suas devidas funcionalidades, podendo assim fornecer informações para construção de vias metabólicas do genoma bacteriano (AZIZ et al., 2008).

Outro tipo de anotação funcional muito utilizada é na predição baseada em *Clusters of Orthologous Groups of Proteins* (COG), pertencente ao NCBI, que é um banco de dados que organiza proteínas de genomas microbianos completos em famílias de genes ortólogos. Cada COG representa um grupo de genes relacionados evolutivamente, permitindo a classificação filogenética e ajudando a identificar funções conservadas entre diferentes organismos. É amplamente utilizado para análises comparativas e anotações funcionais em genômica microbiana (GALPERIN et al., 2019). Já o EggNOG é um banco de dados que fornece dados de ortologia e informações funcionais para microrganismos, oferecendo árvores filogenéticas, permitindo assim estudo evolutivo de espécies além de eventos de duplicações e especiações (HERNÁNDEZ.PLAZA et al., 2023). O KEGG se destaca por ser um banco de dados e software em que podemos fazer uma simulação de comportamentos funcionais como rede de interação de proteínas e reações químicas de organismos a partir de informações genômicas (KANEHISA e SUBRAMANIAM, 2002) através de uma coleção de genes completamente sequenciados com anotações atualizada das funções genéticas (KANEHISA e GOTO, 2000). Esses são apenas alguns exemplos de ferramentas e programas amplamente utilizados no campo da bioinformática para caracterizar possíveis candidatos com potencial probiótico. Entretanto, o desenvolvimento de métodos computacionais e plataformas bioinformáticas para previsão de genes, identificação de cepas microbianas e anotação funcional continua a ser uma vertente crucial de pesquisa em genômica, dada sua relevância para a identificação de isolados bacterianos (GOLYSHEV e KOROTKOV, 2015).

3.2. Avaliação de segurança de probióticos

O status de *Generally Recognised as Safe* (GRAS) é usado pela *Food and Drug Administration* (FDA) para selecionar, por lei, quais substâncias são reconhecidas como seguras e assim podem se tornar aditivos alimentares (FDA GRAS, 2024). Analogamente, o status *Qualified Presumption of Safety* (QPS) é usado pela *European Food Safety Authority* (EFSA), também para regular os aditivos seguros e autorizados para uso alimentar (ROUQUETTE et al., 1998).

Dessa forma, para que uma cepa probiótica possa ser adicionada em alimentos e comercializada, é necessário que ela tenha o status GRAS e QPS, que podem ser obtidos por meio de procedimentos científicos que comprovem a segurança da mesma. Nesse tópico serão abordadas características avaliadas na segurança de probióticos e programas que auxiliam nessa análise probiogenômica.

3.2.1. Fatores de virulência

Os fatores de virulência englobam diferentes características que uma cepa bacteriana pode possuir, como atividade hemolítica, aderência, degradação de mucina, desconjugação de sais biliares, produção de metabólitos tóxicos e indução de quaisquer efeitos adversos no hospedeiro. Todos esses fatores são característicos de bactérias patogênicas, porém as probióticas podem ter algumas dessas características como adaptação ao ambiente intestinal [63]. Dessa forma, cepas probióticas podem ter a presença de algum fator de virulência, mas o mesmo não afetar os benefícios do microrganismo.

3.2.2. Resistência à antibióticos

A resistência à antibióticos podem ser intrínseca, quando a resistência está no cromossomo do organismo, ou adquirida, que acontece quando um microrganismo demonstra resistência a um antimicrobiano ao qual anteriormente era suscetível. Ao realizar a avaliação da segurança de um probiótico é importante investigar a presença ou não de resistência a uma gama dos principais antimicrobianos e, se presente, se ela é intrínseca ou adquirida. Isso porque a resistência intrínseca, por estar em genes não transferíveis, não apresenta risco ao hospedeiro, enquanto a adquirida, por estar em elementos genéticos móveis, pode ser transferida horizontalmente para outros organismos da microbiota do indivíduo (DE JESUS et al., 2024). Dessa forma, essa resistência em probióticos pode trazer um risco à eficiência de tratamentos para infecções, caso a bactéria que o paciente necessita tratar tenha recebido o gene de resistência ao antimicrobiano que está sendo usado (IMPERIAL e IBANA, 2016).

3.2.3. Elementos transponíveis (Fagos e integrases, ilhas genômicas, plasmídeos e crispr-cas)

Um microrganismo pode possuir diferentes tipos de elementos transponíveis, que são sequências de DNA capazes de se mover entre genomas [66]. Fagos são vírus que infectam bactérias e podem integrar material genético heterólogo no seu genoma, o que pode ser perigoso caso esse material genético possua genes de virulência ou resistência antimicrobiana. Integrases são enzimas responsáveis por integrar segmentos diferentes de DNA, causando um maior risco à segurança da bactéria ao integrar genes transmitidos por bacteriófagos, plasmídeos transmitidos horizontalmente com genes de virulência, entre outros (NEMERGUT et al., 2008). Plasmídeos são DNA circular independente do cromossomo do hospedeiro que possuem genes não essenciais (DOGHATHER e GULL, 2019). Por mais que sua presença na cepa não comprove perigo diretamente, se esse material genético possuir genes de resistência ou virulência ele pode facilmente ser transmitido horizontalmente para outros microrganismos da microbiota ou para patógenos ali presentes e assim dificultar o tratamento do paciente (VARANKOVICH et al., 2015).

O sistema CRISPR-Cas é uma ferramenta análoga ao sistema imunológico adaptativo presente em algumas bactérias, que conseguem inativar vírus e plasmídeos que tentem invadi-la (VAN DER OOST et al., 2014). Porém, da mesma forma que os plasmídeos, esse sistema pode ser passado por transferência horizontal e fornecer imunidade a patógenos no organismo do hospedeiro ou para a própria microbiota, acarretando na diminuição da diversidade do microbioma intestinal (MONSHIZADEH et al., 2022). Ilhas genômicas são segmentos de DNA presentes em linhagens ligadas entre si que participam da diversidade e adaptabilidade dos microrganismos (JUHAS et al., 2009). Semelhante aos últimos citados elas podem possuir genes de resistência e virulência, trazendo perigo para outros microrganismos presentes no hospedeiro (JUHAS et al., 2009).

3.2.4. Produção de aminas biogênicas

Aminas biogênicas são compostos orgânicos nitrogenados formados pelos processos de descarboxilação de aminoácidos ou animação e

transaminação de cetonas e aldeídos. Alguns exemplos das mesmas são histamina, cadaverina, tiramina, putrescina e triptamina (WÓJCIK et al., 2021). Esses compostos, em baixas concentrações, não apresentam risco, mas em níveis mais altos podem ser tóxicos (CASTRO-LÓPEZ et al., 2021), causando sintomas como hipertensão, dores de cabeça, vômito, problemas respiratórios, entre outros (EFSA, 2011).

Não existe uma regulação que especifique os limites da quantidade de aminas biogênicas que podem estar presentes nos alimentos, somente para a histamina em produtos com peixes (REGULATION EU, 2013). Dessa forma, é importante o uso da pesquisa para verificar se o microrganismo probiótico produz ou não aminas biogênicas e, se sim, verificar se ela causaria sintomas no paciente, para assim garantir a segurança da mesma.

3.2.5. Ferramentas bioinformáticas que participam da avaliação da segurança

Todos os elementos citados nos tópicos anteriores impactam de alguma forma na avaliação da segurança de uma cepa probiótica e por isso é importante o uso de ferramentas bioinformáticas para auxiliar a análise desses pontos (PENG et al., 2023) para melhor compreender os mecanismos moleculares e genéticos associados na caracterização de linhagens probióticas em relação a sua segurança, na **Tabela 1** está sumarizado as principais ferramentas de bioinformática utilizadas para avaliação desses fatores.

Tabela 1. Principais ferramentas de Bioinformática utilizadas para análise de segurança para a caracterização de microrganismos probióticos.

Ferramenta/banco de dados/programa	Metodologia	Artigo/documentação
VFAnalyzer/VFDB	Plataforma com banco de dados próprio que constrói grupos ortólogos nos genomas de referência e no genoma inserido para análise, realiza um BLAST para verificar os fatores de virulência.	(EFSA, 2011)

PanViTa	Plataforma que usa o banco de dados VFDB para identificar genes de resistência e virulência através do DIAMOND	(RODRIGUES et al., 2023)
Abriicate	Compara a sequência com diversos bancos de dados (como NCBI, CARD, VFDB e Resfinder) a fim de verificar resistência antimicrobiana ou virulência.	(GITHUB, 2024)
VFARs	Ferramenta que usa de diversos bancos de dados e compara os dados genômicos, proteômicos, metabolômicos e transcriptômicos do microrganismo, a fim de visualizar fatores de virulência ou resistência a antimicrobianos.	WASSEEM et al., 2017)
SPAAN	É um modelo de rede neural validado por bancos de dados que calcula vetores relacionados a atributos de sequências de aminoácidos (frequência de aminoácidos, composição de carga, frequência de dipeptídeos, composição hidrofóbica e frequência de múltiplos) e compara a proteína inserida com padrões desses vetores, resultando na probabilidade da proteína ser uma adesina.	(SACHDEVA et al., 2005)
MAAP	Técnica que usa de primers oligonucleotídicos para amplificar regiões do DNA e gerar perfis característicos, que podem ser usados para identificação molecular do microrganismo, mapeamento genético, detecção de DNAs polimórficos, etc.	CAETANO.ANOLLÉS, 1994)
PathoFact	A ferramenta combina um banco de dados de perfis Hidden Markov Model (HMM) e um modelo de aprendizado baseado em Random Forest para comparar se a sequência inserida é de fatores de virulência ou não.	(DE NIES et al., 2021)
DeepVF	Software que identifica fatores de virulência (VFs) em bactérias Gram-negativas. Ele usa uma abordagem chamada "stacking" (ou "empilhamentos"), que combina vários modelos de aprendizado de máquina e deep learning para melhorar a precisão da previsão.	(XIE et al., 2021)
VF-Pred	A ferramenta usa de recursos extraídos como SeqSim e Seg-Alignment para alimentar modelos clássicos de machine learning, que então gera a	(SINGH et al., 2024)

	probabilidade da variante ter ou não alto potencial funcional de virulência.	
PRAP	Pipeline desenvolvido para rápida identificação de genes de resistência a antibióticos a partir de vários formatos de sequências do genoma completos com base no banco de dados CARD ou Resfinder.	(HE et al., 2020)
ResFinder	Banco de dados que usa BLAST ou KMA para alinhar as sequências gênicas inseridas com as sequências de resistência antimicrobiana do próprio banco de dados.	(FLORENSA et al., 2022)
CARD	Banco de dados que combina dados de genes e mutações ligadas a resistência a antimicrobianos.	(ALCOCK et al., 2023)
PATRIC	Plataforma online que combina dados das ômicas, como proteômica, metabolômica. Além de oferecer ferramentas para estudos de filogenia, detecção de resistência a antibióticos e análise comparativa.	(WATTAM et al., 2014)
RAST	Serviço de anotação totalmente automatizado, para genomas arqueanos ou bacterianos. O serviço busca produzir de forma rápida as funções genéticas e uma reconstrução das vias metabólicas iniciais.	(AZIZ et al., 2008)
AMRFINDER	Identifica genes de resistência de genes antimicrobianos. Ele combina buscas BLAST e HMM para identificar genes de resistência	(FELDGARDEN et al., 2019)
MEGARes	Banco de dados de resistência antimicrobiana são adaptados para perfis funcionais de genes em menor escala usando anotações altamente descritivas.	(LAKIN et al., 2017)
VAMPr	Ferramenta que combina grandes volumes de dados genômicos com análise estatística e computacional para melhorar o entendimento e a previsão da resistência a antibióticos.	(KIM et al., 2020)
PHASTER	Ferramenta que realiza um BLAST com bancos de dados de profagos, agrupa genes semelhantes a fagos e pontua cada região de acordo com a quantidade de genes relacionados a fagos.	(ARNDT et al., 2016)

PHASTEST	Ferramenta que usa o software Prodigal para identificar regiões codificantes de proteínas, compara as sequências com bancos de dado e faz um BLAST para determinar as regiões de fago.	(WISHART et al., 2023)
PhaTYP	Método que usa da arquitetura Transformer, identifica as sequências de fago e alinha usando BLAST. O modelo de linguagem Masked LM é treinado pra prever os tokens de proteínas e refinar o desempenho da ferramenta.	(SHANG et al., 2023)
PhaBOX	Servidor web que utiliza o algoritmo PhaMer para identificar contigs que representam profagos, um módulo PhaTYP para prever se o fago é de estilo virulento ou não, e então classifica o mesmo com a combinação das características de similaridade da sequência e as redes convolucionais usando o PhaGCN.	(SHANG et al., 2023)
VirSorter	Ferramenta que usa dos bancos de dados RefSeqABVir e Viromes e os compara com as proteínas por meio de BLAST, para então categorizar as regiões de profago.	(ROUX et al., 2015)
VIBRANT	Ferramenta que utiliza de rede neural treinado com métricas que incluem proporção de proteínas associadas a vírus, assim ele se baseia em pontos de similaridade para identificar e analisar fagos em microrganismos.	(KIEFT et al., 2020)
PlasmidFinder	Ferramenta que identifica plasmídeos em dados de sequenciamento genômico. Ele utiliza uma base de dados com 116 sequências de replicons, representando grupos de incompatibilidade conhecidos. A análise é feita comparando as sequências carregadas com a base usando o algoritmo BLASTn, com critérios mínimos de 80% de identidade e 60% de cobertura.	(CARATTOLI et al., 2020)
MOB-suite	É um conjunto de ferramentas genômicas que permite identificar, reconstruir e classificar plasmídeos com base em suas sequências genômicas.	(ROBERTSON, 2018)

GIPSy	Software que prevê ilhas genômicas de bactérias e archaea a partir de seu desvio de assinatura, podendo resultar na identificação de ilhas de: patogenicidade, virulência, metabólica, resistência e simbiose. Evolução do Pipsy.	(SOARES et al., 2016)
CRISPRFinder	Software que pode filtrar genomas de bactérias e archaea a partir de seu domínio conservado, e identificar e classificar os loci CRISPR-Cas de alta probabilidade, virulência, resistência, recombinação.	(GRISSA et al., 2007)

Os fatores de virulência podem ser analisados pela plataforma VFAnalyzer, que utiliza do seu próprio banco de dados para construir grupos ortólogos nos genomas inseridos e os de referência e realiza um BLAST para filtrar os parálogos, dessa forma, detectando os fatores de virulência da cepa (LIU et al., 2019). O SPAAN é um modelo de rede neural que calcula os atributos de sequências de aminoácidos como vetores e então compara a proteína inserida com os fatores frequência de aminoácidos, de dipeptídeos e de múltiplos, composição hidrofóbica e composição de carga, informando a chance da cadeia inserida ser uma adesina (SACHDEVA et al., 2005).

Os fatores de resistência antimicrobiana podem ser analisados utilizando o banco de dados CARD, que combina diversos dados de genes de resistência (ALCOCK et al., 2023), ou o ResFinder, que utiliza BLAST ou KMA para alinhar as sequência gênicas de referência da própria *database* com a inserida para análise (FLORENSA et al., 2022). Há também ferramentas que conseguem analisar virulência e resistência ao mesmo tempo. É o exemplo do PanViTa, que é uma plataforma que realiza um BLAST com GenBank e identifica os fatores [78], e o Abricate, que faz uma triagem de contigs comparando a sequência com vários bancos de dados (GITHUB, 2024).

A verificação da presença de plasmídeo no genoma bacteriano pode ser visualizada pelas ferramentas PlasmidFinder, que usa critérios de identidade e cobertura para realizar um BLAST com a sequência (CARATTOLI e HASMAN, 2020), e MOB-suite, que é a junção de várias ferramentas que identificam, reconstroem e classificam plasmídeos (ROBERTSON e NASH, 2018). Já o

PHASTER é usado para predição de profagos, com uma metodologia semelhante a outros em que ele realiza um BLAST com o *database*, agrupa genes de semelhança à fagos e pontua-os de acordo com a quantidade desses genes (ARNDT et al., 2016).

O PIPSy e GIPSy são preditores semelhantes de ilhas genômicas de bactérias. A grande diferença entre os softwares é a quantidade de ilhas genômicas que conseguem predizer, sendo que o GIPSy possui uma ampla gama de preditores resultando na identificação de cinco ilhas genômicas (de patogenicidade, de resistência, de virulência, metabólica e de simbiose) (SOARES et al., 2016), enquanto o PIPS possui apenas um preditor, podendo identificar apenas a ilha de virulência (SOARES et al., 2012). Por fim, o CRISPRFinder é um software que filtra genomas e identifica e classifica os loci com alta chance de ser Crispr-cas (GRISSA et al., 2007).

3.3. Avaliação dos mecanismos e propriedades de efeito de probiose

Os microrganismos também apresentam mecanismos de resposta ao estresse, conjuntos de adaptações moleculares, fisiológicas e bioquímicas para lidar com ambientes hostis ou que ameacem sua homeostase. As bactérias podem apresentar diversos mecanismos como resposta ao estresse, os quais são essenciais para a sobrevivência em ambientes adversos (BOOR et al., 2006). Esses mecanismos vão ser ativados quando ocorrer alguma mudança no ambiente, podendo ser mudança de pH, temperatura, nutrientes disponíveis, estresse oxidativo e outros fatores como exposição a antibióticos (ÇAKAR e SAKA, 2024).

Para as bactérias probióticas a resposta ao estresse é algo positivo, por garantir que estas sobrevivam, além de ser funcional e eficaz. Como as bactérias probióticas atuam no intestino elas precisam sobreviver por todo o TGI, um ambiente que passa por mudanças de pH, temperatura e salinidade, ação de enzimas que atuam na digestão, e outras mudanças físico-químicas (BOOR, 2006). Além disso, as bactérias podem competir com outros microrganismos, com a resposta ao estresse podendo conferir vantagens competitivas sobre os outros microrganismos, como patógenos (CASTRO-LÓPEZ et al., 2022).

Outrossim, algumas bactérias probióticas têm seus benefícios relacionados aos mecanismos de resposta ao estresse (KATHIRIYA, 2023).

Através da probiogenômica é possível identificar os genes e os reguladores gênicos que estão associados à resistência ao estresse. Essas ferramentas são muito importantes para a avaliação de possíveis cepas probióticas, além de entender os mecanismos de estresse que determinada bactéria vai ter (BUSTOS et al., 2024). Os mecanismos de estresse são essenciais para a eficiência das bactérias probióticas como uma ferramenta terapêutica, além de garantir que elas sobrevivam na produção de alimentos funcionais. Além dos mecanismos permitirem que as bactérias sobrevivam, potencializa os efeitos para o hospedeiro (BUSTOS et al., 2024).

3.3.1. Atividade antagonista

A atividade antagonista é a capacidade que algumas bactérias têm de inibir ou impedir o crescimento de outros microrganismos, incluindo patógenos. Existem diferentes mecanismos de atividade antagonista, como produção de compostos antimicrobianos, onde podem ser produzidas bacteriocinas, peróxido de hidrogênio e ácido orgânico, estas substâncias vão atuar em outros microrganismos ou criando um ambiente hostil (ZHU et al., 2023). Outros mecanismos são a competição por recursos, modulação do sistema imunológico e a produção de biosurfactantes, esses mecanismos ajudam a inibir a adesão e colonização de possíveis hospedeiros (BERMUDEZ-BRITO et al., 2012). Através da probiogenômica, conseguimos identificar os genes que atuam para promover a inibição ou eliminação de outros microrganismos.

Para bactérias probióticas a atividade antagonista é uma grande vantagem, isso porque inibindo o crescimento de outras bactérias vai garantir que não tenha competição, facilitando o seu crescimento (BERMUDEZ-BRITO et al., 2012). Esse mecanismo é importante para manter a saúde intestinal e ajudar na prevenção de infecções. Além disso, algumas substâncias antimicrobianas são usadas no combate a doenças, como os antibióticos, essas substâncias também podem ser usadas para aumentar a vida dos alimentos, sendo importante na produção de alimentos funcionais (THUY et al., 2024).

3.3.2. Habilidade de adesão

Uma das características mais importantes para as bactérias probióticas é a habilidade de adesão, uma vez que precisam se fixar ao epitélio do intestino (GORREJA e WALKER, 2022). A habilidade de adesão é a capacidade das bactérias de se fixarem a superfícies biológicas, um processo mediado por interações físico-químicas e por moléculas e estruturas próprias das bactérias. A adesão no intestino é importante para a persistência das bactérias no mesmo, prolongando os efeitos delas, e sem essa capacidade seriam rapidamente eliminadas. Sendo assim, a habilidade de adesão é essencial para a eficácia das bactérias probióticas (SAVAGE e FLETCHER, 1985).

Existem diversos benefícios relacionados à habilidade de adesão, entre elas pode-se citar a exclusão de patógenos por competição, já que se uma bactéria probiótica estiver aderida no intestino, outra bactéria, com um patógeno, não vai conseguir aderir aquele local (PECORARO et al., 2023). Outros benefícios são a interação com o sistema imunológico, além de exercer uma influência positiva no equilíbrio da microbiota nativa, promove a integração da barreira intestinal, se ela tiver comprometida (PERA e PIETERS, 2014).

3.3.3. Imunomodulação

A imunomodulação é alguma alteração no sistema imunológico ou alteração de uma resposta imunológica, essa modulação é mediada por agentes que podem ou ativar ou reprimir alguma função (STRZELEC et al., 2023). Entre os agentes que têm capacidade de imunomodulação estão as bactérias probióticas, elas têm a capacidade de interagir com o sistema imunológico do hospedeiro, modulando as respostas inflamatórias e fazendo a regulação de citocinas (MALDONADO GALDEANO et al., 2007). Essa capacidade de bactérias probióticas de modular o sistema imune é essencial e benéfica para garantir os efeitos positivos dos probióticos para o hospedeiro.

Existem diferentes mecanismos de imunomodulação através dos quais as bactérias podem afetar seus hospedeiros. Entre eles, pode-se citar a interação com o epitélio do intestino, em que as bactérias vão interagir com as células do intestino através de receptores, essa interação pode estimular a produção de peptídeos antimicrobianos e fortalecer a barreira intestinal (CRISTOFORI et al.,

2021). Outros mecanismos são a contribuição para ativar as células dendríticas, produção de citocinas e modulação da resposta imune através de linfócitos T e secreção de IgA (SCHAFER et al., 2024). Através desses mecanismos as bactérias conseguem estimular o sistema imune a respostas que são positivas para o hospedeiro, primeiro a tolerância a bactéria probiótica, fazendo com que o organismo não tente eliminá-la, também secretando citocinas e outras substâncias anti-inflamatórias (SINGH et al., 2023).

Sendo assim, a imunomodulação é muito importante, não só para ajudar a bactéria a passar pela barreira do sistema imunológico, mas também pelos benefícios que ela pode trazer para o hospedeiro (TSAI et al., 2012). Além disso, a imunomodulação pelas bactérias benéficas pode trazer diversos benefícios para saúde, como redução de infecções, propriedades anti-inflamatórias, além de contribuir para a melhora da microbiota intestinal (STRZELEC et al., 2023).

3.3.4. Produção de metabólitos secundários

A produção de metabólitos secundários é uma característica em que se tem a produção de compostos orgânicos que não são essenciais para o crescimento, desenvolvimento e reprodução de organismos (GRIESEMER e NAVID, 2023). Essa característica é importante porque, frequentemente possuem propriedades bioativas que vão contribuir para o hospedeiro e para o ambiente microbiológico (YAN et al., 2022). Esses metabólitos são importantes porque podem trazer diversos benefícios, como atividade antibiótica, mediar as relações entre os organismos, proteger contra estresses abióticos e bióticos e proteção contra competidores (FU et al., 2017). Os metabólitos secundários podem ser classificados de acordo com a sua composição química, sendo divididos em Compostos nitrogenados; Compostos fenólicos; Glicosídeos; Taninos; Terpenos; Flavonoides (THIRUMURUGAN et al., 2018).

Alguns exemplos de metabólitos secundários são: as bacteriocinas, inibindo o crescimento de outras bactérias; Ácidos Graxos de Cadeia Curta (SCFAs), pode-se citar exemplos como acetato, butirato, propionato, eles vão contribuir com a regulação do pH e estimulação de células imunológicas (MASSE e LU, 2023). Exopolissacarídeos (EPS), é outro tipo de metabólito, são polímeros

que contribuem para a adesão ao epitélio intestinal, formação de biofilmes e modulação imunológica (NETRUSOV et al., 2023). Por fim, tem algumas vitaminas essenciais que podem ser produzidas por bactérias como biotina e vitamina K. Sendo assim, é possível ver os benefícios que os metabólitos secundários podem trazer para o hospedeiro (GU e LI, 2016).

O programa BAGEL4 (VAN HEEL et al., 2018) é uma ferramenta que é usada para identificar e visualizar clusters gênicos em genomas de procariotos envolvidos na produção de peptídeos sintetizados ribossomalmente e modificados pós-traducionalmente (RiPPs) e bacteriocinas. A metodologia do programa consiste em uma entrada de dados, através de um arquivo com as bases, após isso é feita a tradução e identificação de áreas de interesse (AOIs), em que o DNA vai ser traduzido em quadros de leitura para usar os códons de iniciação, depois se compara as proteínas com um banco de dados e através disso são definidas as áreas de interesse. Após isso vai ser feita uma análise detalhada das AOIs e feito um relatório gráfico e algumas informações como modificações pós-traducionais são incorporadas a partir de banco de dados. Pode ter recursos adicionais, como banco de dados de outras fontes, BLAST e visualização de RNA-Seq.

Já o antiSMASH (BLIN et al., 2023) é uma ferramenta usada para identificar novos compostos naturais com potencial bioativo, que vão ter diversas aplicações como farmacêuticas, controle agrícola e exploração da biodiversidade química microbiana. A metodologia dele consiste em: entrada dos dados, carregando os genomas dos microrganismos que vão ser analisados, após isso o programa vai utilizar um conjunto de regras baseadas em modelos ocultos de Markov (pHMMs) para identificar funções biossintéticas centrais que caracterizam os BGCs. Com isso, ele vai analisar as classes específicas de BGCs, focando em classes específicas como Sintetases de peptídeos não ribossomais (NRPSs) ou Policetídeos tipo I e II (PKSs), depois vai fazer uma previsão da estrutura química dos produtos, e predição de Regulação e Sítios de Ligação de Fatores de Transcrição. Após isso ele vai comparar os peptídeos centrais de precursores identificados com bancos de dados como o antiSMASH-DB e o MIBiG, diferenciando compostos conhecidos de novos. Os resultados vão sair em um gráfico detalhado e em uma tabela interativa.

4. Probiogenômica na Prática: Exemplos de Estudos

A probiogenômica apresenta vários exemplos práticos e aplicáveis que modificaram a forma de realizar investigação nesta área, as quais podem ser demonstradas na **Tabela 2**.

Tabela 2. Métodos de para Análises de Probióticos Baseadas em Abordagem Genômicas

Ferramentas de Bioinformática	Técnicas de Laboratório de Bancada
Análise de Sequenciamento de Nova Geração (NGS)PCR (Reação em Cadeia da Polimerase): sequenciamento rápido de grandes volumes de DNA - identificação de genes e mutações em genomas completos.	PCR (Reação em Cadeia da Polimerase): Técnica para amplificação de pequenas quantidades de DNA para a detecção e análise de genes específicos.
Anotação Genômica: identificação e marcação de genes em uma sequência de DNA por meio de bancos de dados e algoritmos.	Clonagem e Sequenciamento de DNA: Processo de inserção de fragmentos de DNA em vetores plasmidiais para replicação e sequenciamento - análise detalhada de genes.
Predição de Operons: Ferramentas que analisam sequências para identificar grupos de genes transcritos como um único mRNA.	Northern Blot: Técnica de detecção e quantificação de RNA mensageiro (mRNA) específico - indica a expressão de genes.
Modelagem Molecular: simulações computacionais de predição de estrutura tridimensional de proteínas e suas interações com outras moléculas.	Western Blot: Técnica de detecção de proteínas específicas em uma amostra - confirma a expressão de genes a nível proteico.
Análise de Microbioma: Estuda a composição e função de comunidades microbianas em um ambiente específico por meio do sequenciamento de DNA.	Ensaio de Atividade Enzimática: Testes laboratoriais que medem a atividade de enzimas - confirmam a funcionalidade das proteínas codificadas pelos genes.
Comparação Genômica: Compara genomas de diferentes organismos para entender evolução e função de genes - busca identificar similaridades e diferenças.	Mutagenese e Complementação: Técnicas de inativação e reativação de genes específicos para estudar sua função com cópias funcionais dos genes.
Análise de Expressão Gênica: Compara dados de sequenciamento de RNA para medir a expressão gênica em diferentes condições - regulação gênica.	Ensaio de Imunofluorescência: Método que usa anticorpos marcados com fluorescência para detecção de proteínas específicas em células ou tecidos.
Identificação de Clusters de Genes: Ferramentas de análise de sequências para encontrar regiões com múltiplos genes relacionados agrupados.	Análise de Proteômica: Usa técnicas como espectrometria de massa para identificação e quantificação de proteínas em uma amostra - dados de expressão e função das proteínas.

A genômica aplicada ao estudo dos probióticos tornou-se essencial na avaliação da utilização destas substâncias e no desenvolvimento de novas, pois a sua capacidade de identificar cepas ao nível de espécie, munidas desta atividade, de forma mais precisa, possibilitou a realização de estudos mais assertivos, por exemplo, através de métodos como o sequenciamento, que permitiu o reconhecimento do gene 16S rRNA e a aplicação de genes com maior potencial discriminatório, como os codificadores de proteínas de cópia única (pheS) (SANDERS et al., 2010).

Deste modo, as abordagens probióticas baseadas na genômica proporcionaram muitos avanços na pesquisa de cepas específicas de *Lactobacillus*, permitindo a identificação de genes importantes para a adaptação destes microrganismos ao ambiente característico da mucosa intestinal, bem como para a sua capacidade de produzir substâncias probióticas. Além disso, possibilitaram a seleção de propriedades de interesse específicas (LEBEER et al., 2008). Por exemplo, cepas que possuem maior resistência a estresses oxidativos, osmóticos e térmicos, como é o caso de *Lactiplantibacillus plantarum* (BUSTOS et al., 2024; NYABAKO et al., 2020), ou cepas que apresentam capacidade de adesão à mucosa intestinal, garantindo a sobrevivência e efeitos benéficos para a saúde do hospedeiro, como em casos de *Lactobacillus acidophilus* (KRAUSOVA et al., 2019).

L. plantarum é uma cepa muito destacada na literatura, servindo como um bom exemplo da aplicabilidade de abordagens genômicas no desenvolvimento de probióticos. Esta cepa apresenta inúmeras propriedades de interesse capazes de contribuir para a promoção da saúde, determinadas por meio do sequenciamento do seu genoma completo possibilitando identificar a presença de peptídeos antimicrobianos e bacteriocinas, que atualmente são utilizados como conservantes alimentares e coadjuvantes no tratamento com antibióticos, contribuindo para a saúde única (BAUTISTA e BARRADO et al., 2023).

Outras contribuições de análises genômicas estão relacionadas principalmente com a segurança da utilização destes probióticos, dado que estes microrganismos são capazes de desenvolver resistência. De facto, o *L. plantarum* possui uma acidez e bile que facilitam a sua colonização no TGI (HAMON et al., 2011). A fim de garantir essa segurança, são imprescindíveis

análises genômicas para confirmar a ausência de genes de virulência e identificar esses genes de resistência, bem como para verificar se a bactéria tem capacidade de transmitir essa resistência a outras bactérias patogênicas. Na área da saúde, são necessários testes experimentais específicos, como a PCR, para detectar a presença destes genes utilizando certos primers (SEDDIK et al., 2017).

Os *Lacticaseibacillus rhamnosus* são bactérias com potencial probiótico que apresentam operons específicos identificados por meio de análises computacionais de bioinformática, como o spaCBA-srtC1 e spaFED-srtC2, responsáveis por regular certos processos biológicos, bem como clusters de genes que codificam proteínas secretoras de bacteriocinas, como Bact1 e Bact2, e genes de imunidade que protegem a bactéria da sua própria bacteriocina (DUTRA-SILVA et al., 2023). O conhecimento do código genético de qualquer ser vivo e das suas espécies, bem como a dominância de técnicas laboratoriais como a PCR, a clonagem, a CRISPR-CAS, entre outras, possibilitou que os investigadores pudessem mapear a distribuição de genes e reconhecer padrões intra-espécies, bem como promover uma maior compreensão dos mecanismos de ação dos probióticos e do seu comportamento em ambientes específicos onde desempenham funções benéficas no organismo do hospedeiro (DUTRA-SILVA et al., 2023).

A análise de cepas de *Bacillus clausii* e *Bacillus coagulans* é um exemplo notável da aplicabilidade de estudos de probióticos baseados em abordagens genômicas, dado que revelaram a presença de genes de resistência a antibióticos e de produção de enzimas antioxidantes. Estes genes são essenciais para a adaptação das cepas de *Bacillus clausii* a ambientes intestinais adversos e para a sua proteção contra o stress oxidativo. Além disso, foram encontrados genes produtores de enzimas digestivas e de vitaminas, como a B12. Por outro lado, foi também indicada a ausência de genes patogênicos (SADRIMOVAHED et al., 2024). Em relação ao *Bacillus coagulans*, esta bactéria é uma enorme fonte de riboflavina (vitamina B12) e apresenta proteínas essenciais para a esporulação que estão conservadas em todas as suas cepas probióticas. Além disso, apresenta vários domínios funcionais conservados associados à resistência a antibióticos, o que constitui características com um

potencial significativo e diferenciador para melhorar a eficácia e a aplicabilidade dos probióticos de *Bacillus* em comparação com os de *Lactobacillus* (KHATRI et al., 2016).

A união de todas essas informações permitiu maximizar os benefícios dos produtos comerciais que contêm cepas das duas espécies de *Bacillus* citadas, produzindo formulações com menores riscos de efeitos adversos, graças aos testes laboratoriais que comprovam estes benefícios na prática. Por exemplo, são realizados experimentos que simulam condições intestinais adversas com isolados bacterianos, a fim de confirmar a capacidade dessas cepas de aderir ao muco intestinal e sobreviver, reafirmando a eficácia destes probióticos no TGI (MAZZANTINI et al., 2022).

Além disso, abordagens genômicas têm aplicações clínicas avançadas como o desenvolvimento de probióticos personalizados (CUNNINGHAM et al., 2021). A análise do genoma bacteriano associado ao microbioma de um indivíduo permite a identificação de cepas específicas mais adequadas às suas necessidades metabólicas e imunológicas (KESHET e SEGAL, 2024). Por exemplo, estudos recentes podem explorar como a integração de conjuntos de dados genômicos e metabolômicos pode ajudar na formulação de probióticos direcionados para condições específicas, como síndrome do intestino irritável, doenças inflamatórias intestinais e até mesmo distúrbios metabólicos como obesidade e diabetes (PUIG-CASTELLVÍ et al., 2023; ZHAO et al., 2023). Outra técnica promissora atualmente é o uso da tecnologia CRISPR-Cas para modificar geneticamente cepas probióticas, aumentando sua funcionalidade ou segurança. Essa abordagem já foi usada para otimizar a produção de bacteriocinas, substâncias antimicrobianas naturais que podem servir como alternativas aos antibióticos (MU et al., 2022). A engenharia de probióticos com propriedades imunomoduladoras também está surgindo como um foco significativo, especialmente para o tratamento ou prevenção de condições inflamatórias crônicas (MU et al., 2022).

Perspectivas futuras da probiogenômica podem ajudar a incorporar a aplicação de técnicas multiômicas, combinando genômica, transcriptômicas, proteômica e metabolômica para fornecer uma compreensão holística do impacto dos probióticos na saúde do hospedeiro. Esses avanços podem levar à

criação de formulações probióticas altamente personalizadas e específicas, minimizando riscos e maximizando benefícios.

5. Conclusão

Neste capítulo, foram explorados os principais tópicos relacionados ao estudo de probióticos utilizando a metodologia de probiogenômica, uma abordagem interdisciplinar que combina testes *in vitro*, *in vivo* e *in silico* para caracterizar e avaliar microrganismos tradicionais e de nova geração como potenciais probióticos. Essa metodologia integra análises multiômicas, abrangendo desde genômica até o estudo de metabólitos secundários, para compreender, em nível genético e molecular, a complexa relação simbiótica entre probiótico e hospedeiro. Tal compreensão é essencial para elucidar como essa interação promove a saúde do hospedeiro, contribuindo para a prevenção de doenças, modulação imunológica e controle da microbiota intestinal por meio de mecanismos antagonistas.

O capítulo aprofundou as abordagens ômicas empregadas na pesquisa de probióticos, destacando tecnologias avançadas para identificar potenciais candidatos a probióticos em nível molecular. Foram discutidos aspectos fundamentais da avaliação de segurança, que inclui a análise de fatores de virulência, resistência a antibióticos, presença de elementos transponíveis (como fagos, integrases, ilhas genômicas, plasmídeos e sistemas CRISPR-Cas), bem como a capacidade de produzir aminas biogênicas. Além disso, o texto apresentou uma visão detalhada sobre as ferramentas bioinformáticas que facilitam essas avaliações, proporcionando maior precisão e confiabilidade no processo de seleção e validação de linhagens.

Outro aspecto abordado foi a avaliação dos mecanismos de probiose, incluindo propriedades como a atividade antagonista contra patógenos, a habilidade de adesão às células do hospedeiro, a capacidade imunomoduladora e a produção de metabólitos secundários com impacto direto na saúde intestinal. Essas características são cruciais para determinar a eficácia de um microrganismo como probiótico e seu potencial de aplicação clínica.

Por fim, foram apresentados exemplos práticos de estudos que ilustram a aplicação da probiogenômica em diferentes contextos. Esses estudos destacam como a integração de abordagens multiômicas e técnicas avançadas pode acelerar a identificação e o desenvolvimento de probióticos utilizando probiogenômica, sendo mais seguros e eficazes, alinhados às demandas da ciência e da saúde moderna. Assim, o capítulo oferece uma visão abrangente e aprofundada sobre as etapas críticas no estudo de probióticos, desde a identificação molecular até a avaliação de segurança e funcionalidade, consolidando a probiogenômica como uma ferramenta essencial no campo da microbiologia aplicada e da biotecnologia.

6. Referências

- AL DOGHAITHER, H.; GULL, M. Plasmids as Genetic Tools and Their Applications in Ecology and Evolution. *Plasmid*, 2019. DOI: <https://doi.org/10.5772/INTECHOPEN.85705>.
- 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, p. D690–D699, 2023. DOI: <https://doi.org/10.1093/NAR/GKAC920>.
- ARNDT, D. et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, v. 44, n. W1, p. W16–W21, 2016. DOI: <https://doi.org/10.1093/NAR/GKW387>.
- AUCH, A.F.; VON JAN, M.; KLENK, H.P.; GÖKER, M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Standards in Genomic Sciences*, v. 2, n. 1, p. 117–134, 2010. DOI: <https://doi.org/10.4056/SIGS.531120/FIGURES/8>.
- AW, W.; FUKUDA, S. An Integrated Outlook on the Metagenome and Metabolome of Intestinal Diseases. *Diseases*, v. 3, n. 4, p. 341–359, 2015. DOI: <https://doi.org/10.3390/DISEASES3040341>.
- AZIZ, R.K. et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, v. 9, n. 1, p. 1–15, 2008. DOI: <https://doi.org/10.1186/1471-2164-9-75/TABLES/3>.
- BAUTISTA, A.G.; BARRADO, A.G. Bacteriocinas como bioconservador alimentario: características generales y aplicación en alimentos. *PubSaúde*, v. 12, 2023, 1–9. DOI: <https://doi.org/10.31533/PUBSAUDE12.A366>.

- BERMUDEZ-BRITO, M.; PLAZA-DÍAZ, J.; MUÑOZ-QUEZADA, S.; GÓMEZ-LLORENTE, C.; GIL, A. Probiotic Mechanisms of Action. *Annals of Nutrition and Metabolism*, v. 61, n. 2, p. 160–174, 2012. DOI: <https://doi.org/10.1159/000342079>.
- BLIN, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research*, v. 51, n. W1, p. W46–W50, 2023. DOI: <https://doi.org/10.1093/NAR/GKAD344>.
- BOOR, K.J. Bacterial Stress Responses: What Doesn't Kill Them Can Make Them Stronger. *PLoS Biology*, v. 4, n. 2, 2006, e23. DOI: <https://doi.org/10.1371/JOURNAL.PBIO.0040023>.
- BUSTOS, A.Y.; TARANTO, M.P.; GEREZ, C.L.; AGRIOPOULOU, S.; SMAOUI, S.; VARZAKAS, T.; EL ENSHASY, H.A. Recent Advances in the Understanding of Stress Resistance Mechanisms in Probiotics: Relevance for the Design of Functional Food Systems. *Probiotics and Antimicrobial Proteins*, 2024, p. 1–21. DOI: <https://doi.org/10.1007/S12602-024-10273-9>.
- CAETANO-ANOLLES, G. MAAP: a versatile and universal tool for genome analysis. *Plant Molecular Biology*, v. 25, n. 6, p. 1011–1026, 1994. DOI: <https://doi.org/10.1007/BF00014674/METRICS>.
- ÇAKAR, Z.P.; SAKA, H.A.; ECHENIQUE, J. Editorial: New insights in microbial stress tolerance mechanisms. *Frontiers in Microbiology*, v. 15, 2024, 1513485. DOI: <https://doi.org/10.3389/FMICB.2024.1513485>.
- CARATTOLI, A.; HASMAN, H. PlasmidFinder and In Silico pMLST: Identification and Typing of Plasmid Replicons in Whole-Genome Sequencing (WGS). *Methods in Molecular Biology*, v. 2075, p. 285–294, 2020. DOI: https://doi.org/10.1007/978-1-4939-9877-7_20.
- CARVALHO, R.D.O.; GUÉDON, E.; ABURJAILE, F.F.; AZEVEDO, V. Editorial: Probiogenomics of classic and next-generation probiotics. *Frontiers in Microbiology*, v. 13, 2022, 982642. DOI: <https://doi.org/10.3389/FMICB.2022.982642/BIBTEX>.
- CASTRO-LÓPEZ, C. et al. Genomics-based approaches to identify and predict the health-promoting and safety activities of promising probiotic strains – A probiogenomics review. *Trends in Food Science & Technology*, v. 108, p. 148–163, 2021. DOI: <https://doi.org/10.1016/J.TIFS.2020.12.017>.
- CASTRO-LÓPEZ, C. et al. Key Stress Response Mechanisms of Probiotics During Their Journey Through the Digestive System: A Review. *Probiotics and Antimicrobial Proteins*, v. 15, n. 5, p. 1250–1270, 2022. DOI: <https://doi.org/10.1007/S12602-022-09981-X>.
- CHEN, E. J. et al. Effects of different initial pH values on freeze-drying resistance of *Lactiplantibacillus plantarum* LIP-1 based on transcriptomics and proteomics. *Food Research International*, v. 149, 2021, 110694. DOI:

<https://doi.org/10.1016/J.FOODRES.2021.110694>.

- CLAESSON, M.J.; VAN SINDEREN, D.; O'TOOLE, P.W. *Lactobacillus* phylogenomics - Towards a reclassification of the genus. *International Journal of Systematic and Evolutionary Microbiology*, 1 v. 58, n. 12, p. 2945–2954, 2008. DOI: <https://doi.org/10.1099/IJS.0.65848-0/CITE/REFWORKS>.
- CRISTOFORI, F. et al. Anti-Inflammatory and Immunomodulatory Effects of Probiotics in Gut Inflammation: A Door to the Body. *Frontiers in Immunology*, v. 12, 2021, 578386. DOI: <https://doi.org/10.3389/FIMMU.2021.578386/BIBTEX>.
- CUNNINGHAM, M. et al. Shaping the Future of Probiotics and Prebiotics. *Trends in Microbiology*, v. 29, n. 8, p. 667–685, 2021. DOI: <https://doi.org/10.1016/j.tim.2021.04.003>.
- DA SILVA, T.F. et al. Unlocking the Potential of Probiotics: A Comprehensive Review on Research, Production, and Regulation of Probiotics. *Probiotics and Antimicrobial Proteins*, v. 16, n. 5, p. 1687–1723, 2024. DOI: <https://doi.org/10.1007/S12602-024-10247-X>.
- DE JESUS, L.C.L. et al. Health-promoting effects and safety aspects of *Lactobacillus delbrueckii*: A food industry species. *Trends in Food Science & Technology*, v. 150, 2024, 104605. DOI: <https://doi.org/10.1016/J.TIFS.2024.104605>.
- DE MELO PEREIRA, G.V. et al. How to select a probiotic? A review and update of methods and criteria. *Biotechnology Advances*, v. 36, n. 8, p. 2060–2076, 2018. DOI: <https://doi.org/10.1016/J.BIOTECHADV.2018.09.003>.
- DE NIES, L. et al. PathoFact: a pipeline for the prediction of virulence factors and antimicrobial resistance genes in metagenomic data. *Microbiome*, v. 9, n. 1, p. 1–14, 2021. DOI: <https://doi.org/10.1186/S40168-020-00993-9/FIGURES/6>.
- DE SIMONE, C. The Unregulated Probiotic Market. *Clinical Gastroenterology and Hepatology*, v. 17, n. 5, p. 809–817, 2019. DOI: <https://doi.org/10.1016/j.cgh.2018.01.018>.
- DUTRA-SILVA, L.; MATTEOLI, F.P.; ARISI, A.C.M. Distribution of Genes Related to Probiotic Effects Across *Lactobacillus rhamnosus* Revealed by Population Structure. *Probiotics and Antimicrobial Proteins*, v. 15, n. 3, p. 548–557, 2023. DOI: <https://doi.org/10.1007/S12602-021-09868-3/FIGURES/6>.
- ECHEGARAY, N. et al. A novel approach to *Lactiplantibacillus plantarum*: From probiotic properties to the omics insights. *Microbiological Research*, v. 268, 2023, 127289. DOI: <https://doi.org/10.1016/J.MICRES.2022.127289>.
- EDLUND, A. et al. Uncovering complex microbiome activities via metatranscriptomics during 24 hours of oral biofilm assembly and

- maturation. *Microbiome*, v. 6, n. 1, p. 1–22, 2018. DOI: <https://doi.org/10.1186/S40168-018-0591-4/FIGURES/4>.
- EFSA Panel on Biological Hazards (BIOHAZ). Scientific Opinion on risk based control of biogenic amine formation in fermented foods. *EFSA Journal*, v. 9, n. 10, 2393, 2011. DOI: <https://doi.org/10.2903/J.EFSA.2011.2393>.
- ESCOBAR-SÁNCHEZ, M. et al. Probiotic Properties and Proteomic Analysis of *Pediococcus pentosaceus* 1101. *Foods*, v. 12, n. 1, 2023, 46. DOI: <https://doi.org/10.3390/FOODS12010046/S1>.
- FEDOROFF, N.V. Transposable elements, epigenetics, and genome evolution. *Science*, v. 338, n. 6108, p. 758–767, 2012. DOI: https://doi.org/10.1126/SCIENCE.338.6108.758/ASSET/4D4639DF-CF55-418E-9680-EFB443D9A854/ASSETS/GRAPHIC/338_758_F9.JPEG.
- FELDGARDEN, M. et al. Validating the AMRFINDER tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrobial Agents and Chemotherapy*, v. 63, 2019. DOI: https://doi.org/10.1128/AAC.00483-19/SUPPL_FILE/AAC.00483-19-S0004.XLSX.
- FLORENZA, A.F.; KAAS, R.S.; CLAUSEN, P.T.L.C.; AYTAN-AKTUĞ, D.; AARESTRUP, F.M. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics*, v. 8, n. 1, 2022, 000748. DOI: <https://doi.org/10.1099/MGEN.0.000748/CITE/REFWORKS>.
- FONTANA, F. et al. Probiogenomics Analysis of 97 *Lactobacillus crispatus* Strains as a Tool for the Identification of Promising Next-Generation Probiotics. *Microorganisms*, v. 9, n. 1, 2020, 73. DOI: <https://doi.org/10.3390/MICROORGANISMS9010073>.
- FU, R.H.; WU, D.C.; YANG, W.; SCOFFIELD, J.; BENVENISTE, E.N. Probiotic Metabolites Promote Anti-inflammatory Functions of Immune Cells. *FASEB Journal*, v. 31, n. 1_supplement, p. 1048.1-1048.1, 2017. DOI: https://doi.org/10.1096/FASEBJ.31.1_SUPPLEMENT.1048.1.
- GALDEANO, C.M.; DE LEBLANC, A.D.M.; VINDEROLA, G.; BONET, M.E.B.; PERDIGÓN, G. Proposed model: Mechanisms of immunomodulation induced by probiotic bacteria. *Clinical and Vaccine Immunology*, v. 14, n. 5, p. 485–492, 2007. DOI: <https://doi.org/10.1128/CVI.00406-06/ASSET/1D11DF3F-7794-4E66-8623-2BEE10AE2854/ASSETS/GRAPHIC/ZCD0050729630004.JPEG>.
- GALPERIN, M.Y.; KRISTENSEN, D.M.; MAKAROVA, K.S.; WOLF, Y.I.; KOONIN, E.V. Microbial genome analysis: the COG approach. *Briefings in Bioinformatics*, v. 20, n. 4, p. 1063–1070, 2019. DOI: <https://doi.org/10.1093/BIB/BBX117>.

- GARCÍA-HERNÁNDEZ, Y. et al. Isolation, characterization and evaluation of probiotic lactic acid bacteria for potential use in animal production. *Research in Veterinary Science*, v. 108, p. 125–132, 2016. DOI: <https://doi.org/10.1016/J.RVSC.2016.08.009>.
- GENERALLY Recognized as Safe (GRAS) | FDA. [S. l.], [s. d.]. Disponível em: <https://www.fda.gov/food/food-ingredients-packaging/generally-recognized-safe-gras>. Acesso em: 10 dez. 2024.
- GITHUB - tseemann/abricate: :mag_right: Mass screening of contigs for antimicrobial and virulence genes. [S. l.], [s. d.]. Disponível em: <https://github.com/tseemann/abricate>. Acesso em: 13 dez. 2024.
- GOLYSHEV, M.A.; KOROTKOV, E.V. Developing of the Computer Method for Annotation of Bacterial Genes. *Advances in Bioinformatics*, 2015, 635437. DOI: <https://doi.org/10.1155/2015/635437>.
- GORREJA, F.; WALKER, W.A. The potential role of adherence factors in probiotic function in the gastrointestinal tract of adults and pediatrics: a narrative review of experimental and human studies. *Gut Microbes*, v. 14, 2022. DOI: <https://doi.org/10.1080/19490976.2022.2149214>.
- GRIESEMER, M.; NAVID, A. Uses of Multi-Objective Flux Analysis for Optimization of Microbial Production of Secondary Metabolites. *Microorganisms*, v. 11, n. 9, 2023, 2149. DOI: <https://doi.org/10.3390/MICROORGANISMS11092149/S1>.
- GRISSA, I.; VERGNAUD, G.; POURCEL, C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, v. 35, n. Web Server issue, p. W52–W57, 2007. DOI: <https://doi.org/10.1093/NAR/GKM360>.
- GU, Q.; LI, P. Biosynthesis of Vitamins by Probiotic Bacteria. In: *Probiotics and Prebiotics in Human Nutrition and Health*. [S. l.]: IntechOpen, 2016. DOI: <https://doi.org/10.5772/63117>.
- HAMON, E. et al. Comparative proteomic analysis of *Lactobacillus plantarum* for the identification of key proteins in bile tolerance. *BMC Microbiology*, v. 11, 2011, 63. DOI: <https://doi.org/10.1186/1471-2180-11-63/TABLES/3>.
- HE, Y. et al. PRAP: Pan Resistome analysis pipeline. *BMC Bioinformatics*, v. 21, n. 1, p. 1–8, 2020. DOI: <https://doi.org/10.1186/S12859-019-3335-Y/FIGURES/5>.
- HERNÁNDEZ-PLAZA, A. et al. eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Research*, v. 51, n. D1, p. D389–D394, 2023. DOI: <https://doi.org/10.1093/NAR/GKAC1022>.
- HILL, C. et al. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nature Reviews Gastroenterology & Hepatology*, v. 11, n. 8, p. 506–514, 2014. DOI: <https://doi.org/10.1038/nrgastro.2014.66>.

- IMPERIAL, I.C.V.J.; IBANA, J.A. Addressing the Antibiotic Resistance Problem with Probiotics: Reducing the Risk of Its Double-Edged Sword Effect. *Frontiers in Microbiology*, v. 7, 2016. DOI: <https://doi.org/10.3389/FMICB.2016.01983>.
- JIANG, S. et al. Multi-omics analysis of the effects of dietary changes and probiotics on diet-induced obesity. *Current Research in Food Science*, v. 6, 2023, 100435. DOI: <https://doi.org/10.1016/J.CRFS.2023.100435>.
- JUHAS, M.; VAN DER MEER, J.R.; GAILLARD, M.; HARDING, R.M.; HOOD, D.W.; CROOK, D.W. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews*, v. 33, n. 2, p. 376–393, 2009. DOI: <https://doi.org/10.1111/J.1574-6976.2008.00136.X>.
- KANEHISA, M. The KEGG Database. *Novartis Foundation Symposium*, v. 247, p. 91–103, 2002. DOI: <https://doi.org/10.1002/0470857897.CH8>.
- KANEHISA, M.; GOTO, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, v. 28, n. 1, p. 27–30, 2000. DOI: <https://doi.org/10.1093/NAR/28.1.27>.
- KATHIRIYA, M.R.; VEKARIYA, Y.V.; HATI, S. Understanding the Probiotic Bacterial Responses Against Various Stresses in Food Matrix and Gastrointestinal Tract: A Review. *Probiotics and Antimicrobial Proteins*, v. 15, n. 4, p. 1032–1048, 2023. DOI: <https://doi.org/10.1007/S12602-023-10104-3>.
- KESHET, A.; SEGAL, E. Identification of gut microbiome features associated with host metabolic health in a large population-based cohort. **Nature Communications**, v. 15, n. 1, 2024, 1–13. DOI: [\[https://doi.org/10.1038/s41467-024-53832-y\]](https://doi.org/10.1038/s41467-024-53832-y)(<https://doi.org/10.1038/s41467-024-53832-y>).
- KHATRI, I.; SHARMA, S.; RAMYA, T.N.C.; SUBRAMANIAN, S. Complete Genomes of *Bacillus coagulans* S-lac and *Bacillus subtilis* TO-A JPC, Two Phylogenetically Distinct Probiotics. *PLoS One*, v. 11, n. 6, 2016, e0156745. DOI: <https://doi.org/10.1371/JOURNAL.PONE.0156745>.
- KIEFT, K.; ZHOU, Z.; ANANTHARAMAN, K. VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome*, v. 8, n. 1, p. 1–23, 2020. DOI: <https://doi.org/10.1186/S40168-020-00867-0/TABLES/2>.
- KIM, J. et al. VAMPr: VArant Mapping and Prediction of antibiotic resistance via explainable features and machine learning. *PLOS Computational Biology*, v. 16, n. 12, 2020, e1007511. DOI: <https://doi.org/10.1371/JOURNAL.PCBI.1007511>.
- KIU, R. et al. *Bifidobacterium breve* UCC2003 Induces a Distinct Global Transcriptomic Program in Neonatal Murine Intestinal Epithelial Cells. *iScience*, v. 23, 2020. DOI:

<https://doi.org/10.1016/J.ISCI.2020.101336/ATTACHMENT/F9F94FE8-4166-4144-A2CA-5EECB6AA4A8A/MMC5.XLSX>.

- KOLLI, U. et al. Multi-omics analysis revealing the interplay between gut microbiome and the host following opioid use. *Gut Microbes*, v. 15, 2023. DOI: <https://doi.org/10.1080/19490976.2023.2246184>.
- KOPONEN, J.; LAAKSO, K.; KOSKENNIEMI, M.; KANKAINEN, M.; SAVIJOKI, K.; NYMAN, T.A.; DE VOS, W.M.; TYNKKYNNEN, S.; KALKKINEN, N.; VARMANEN, P. Effect of acid stress on protein expression and phosphorylation in *Lactobacillus rhamnosus* GG. *Journal of Proteomics*, v. 75, p. 1357–1374, 2012. DOI: <https://doi.org/10.1016/J.JPROT.2011.11.009>.
- KRAUSOVA, G.; HYRSLOVA, I.; HYNSTOVA, I. In Vitro Evaluation of Adhesion Capacity, Hydrophobicity, and Auto-Aggregation of Newly Isolated Potential Probiotic Strains. *Fermentation*, v. 5, n. 4, 2019, 100. DOI: <https://doi.org/10.3390/FERMENTATION5040100>.
- KWOJI, I.D.; AIYEGORO, O.A.; OKPEKU, M.; ADELEKE, M.A. ‘Multi-omics’ data integration: applications in probiotics studies. *Npj Science of Food*, v. 7, n. 1, p. 1–10, 2023. DOI: <https://doi.org/10.1038/s41538-023-00199-x>.
- LAKIN, S.M. et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Research*, v. 45, n. D1, p. D574–D580, 2017. DOI: <https://doi.org/10.1093/NAR/GKW1009>.
- LEBEER, S.; VANDERLEYDEN, J.; DE KEERSMAECKER, S.C.J. Genes and Molecules of *Lactobacilli* Supporting Probiotic Action. *Microbiology and Molecular Biology Reviews*, v. 72, n. 4, p. 728–764, 2008. DOI: <https://doi.org/10.1128/MMBR.00017-08/ASSET/436254AA-E677-4D30-9B6C-C8821F01B7C4/ASSETS/GRAPHIC/ZMR0040821960004.JPEG>.
- LEE, Y. et al. Characterization of a Potential Probiotic *Lactiplantibacillus plantarum* LRCC5310 by Comparative Genomic Analysis and its Vitamin B6 Production Ability. *Journal of Microbiology and Biotechnology*, v. 33, n. 3, p. 644–655, 2023. DOI: <https://doi.org/10.4014/JMB.2211.11016>.
- LEE, Y. et al. Mechanisms underlying probiotic effects on neurotransmission and stress resilience in fish via transcriptomic profiling. *Fish & Shellfish Immunology*, v. 141, 2023, 109063. DOI: <https://doi.org/10.1016/J.FSI.2023.109063>.
- LIN, T.-L. et al. Investiture of next generation probiotics on amelioration of diseases – Strains do matter. *Medical Microecology*, v. 1-2, 2019, 100002. DOI: <https://doi.org/10.1016/J.MEDMIC.2019.100002>.
- LINDSEY, R.L. et al. Rapid identification of enteric bacteria from whole genome sequences using average nucleotide identity metrics. *Frontiers in Microbiology*, v. 14, 2023, 1225207. DOI: <https://doi.org/10.3389/FMICB.2023.1225207/BIBTEX>.

- LIU, B.; ZHENG, D.; JIN, Q.; CHEN, L.; YANG, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research*, v. 47, n. D1, p. D687–D692, 2019. DOI: <https://doi.org/10.1093/NAR/GKY1080>.
- LOBB, B.; TREMBLAY, B.J.M.; MORENO-HAGELSIEB, G.; DOXEY, A.C. An assessment of genome annotation coverage across the bacterial tree of life. *Microbial Genomics*, v. 6, n. 1, 2020, e000341. DOI: <https://doi.org/10.1099/MGEN.0.000341/CITE/REFWORKS>.
- LUGLI, G.A. et al. The Probiotic Identity Card: A Novel “Probiogenomics” Approach to Investigate Probiotic Supplements. *Frontiers in Microbiology*, v. 12, 2022, 790881. DOI: <https://doi.org/10.3389/FMICB.2021.790881/BIBTEX>.
- MAAKE, T.W.; AIYEGORO, O.A.; ADELEKE, M.A. Effects of *Lactobacillus rhamnosus* and *Enterococcus faecalis* supplementation as direct-fed microbials on rumen microbiota of boer and speckled goat breeds. *Veterinary Sciences*, v. 8, n. 6, 2021, 103. DOI: <https://doi.org/10.3390/VETSCI8060103/S1>.
- MAKAROVA, K. et al. Comparative genomics of the lactic acid bacteria. *Proceedings of the National Academy of Sciences*, v. 103, n. 42, p. 15611–15616, 2006. DOI: https://doi.org/10.1073/PNAS.0607117103/SUPPL_FILE/07117TABLE1.DOC.
- MARCO, M.L.; PAVAN, S.; KLEEREBEZEM, M. Towards understanding molecular modes of probiotic action. *Current Opinion in Biotechnology*, v. 17, n. 2, p. 204–210, 2006. DOI: <https://doi.org/10.1016/J.COPBIO.2006.02.005>.
- MASSE, K.E.; LU, V.B. Short-chain fatty acids, secondary bile acids and indoles: gut microbial metabolites with effects on enteroendocrine cell function and their potential as therapies for metabolic disease. *Frontiers in Endocrinology (Lausanne)*, v. 14, 2023, 1169624. DOI: <https://doi.org/10.3389/FENDO.2023.1169624/BIBTEX>.
- MAZZANTINI, D.; CALVIGIONI, M.; CELANDRONI, F.; LUPETTI, A.; GHELARDI, E. In vitro assessment of probiotic attributes for strains contained in commercial formulations. *Scientific Reports*, v. 12, n. 1, 2022, 1–11. DOI: <https://doi.org/10.1038/s41598-022-25688-z>.
- MEIER-KOLTHOFF, J.P.; CARBASSE, J.S.; PEINADO-OLARTE, R.L.; GÖKER, M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Research*, v. 50, n. D1, p. D801–D807, 2022. DOI: <https://doi.org/10.1093/NAR/GKAB902>.
- MEIER-KOLTHOFF, J.P.; GÖKER, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature*

Communications, v. 10, 2019. DOI: <https://doi.org/10.1038/S41467-019-10210-3>.

- MILWARD, E.A.; SHAHANDEH, A.; HEIDARI, M.; JOHNSTONE, D.M.; DANESHI, N.; HONDERMARCK, H. Transcriptomics. *Encyclopedia of Cell Biology*, v. 4, p. 160–165, 2016. DOI: <https://doi.org/10.1016/B978-0-12-394447-4.40029-5>.
- MONSHIZADEH, M.; ZOMORODI, S.; MORTENSEN, K.; YE, Y. Revealing bacteria-phage interactions in human microbiome through the CRISPR-Cas immune systems. *Frontiers in Cellular and Infection Microbiology*, v. 12, 2022, 933516. DOI: <https://doi.org/10.3389/FCIMB.2022.933516/BIBTEX>.
- MU, Y. et al. Development and Applications of CRISPR/Cas9-Based Genome Editing in *Lactobacillus*. *International Journal of Molecular Sciences*, v. 23, n. 21, 2022, 12852. DOI: <https://doi.org/10.3390/ijms232112852>.
- NEMERGUT, D.R.; ROBESON, M.S.; KYSELA, R.F.; MARTIN, A.P.; SCHMIDT, S.K.; KNIGHT, R. Insights and inferences about integron evolution from genomic data. *BMC Genomics*, v. 9, 2008, p. 1–12. DOI: <https://doi.org/10.1186/1471-2164-9-261/FIGURES/5>.
- NETRUSOV, A.I.; LIYASKINA, E.V.; KURGAEVA, I.V.; LIYASKINA, A.U.; YANG, G.; REVIN, V.V. Exopolysaccharides Producing Bacteria: A Review. *Microorganisms*, v. 11, n. 6, 2023, 1541. DOI: <https://doi.org/10.3390/MICROORGANISMS11061541>.
- NYABAKO, B.A.; FANG, H.; CUI, F.J.; LIU, K.Y.; TAO, T.L.; ZAN, X.Y.; SUN, W.J. Enhanced Acid Tolerance in *Lactobacillus acidophilus* by Atmospheric and Room Temperature Plasma (ARTP) Coupled with Adaptive Laboratory Evolution (ALE). *Applied Biochemistry and Biotechnology*, v. 191, n. 4, p. 1499–1514, 2020. DOI: <https://doi.org/10.1007/S12010-020-03264-3/TABLES/2>.
- O'HARA, A.M.; SHANAHAN, F. Mechanisms of Action of Probiotics in Intestinal Diseases. *The Scientific World Journal*, v. 7, 2007, p. 31–46. DOI: <https://doi.org/10.1100/TSW.2007.26>.
- PECORARO, C.; CARBONE, D.; PARRINO, B.; CASCIOFERRO, S.; DIANA, P. Recent Developments in the Inhibition of Bacterial Adhesion as Promising Anti-Virulence Strategy. *International Journal of Molecular Sciences*, v. 24, n. 5, 2023, 4872. DOI: <https://doi.org/10.3390/IJMS24054872>.
- PENG, X.; ED-DRA, A.; YUE, M. Whole genome sequencing for the risk assessment of probiotic lactic acid bacteria. *Critical Reviews in Food Science and Nutrition*, v. 63, n. 28, p. 11244–11262, 2023. DOI: <https://doi.org/10.1080/10408398.2022.2087174>.
- PERA, N.P.; PIETERS, R.J. Towards bacterial adhesion-based therapeutics and detection methods. *MedChemComm*, v. 5, n. 7, p. 1027–1035, 2014.

DOI: <https://doi.org/10.1039/C3MD00346A>.

PRITCHARD, L. et al. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods*, v. 8, n. 1, p. 12–24, 2015. DOI: <https://doi.org/10.1039/C5AY02550H>.

PUIG-CASTELLVÍ, F. et al. Advances in the integration of metabolomics and metagenomics for human gut microbiome and their clinical applications. **TrAC Trends in Analytical Chemistry**, v. 167, 2023, 117248. DOI: <https://doi.org/10.1016/j.trac.2023.117248>(<https://doi.org/10.1016/j.trac.2023.117248>).

QUALIFIED presumption of safety (QPS) | EFSA. [S. l.], [s. d.]. Disponível em: <https://www.efsa.europa.eu/en/topics/topic/qualified-presumption-safety-qps>. Acesso em: 13 dez. 2024.

Regulation (EU) No 1019/2013 of 23 October 2013 amending Annex I to Regulation (EC) No 2073/2005 as regards histamine in fishery products. 1 [S. l.], [s. d.]. Disponível em: <https://eur-lex.europa.eu/eli/reg/2013>.

RICHTER, M.; ROSSELLÓ-MÓRA, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, v. 106, n. 45, p. 19126–19131, 2009. DOI: https://doi.org/10.1073/PNAS.0906412106/SUPPL_FILE/0906412106SI.PDF.

ROBERTSON, J.; NASH, J.H.E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics*, v. 4, 2018. DOI: <https://doi.org/10.1099/MGEN.0.000206>.

RODRIGUES, D.L.N. Desafios na padronização da anotação genômica. *BIOINFO*, v. 3, 2023, 05. DOI: <https://doi.org/10.51780/BIOINFO-03-05>.

RODRIGUES, D.L.N.; ARIUTE, J.C.; RODRIGUES DA COSTA, F.M.; BENKO-ISEPPON, A.M.; BARH, D.; AZEVEDO, V.; ABURJAILE, F. PanViTa: Pan Virulence and resisTance analysis. *Frontiers in Bioinformatics*, v. 3, 2023, 1070406. DOI: <https://doi.org/10.3389/FBINF.2023.1070406>.

RODRIGUES, R.R. et al. Transkingdom interactions between *Lactobacilli* and hepatic mitochondria attenuate western diet-induced diabetes. *Nature Communications*, v. 12, n. 1, p. 1–15, 2021. DOI: <https://doi.org/10.1038/s41467-020-20313-x>.

ROUQUETTE, C.; DE CHASTELLIER, C.; NAIR, S.; BERCHE, P. The ClpC ATPase of *Listeria monocytogenes* is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Molecular Microbiology*, v. 27, n. 6, p. 1235–1245, 1998. DOI: <https://doi.org/10.1046/J.1365-2958.1998.00775.X>.

ROUX, S.; ENAULT, F.; HURWITZ, B.L.; SULLIVAN, M.B. VirSorter: Mining viral signal from microbial genomic data. *PeerJ*, 2015, e985. DOI: <https://doi.org/10.7717/PEERJ.985/SUPP-2>.

- RUIZ, L.; HIDALGO, C.; BLANCO-MÍGUEZ, A.; LOURENÇO, A.; SÁNCHEZ, B.; MARGOLLES, A. Tackling probiotic and gut microbiota functionality through proteomics. *Journal of Proteomics*, v. 147, p. 28–39, 2016. DOI: <https://doi.org/10.1016/J.JPROT.2016.03.023>.
- SACHDEVA, G.; KUMAR, K.; JAIN, P.; RAMACHANDRAN, S. SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics*, v. 21, n. 3, p. 483–491, 2005. DOI: <https://doi.org/10.1093/BIOINFORMATICS/BI028>.
- SADRIMOVAHED, M.; ULUSOY, B.H. *Bacillus clausii*: A Review into Story of Its Probiotic Success and Potential Food Applications. *Fermentation*, v. 10, n. 10, 2024, 522. DOI: <https://doi.org/10.3390/FERMENTATION10100522>.
- SALMINEN, S. et al. The International Scientific Association of Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of postbiotics. *Nature Reviews Gastroenterology & Hepatology*, v. 18, n. 9, p. 649–667, 2021. DOI: <https://doi.org/10.1038/s41575-021-00440-6>.
- SÁNCHEZ, B. et al. Omics for the study of probiotic microorganisms. *Food Research International*, v. 54, n. 1, p. 1061–1071, 2013. DOI: <https://doi.org/10.1016/J.FOODRES.2013.01.029>.
- SANDERS, M.E. et al. Safety assessment of probiotics for human use. *Gut Microbes*, v. 1, n. 3, p. 164–185, 2010. DOI: <https://doi.org/10.4161/GMIC.1.3.12127>.
- SAVAGE, D.C.; FLETCHER, M. (Eds.). *Bacterial Adhesion*. [S. l.]: Springer US, 1985. DOI: <https://doi.org/10.1007/978-1-4615-6514-7>.
- SAYERS, E.W. et al. Database resources of the National Center for Biotechnology Information in 2023. *Nucleic Acids Research*, v. 51, n. D1, p. D29–D38, 2023. DOI: <https://doi.org/10.1093/NAR/GKAC1032>.
- SCHAFER, S.; CHEN, K.; MA, L. Crosstalking with dendritic cells: a path to engineer advanced T Cell immunotherapy. *Frontiers in Systems Biology*, v. 4, 2024, 1372995. DOI: <https://doi.org/10.3389/FSYS>
- SEDDIK, H.A.; BENDALI, F.; GANCEL, F.; FLISS, I.; SPANO, G.; DRIDER, D. *Lactobacillus plantarum* and Its Probiotic and Food Potentialities. *Probiotics and Antimicrobial Proteins*, v. 9, n. 2, p. 111–122, 2017. DOI: <https://doi.org/10.1007/S12602-017-9264-Z/TABLES/2>.
- SEEMANN, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, v. 30, n. 14, p. 2068–2069, 2014. DOI: <https://doi.org/10.1093/BIOINFORMATICS/BTU153>.
- SHANG, J.; PENG, C.; LIAO, H.; TANG, X.; SUN, Y. PhaBOX: a web server for identifying and characterizing phage contigs in metagenomic data. *Bioinformatics Advances*, v. 3, 2023. DOI: <https://doi.org/10.1093/BIOADV/VBAD101>.

- SHANG, J.; TANG, X.; SUN, Y. PhaTYP: predicting the lifestyle for bacteriophages using BERT. *Briefings in Bioinformatics*, v. 24, 2023. DOI: <https://doi.org/10.1093/BIB/BBAC487>.
- SHARMA, A.; LEE, S.; PARK, Y.S. Molecular typing tools for identifying and characterizing lactic acid bacteria: a review. *Food Science and Biotechnology*, v. 29, n. 6, p. 1301–1318, 2020. DOI: <https://doi.org/10.1007/S10068-020-00802-X/TABLES/1>.
- SINGH, D.; SINGH, A.; KUMAR, S. Probiotics: friend or foe to the human immune system. *Bulletin of the National Research Centre*, v. 47, n. 1, 2023, 47. DOI: <https://doi.org/10.1186/S42269-023-01098-7>.
- SINGH, S.; LE, N.Q.K.; WANG, C. VF-Pred: Predicting virulence factor using sequence alignment percentage and ensemble learning models. *Computational Biology and Medicine*, v. 168, 2024, 107662. DOI: <https://doi.org/10.1016/J.COMPBIOMED.2023.107662>.
- SOARES, S.C.; ABREU-VINÍ, V.A.C.; RAMOS, R.T.J.; CERDEIRA, L.; SILVA, A.; BAUMBACH, J.; TROST, E.; TAUCH, A.; HIRATA, R.; MATTOS-GUARALDI, A.L.; MIYOSHI, A.; AZEVEDO, V. PIPS: Pathogenicity Island Prediction Software. *PLoS One*, v. 7, n. 1, 2012, e30848. DOI: <https://doi.org/10.1371/JOURNAL.PONE.0030848>.
- SOARES, S.C.; GEYIK, H.; RAMOS, R.T.J.; DE SÁ, P.H.C.G.; BARBOSA, E.G.V.; BAUMBACH, J.; FIGUEIREDO, H.C.P.; MIYOSHI, A.; TAUCH, A.; SILVA, A.; AZEVEDO, V. GIPSy: Genomic island prediction software. *Journal of Biotechnology*, v. 232, p. 2–11, 2016. DOI: <https://doi.org/10.1016/J.JBIOTEC.2015.09.008>.
- SONG, Z.; DU, H.; ZHANG, Y.; XU, Y. Unraveling core functional microbiota in traditional solid-state fermentation by high-throughput amplicons and metatranscriptomics sequencing. *Frontiers in Microbiology*, v. 8, 2017, 262958. DOI: <https://doi.org/10.3389/FMICB.2017.01294/BIBTEX>.
- STEIN, L. Genome annotation: from sequence to biology. *Nature Reviews Genetics*, v. 2, n. 7, p. 493–503, 2001. DOI: <https://doi.org/10.1038/35080529>.
- STRZELEC, M.; DETKA, J.; MIESZCZAK, P.; SOBOCIŃSKA, M.K.; MAJKA, M. Immunomodulation—a general review of the current state-of-the-art and new therapeutic strategies for targeting the immune system. *Frontiers in Immunology*, v. 14, 2023, 1127704. DOI: <https://doi.org/10.3389/FIMMU.2023.1127704/BIBTEX>.
- THIRUMURUGAN, D.; CHOLARAJAN, A.; VIJAYAKUMAR, R. An Introductory Chapter: Secondary Metabolites. In: *Secondary Metabolites - Sources and Applications*. [S. l.]: IntechOpen, 2018. DOI: <https://doi.org/10.5772/INTECHOPEN.79766>.
- THUY, T.T.D. et al. Characterization of the broad-spectrum antibacterial activity of bacteriocin-like inhibitory substance-producing probiotics isolated from

- fermented foods. *BMC Microbiology*, v. 24, 2024, 1–15. DOI: <https://doi.org/10.1186/S12866-024-03245-0/FIGURES/6>.
- TSAI, Y.T.; CHENG, P.C.; PAN, T.M. The immunomodulatory effects of lactic acid bacteria for improving immune functions and benefits. *Applied Microbiology and Biotechnology*, v. 96, n. 4, p. 853–862, 2012. DOI: <https://doi.org/10.1007/S00253-012-4407-3/FIGURES/3>.
- VAN DER OOST, J.; WESTRA, E.R.; JACKSON, R.N.; WIEDENHEFT, B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nature Reviews Microbiology*, v. 12, n. 7, p. 479–492, 2014. DOI: <https://doi.org/10.1038/nrmicro3279>.
- VAN HEEL, A.J.; DE JONG, A.; SONG, C.; VIEL, J.H.; KOK, J.; KUIPERS, O.P. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Research*, v. 46, n. W1, p. W278–W281, 2018. DOI: <https://doi.org/10.1093/NAR/GKY383>.
- VARANKOVICH, N.V.; NICKERSON, M.T.; KORBER, D.R. Probiotic-based strategies for therapeutic and prophylactic use against multiple gastrointestinal diseases. *Frontiers in Microbiology*, v. 6, 2015, 133694. DOI: <https://doi.org/10.3389/FMICB.2015.00685/BIBTEX>.
- VENTURA, M. et al. Genome-scale analyses of health-promoting bacteria: probiogenomics. *Nature Reviews Microbiology*, v. 7, n. 1, p. 61–71, 2008. DOI: <https://doi.org/10.1038/nrmicro2047>.
- VENTURA, M.; TURRONI, F.; VAN SINDEREN, D. Probiogenomics as a tool to obtain genetic insights into adaptation of probiotic bacteria to the human gut. *Bioengineered*, v. 3, p. 73–79, 2012. DOI: <https://doi.org/10.4161/bbug.18540>. (Note: References 25 and 27 are identical.)
- WASEEM, H. et al. Virulence factor activity relationships (VFARs): a bioinformatics perspective. *Environmental Science: Processes & Impacts*, v. 19, n. 2, p. 247–260, 2017. DOI: <https://doi.org/10.1039/C6EM00689B>.
- WATTAM, A.R. et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, v. 42, n. D1, p. D581–D591, 2014. DOI: <https://doi.org/10.1093/NAR/GKT1099>.
- WINAND, R. et al. Targeting the 16S rRNA Gene for Bacterial Identification in Complex Mixed Samples: Comparative Evaluation of Second (Illumina) and Third (Oxford Nanopore Technologies) Generation Sequencing Technologies. *International Journal of Molecular Sciences*, v. 21, n. 1, 2019, 298. DOI: <https://doi.org/10.3390/IJMS21010298>.
- WISHART, D.S. et al. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Research*, v. 51, n. W1, p. W443–W450, 2023. DOI: <https://doi.org/10.1093/NAR/GKAD382>.
- WÓJCIK, W.; ŁUKASIEWICZ, M.; PUPPEL, K. Biogenic amines: formation,

- action and toxicity – a review. *Journal of the Science of Food and Agriculture*, v. 101, n. 7, p. 2634–2640, 2021. DOI: <https://doi.org/10.1002/JSFA.10928>.
- XIE, R. et al. DeepVF: a deep learning-based hybrid framework for identifying virulence factors using the stacking strategy. *Briefings in Bioinformatics*, v. 22, 2021. DOI: <https://doi.org/10.1093/BIB/BBAA125>.
- YAN, S.; ZENG, M.; WANG, H.; ZHANG, H. Micromonospora: A Prolific Source of Bioactive Secondary Metabolites with Therapeutic Potential. *Journal of Medicinal Chemistry*, 1 v. 65, n. 6, p. 8735–8771, 2022. DOI: https://doi.org/10.1021/ACS.JMEDCHEM.2C00626/SUPPL_FILE/JM2C00626_SI_001.PDF.
- YETIMAN, A.E.; ORTAKCI, F. Genomic, probiotic, and metabolic potentials of *Liquorilactobacillus nagelii* AGA58, a novel bacteriocinogenic motile strain isolated from lactic acid-fermented shalgam. *Journal of Bioscience and Bioengineering*, v. 135, p. 34–43, 2023. DOI: <https://doi.org/10.1016/J.JBIOSC.2022.10.008>.
- YIN, R. et al. Gut Microbiota, Dietary Phytochemicals, and Benefits to Human Health. *Current Pharmacology Reports*, v. 5, n. 5, p. 332–344, 2019. DOI: <https://doi.org/10.1007/S40495-019-00196-3>.
- ZHANG, P. et al. A dynamic mouse peptidome landscape reveals probiotic modulation of the gut-brain axis. *Science Signaling*, v. 13, 2020. DOI: https://doi.org/10.1126/SCISIGNAL.ABB0443/SUPPL_FILE/ABB0443_S M.PDF.
- ZHANG, X. et al. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*, v. 9, n. 1, p. 1–14, 2018. DOI: <https://doi.org/10.1038/s41467-018-05357-4>.
- ZHAO, Q. et al. Drug-microbiota interactions: an emerging priority for precision medicine. **Signal Transduction and Targeted Therapy**, v. 8, n. 1, 2023, 1–27. DOI: <https://doi.org/10.1038/s41392-023-01619-w>
- ZHU, J. et al. Mechanisms of probiotic *Bacillus* against enteric bacterial infections. *One Health Advances*, v. 1, n. 1, 2023, 1–22. DOI: <https://doi.org/10.1186/S44280-023-00020-0>.
- ZMORA, N. et al. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell*, v. 174, n. 6, p. 1388-1405.e21, 2018. DOI: <https://doi.org/10.1016/J.CELL.2018.08.041/ASSET/4F76245A-B166-475C-B964-7AA2683CA6DE/MAIN.ASSETS/GR7.JPG>.

Autores

Eduarda Guimarães Sousa¹, Lucas Gabriel Rodrigues Gomes¹, Arun Kumar Jaiswal¹, Laura Diniz Freitas¹, Rafael Júnio de Oliveira¹, Letícia Isadora Rodrigues¹, Luiza Pereira Reis¹, Aline Ferreira Maciel de Oliveira¹, Siomar de Castro Soares² e Vasco Ariston de Carvalho Azevedo¹

1. Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brasil.
2. Laboratório de Bioinformática, Departamento de Ciências Biológicas, Universidade Federal do Triângulo Mineiro, Uberaba, Minas Gerais, Brasil.

CAPÍTULO 4

Microalgal biotechnology: a new green economy?

Luis G. Ramírez Mérida

<https://doi.org/10.69570/mp.978-65-84548-32-9.c4>

Resumo

The green economy emerges as a process that seeks to improve society based on natural methodologies that generate environmentally friendly technologies. Recently, it has been possible to extrapolate this concept to biological processes, one of which is microalgae cultivation. Microalgae have great potential to produce biofuels and bioactive compounds; this technology can contribute to the development of clean energy and high value-added compounds. Although microalgae biotechnology is presented as a promising technology, improvements are needed in the production processes. Thus, microalgal biotechnology seeks to maintain high yields in biomass production, which makes the microalgal industry an economically attractive company for investors.

Palavras-chave: bioproducts, cultivation, photobioreactor, microalgae, technology

1. Introduction

The green economy emerged as a process to generate improvements and sustainability in the environment. Although a relatively recent terminology, organizations, institutions, processes and technologies are included in this term, ranging from renewable energy to waste treatment, to natural resources management and water treatment, among others. Recently the green economy emphasizes the exploitation of biological systems for the benefit of mankind. This is where the microalgae biotechnological attracts the interest of many researchers and entrepreneurs in diverse areas including agricultural, food, pharmaceutical, environmental, in view of the great potentiality and attributes

presented by microalgal biomass for the production of nutraceuticals, animal and human food, biofuels, fertilizers, chemicals of high added value, among others (YAACOB et al., 2022).

There are many variations in the microalgae biochemical composition which depend on the strain type as well as the culture conditions and downstream processing of the biomass. Thus, for example, when subjecting strains of *Chlorella vulgaris* to different irradiations the growth rate varies from 0.08 d⁻¹ to 1.13 d⁻¹ when irradiated in a photoperiod of 16: 8 ha 37.5 and 100 Mmol photons m⁻²s⁻¹ respectively (SEYFABADI et al. 2011), . On the other hand, seasonal variations may interfere with cellular performance. Open culture systems of *Nannochloropsis oculata* showed fluctuations in lipid yield. In summer and autumn were the most favorable seasons for the production of fatty acids more suitable for biodiesel production (MINHAS et al., 2023).

Microalgae present great potential to produce metabolic compounds of industrial interest. However, some aspects should be considered to make microalgal biotechnology an economically viable industry such as: (i) improvement in the strains selection, particularly those with high biotechnological potential, (ii) production of differentiated culture media for isolation, (iii) standardize biochemical modification methods, (iv) optimization of growth and biomass productivity, (v) design of photobioreactors suitable for large scale production, (vi) improvement in biorefinery processes. There are industries that develop biotechnological processes with microalgae distributed in the world, the considerations developed by them would contribute in the construction of an innovative technological platform for industries based on microalgae (YAP et al., 2021).

Microalgal culture has been used on industrial scale for decades to produce compounds that have repercussions on economic processes. In recent years, there has been an overall increase in microalgal biotechnology for commercial applications, such as the production of biofuels, bioactive compounds and bioremediation. However, few strains are used on an industrial scale, *Chlorella vulgaris*, *Dunaliella salina*, *Haematococcus pluvialis*, *Isochrysis galbana*, *Spirulina maxima* since their development and cultivation characteristics help to obtain better yields and to avoid cross contamination (ELADL et al., 2024).

Microalgae biotechnology presents advantages that catch the attention of many investors: (i) microalgal culture for biofuel production, which do not compromise the food chain, (ii) autotrophic and heterotrophic growth capacity, (iii) relatively easy obtaining of metabolic compounds, (iv) production of food compounds with nutritional and health benefits, (v) carbon biofixation, product of the microalgae photosynthetic metabolism. The production of microalgal biomass can be improved by introducing carbon dioxide (CO_2) from industrial effluents, which in addition to contributing and having high biomass productivity reduce process costs, which could help reduce greenhouse gas emissions (CHIA et al., 2021).

The process of microalgal biotechnology involves basically two steps: upstream processing and downstream processing, each of which has several established phases ranging from nutrient composition to culture systems, types of lighting, harvesting and biorefinery (Figure 1).

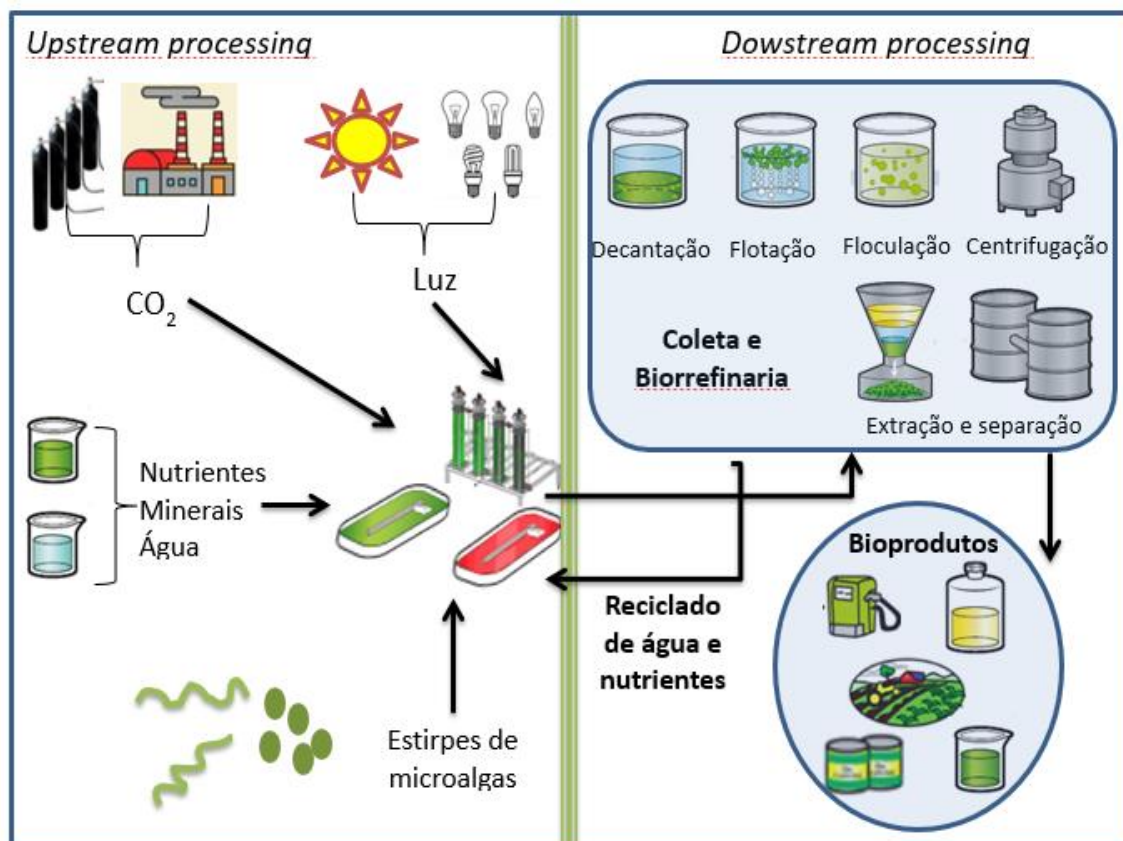


Figure 1. Scheme of the formation process of microalgae biomass and bioproducts.

In view of the fact that microalgae biotechnology offers great investment opportunities, the objective is to obtain a high production and good yield of microalgal biomass, which will serve as a source of a bioproducts diversity. The world market value of carotenoids can reach up to UD\$ 1 billion per year, and bioproducts such as astaxanthin may have higher prices (SHOW, 2022). Although these types of biotech companies have great potential for profit, many challenges need to be solved. One of the major challenges for the commercial operations of microalgae is the large-scale production, since at the present time the productive processes are carried out in large measure with open photobioreactors, which limits the growth and the production to few strains types and bioproducts. The closed photobioreactors available on the market do not represent an optimal solution for the companies'needs, although the volumetric and surface productivity is higher than in the open photobioreactors, the performance does not approximate to yield obtained in laboratory scale (SARKER and KAPARAJU 2023). In this paper, we will discuss the need to optimize strains, culture systems, production costs, genetic modification, to make microalgal biotechnology economically viable.

2. Microalgae culture

Despite the great diversity of microalgae in the world, estimated at three hundred thousand, only thirty thousand have been identified and of them about fifty species are used in biotechnological research, only six to ten are used for commercial production of compounds and food additives, but not yet the number of microalgae species that can be used to produce biofuels has been defined (AHMAD et al., 2022). To initiate a microalgal culture, it is necessary to start with a mother culture which must be in adequate conditions to avoid problems of deterioration or contamination in the time. Lyophilization techniques are advantageous. Also the preservation of mother cultures in tubes or Petri plates with minimum medium is used. However, in the latter, cell degeneration can be present, which can be reduced with the addition of a new organic base and continuous peals (SARAVANA, 2023). In addition, manipulation of culture conditions in terms of lighting intensity, temperature and nutrients, produces improved cellular yields (MASOJÍDEK et al., 2021).

The culture conditions, medium type, lighting, coadjutants, temperature, directly influence the costs. For autotrophic culture there are approximately forty types of medium consisting basically of mineral salts, which depending on the culture system should be supplemented with a carbon source. Some commercial medium supplemented with vitamins, amino acids and/or phyto-hormones are used for example the F/2 medium costing UD\$ 40 - 60 per liter (CHAKRABORTY et al., 2023).

The microalgal photosynthetic culture depends basically on the light energy in the cells. Inadequate illumination produces biochemical consequences, the product of photolimitation and photoinhibition that affects the kinetic performance in the reactors. The source and type of lighting in the systems is a factor to consider since, based on the characteristics of the process can use natural or artificial lighting. Other aspects that should be considered are the place of establishment of the culture, seasonal variations and variations of the photoperiod (SEYFABADI et al., 2011). Temperature is another important factor associated with microalgal cultures. In general, the optimum growth temperature is in the mesophilic region (25-35°C), although some thermophilic strains resistant to temperatures in the range of 60 °C. The great majority of culture systems assume the temperature changes as a result of the variation of the environment, although the use of heating mantles, coils and external heat exchangers can be installed in the reactors for temperature control (Satiro et al., 2016).

Regarding pH, microalgal culture is normally favored in a range that varies from neutral to alkaline (7-9), although some species show optimum growth at pH 4 (*Chlorococcum littorale*) and pH 10 (*Spirulina platensis*). There is a complex relationship between the concentration and type of the carbon and nitrogen source, since the compounds of NH_3 and NH_4^+ are strongly affected by the pH of the culture medium. Thus, pH control mechanisms must be integrated to the culture system to maintain the optimal conditions of this parameter in the process (MARKOU et al., 2016).

In photosynthetic culture systems, CO_2 is the main source of carbon used in culture; this gas has an atmospheric concentration of 0.034% which is not sufficient to maintain high biomass yields. CO_2 must be supplied through

concentrated stationary sources (compressed carbon dioxide, primary standard mixtures and industrial combustion gases) which provide inlet concentrations in levels between 3 and 15%. Ben-Amoz (2009) showed that *Dunaliella* cultures fed with flue gases increases biomass yield up to 10 times, showing to be a source of economic carbon, besides contributing to the decrease of greenhouse gases. The approximate costs of dry microalgae biomass in the 1950s were UD \$ 1,000 kg⁻¹, nowadays thanks to biotechnological development can be found values up to UD \$ 0.20 kg⁻¹ when grown in wastewater (DE PAUW et al., 1984).

However, there is a broad expectation in the availability and use of CO₂ from industrial combustion gases to support photosynthetic microalgae cultures. However, there have been difficulties in using this technique, since these gases have high indices of impurity (CO, CH₄, NO_x, SO_x, H₂ and particulate matter), and the temperature of these gases is very high (JACOB-LOPES & FRANCO, 2013). In the same way, heterotrophic and mixotrophic culture have been used in environmental research and has been incorporated in studies of the industrial waste use (STREIT et al., 2015).

Microalgae culture used on large scale include two types: open and closed systems. Open systems are easier to build, more economical, and relatively simple control that closed systems. There are two types of open systems that are best known: (i) circular tanks and (ii) raceway tanks (RAMÍREZ-MÉRIDA et al., 2013). Although simple to construct, these systems have limitations in the transfer of CO₂, generate high evaporation rates producing significant variations in the composition of the culture medium, besides allowing the access of external contaminants. All this makes the use of open systems limited to extremophilic microalgae strains, which are used at industrial level basically for the production of food derived bioproducts (ROJAS-VILLALTA et al., 2024).

The closed systems provide a greater control of the physicochemical parameters during the cycles of microalgae growth, besides allowing the conduction of photosynthetic, mixotrophic and heterotrophic cultures. The main configurations of the closed systems are formed by photobioreactors type: tubular (horizontal, vertical, bubble column and air-lift), flat plate, hybrids and fermenters.

Tubular photobioreactors are the most used at laboratory, pilot and industrial level. The tubular photobioreactors (horizontal and vertical), so called based on the arrangement of their transparent parallel tubes with diameters of 0.1 - 0.2 m. The culture medium circulates through them with gravitational aid or driven with a pump. The greater the number of tubes, the greater the light energy arrangement in the medium, and therefore the surface/volume (S/V) ratio improves. Disadvantages arise when oxygen accumulates, which is toxic to cells, as well as biofilm formation in the walls of the tubes, which decreases light input and increases the cost of cleaning the photobioreactor. In addition, constant recirculation of medium can cause cellular stress (SEGURA-MORALES et al. 2024).

The bubble column and airlift photobioreactors have a cylindrical shape, wide S/V ratio and a gas diffuser at the bottom of the column. They have good kinetic performance at low volume. However, their scaling is difficult because of the low reaction volume (PALADINO AND NEVIANI et al., 2021).

Flat plate photobioreactors are designed to make efficient use of sunlight; they have a high S/V ratio. This type of photobioreactors can be organized in such a way that they are oriented towards the sun, which allows better efficiency in terms of absorbed energy (PALADINO AND NEVIANI et al., 2021). The photobioreactor presents good kinetic performance and can be scaled by arranging several surface plates to gain area, but the plate elongation is not recommended for scaling, making it expensive to install on an industrial scale (Padrón and Mérida, 2025) .

The fermenters, for heterotrophic microalgae cultures, has been used in the last years, with advantages since they use substrates of industrial waste. They present high kinetic performance and are easily scalable when using configurations with 1.0-1.5 A/D ratios (RAMIREZ-MÉRIDA et al., 2015).

A comparative study of production costs per liter of microalgae oil in open and closed culture systems showed that in open systems the cost is almost three times lower than in closed autotrophic systems. On the other hand, heterotrophic crops proved to be even more economical. Production costs in autotrophic culture range from UD \$ 0.1 kg⁻¹ to UD \$ 32 kg⁻¹ compared to the heterotrophic culture UD \$ 2.0 kg⁻¹ to UD \$ 12 kg⁻¹ (ALABI, 2009). Although it is not currently possible

to compete with fossil fuels, it is expected that in the next few years, with the help of biotechnology, microalgae will produce economically viable biofuel.

3. Microalgae strains

The search for and collection of microalgae is relatively easy, although finding microalgae of biotechnological interest is more complicated. Microalgal strains must present characteristics applicable to biotechnological processes that are attractive to entrepreneurs. Thus, strains that present high growth kinetics, with high biomass yield, that are cultivated in wastewater or under natural lighting and temperature conditions, that produce metabolites of interest and that are easily separable would be desirable characteristics for the optimization of production processes. Thus, the choice of commercial microalgae strains is of utmost importance and deserves rigorous investigation. In their natural habitat, the species are well adapted to local environmental conditions and their usefulness contributes to more successful cultivation than non-native species. The cultivation of microalgae in wastewater is profitable in the production of microalgal biomass, and helps in the removal of nutrients (RAMÍREZ-MÉRIDA and RODRÍGUEZ-PADRÓN, 2023). Many research groups show attention to extremophilous strains. They have certain advantages, since they can withstand extreme conditions of temperature, toxic substances, lighting, pressure, among others, which guarantee their usefulness in biotechnological areas of interest. For example, they could directly use combustion gases from industrial furnaces, present high desiccation and photophysiological tolerance, tolerate high concentrations of heavy metals for bioremediation, present resistance to photoinhibition, produce a greater quantity of metabolites under certain conditions, allowing their use to have great potential in biotechnological processes.

4. Biochemical manipulation and high yield

The biochemical manipulation of microalgae is reflected in the composition of the chemical constituents. Several experiments have been reported, when manipulating the nutrients in *Scenedesmus obliquus* cultures, the lipid content

was five to ten times higher than the controls (MANDAL and MALLICK, 2009). On the other hand, it is interesting to note that the carotenoid content in *Dunaliella salina* cultures increased in direct proportion to the increase in salinity (Reshma et al., 2021). Takagi et al., (2006) showed that the salt content of the medium can also be a stressor in *Dunaliella*. In the early stages of the cultures, when NaCl was increased from 0.5 M (equivalent to seawater) to 1.0 M, lipids increased by 67%; when mid- or late-stage cultures were subjected to a similar stress, lipids increased by up to 70%. Thus, when selecting cells for biotechnological applications, the strain and physiological state of the microalgae play critical roles in determining production. The culture medium used influences the cellular biochemical profile as well as the rate of metabolic production. Through biochemical manipulation, lipid content can be regulated, since by placing nutrients in a limiting manner, physiological stress is generated, forcing the cell to alter its metabolism, leading to the accumulation of compounds necessary for maintenance. Some experiments show that in *Neochloris oleoabundans* cultures enriched with sodium nitrate, urea and ammonium bicarbonate as a nitrogen source, only the lowest levels of sodium nitrate showed an increase in lipids (WU et al., 2023). A limitation of inorganic phosphorus and CO₂ in *Chlamydomonas acidophila* cultures resulted in high rates (BOSSA et al., 2024). In *Monodus subterraneus* cultures, as the phosphate content decreases, the concentration of lipids in the cell increases (KHOZIN-GOLDBERG AND COHEN, 2006). Nitrogen limitation in *Scenedesmus obliquus* crops produces an increase in cellular lipids from 12.7% to 43% of dry weight (MANDAL and MALLICK, 2009); similarly, an increase in CO₂ concentration from 2% to 10% produces an increase in lipid content of up to 170% in seven days (MURADYAN et al., 2004; Tambat et al., 2023).

Therefore, several experiments must continue to advance in order to understand and establish optimal formulations and conditions of nutrients, temperature, and lighting that can guarantee high yields of microalgal biomass.

5. Genetic modification of microalgae

Genomic analyses of microalgae are available for only a few species. Genetic modification of microalgae is promising as a strategy to achieve higher

yields in the production of bioproducts for industrial purposes. Although several hundred microalgal strains have been cultivated, detailed investigation of cellular physiology and biochemistry is limited to less than one percent of the species, and even fewer in those whose genomes have been studied. Genetic transformation of microalgae has been limited by the presence of rigid cell walls (Shivakumar et al., 2024). However, using various techniques such as bombardment, electroporation and treatment with glass beads, several species have been genetically modified (Mosey et al., 2021).

Genetic modifications can confer properties to improve yield. For example, Li and Tsai (2008) demonstrated that the microalgae *Nannochloropsis oculata* was optimized to produce bovine lactoferricin (BLF) fused with a red fluorescent protein (DsRed), which has a bactericidal effect against *V. parahaemolyticus* infection in the digestive tract of shrimp. Microalgae engineering processes such as conversion of autotrophic systems into heterotrophic ones, improvement in photosynthetic efficiency, improvement of cultivation processes, among others, aiming at increasing cellular lipid synthesis, have generated promising results. However, the increase in lipid synthesis yield with genetic modifications has not yet been achieved at the expected levels. There are three possible strategies to improve lipid production: biochemical engineering, genetic engineering and transcription factor engineering. Biochemical engineering approaches are currently the most widely established in microalgae lipid production (BEHERA et al., 2021). Improvements in photosynthesis efficiency can be achieved by manipulating photosynthetic antennae. Since smaller antennae lead to greater photosynthetic efficiency, mutations in the genes that control their biogenesis are a possible mechanism to increase photosynthetic efficiency (WANG et al., 2022).

There is the possibility of improving the efficiency of solar energy conversion from 1-4% to 8-12%, to fully perform the metabolic process in the microalgae *Chlamydomonas perigranulata*, *C. reinhardtii*, *Chlorella vulgaris*, *Cyclotella* sp., *Dunaliella salina*, *Scenedesmus obliquus* and *Synechocystis* PCC 6714 (STEPHENS et al., 2010).

Microalgal biotechnological processes present an opportunity to generate methodologies and products that may be useful to society. Cell strains that have all the properties required for large-scale biotechnological processes must be

developed, and further research must be conducted in parallel with natural strains to fully understand their physiological functioning. Such modifications may confer properties to improve their yield.

6. Harvesting

There is no single best method for harvesting microalgae and reducing their water content. Cost-effective and energy-efficient harvesting methods are needed to make the biomass and bioproducts production process economical. Microalgal biomass recovery represents 20-30% of total production costs according to one source. The most common harvesting processes are flocculation, filtration, centrifugation, gravity sedimentation, and flotation. The selection of the harvesting technique will depend on the properties of the microalgae, such as density, size, and the value of the desired products (JACOB-LOPES et al., 2015).

Flocculation is a process in which the dispersion of particles from the medium is induced with the help of chemical compounds that produce aggregation in the microalgal cells. Flocculants stimulate flocculation by generating colloids and other particles in suspension in the solution to form floccules. The most commonly used for microalgae harvesting include ferric chloride, aluminum sulfate, aluminum, ferric sulfate, polyferric sulfate, as well as cationic polymers (polyelectrolytes), organic flocculants (chitosan). Bioflocculants such as *Paenibacillus* + aluminum sulfate, and also co-bioflocculation (*Nannochloris* + diatoms) are used, but require extra effort to cultivate other microalgae (MOLINA-GRIMA et al., 2003).

Pressure or vacuum filtration is satisfactory for recovering relatively large (> 70 μm) and/or filamentous microalgae, but are less effective for separating microalgae species with dimensions close to prokaryotes. Membrane microfiltration and ultrafiltration processes may be an option for recovering microalgae biomass below 30 μm (MOLINA-GRIMA et al., 2003). Cross-flow filtration allowed the recovery of 70-89% of microalgae. This procedure has the advantage of maintaining the integrity of the microalgae biomass. Filtration is an

expensive process due to membrane exchange and pumping (ZHANG et al., 1995).

The centrifugation technique is a methodology that is based on the use of centrifugal acceleration to achieve the sedimentation of microalgae in heterogeneous mixtures. The size and density of the structure determine the centrifugal separation of the element. The supernatant is a liquid located in the upper layer of the centrifuge tube, once decanted. The remaining solid represents the microalgae concentrate. The process is fast and intensive and depends on the sedimentation of the cells for the recovery of the biomass, the residence time of the cell suspension in the centrifuge and the depth of settling (MOLINA-GRIMA et al., 2003). Recovery by centrifugation is an efficient methodology for use with low volumes of fluid and high energy consumption. The disadvantages of centrifugation are high initial investment costs, noise generated during operation and electricity costs (MOLINA-GRIMA et al., 2003).

Gravity sedimentation is widely used for the separation of microalgae in aqueous solution and wastewater treatment. The sedimentation rates of microalgae are influenced by the sedimentation rate of solids and determined by the density and area of the microalgae cells (HONG et al., 2022). Gravity sedimentation preceded by flocculation is one of the most widely used techniques for harvesting microalgae biomass. The disadvantages are the time-consuming method (0.1 to 2.6 m h^{-1}) and the biomass may undergo decomposition under high ambient temperature conditions. In addition, the technique is good for use with large microalgae or with filamentous morphology (REARTE et al., 2021).

Flotation methods are based on the binding of microalgae cells using air bubbles. The resulting flocs rise to the surface of the liquid and are recovered by physical or chemical procedures. Particles as small as $500 \text{ }\mu\text{m}$ can be recovered by flotation. Some strains have gas vacuoles and float on the surface of the water. The incorporation of air bubbles depends on several aspects, such as air, solids, and contact angle with the aqueous phase (FUAD et al., 2021).

6. Final considerations

In recent years, new economic and environmental opportunities have emerged thanks to genetic and process engineering that has stimulated research on microalgae. Innovative ideas emerge in research centers and are implemented on a large scale in industries.

Microalgae biotechnology is making rapid progress in the mass cultivation of microalgae and their application in various industrial processes. However, less than fifty species are used, while thousands remain unexplored. The potential roles of microalgae in genetic engineering and nanotechnology, as well as the examples of entrepreneurial researchers, have increased the prospects for the next generation of establishing biotechnological processes with microalgae to form economically viable companies. Concerted research is needed to (i) develop inexpensive media through the enrichment of wastewater; (ii) isolate and cultivate new varieties of microalgae that present high yields; (iii) improve production systems; (iv) improve biochemical and metabolic pathways through genetic engineering; and (v) improve cultivation techniques.

Furthermore, attention should be paid to high-value natural and recombinant products that can be extracted from microalgae to increase the profitability of commercial operations. Simulation and pilot-scale models will serve as a basis for optimizing processes for wastewater treatment, eutrophication, gas capture, reactor design, and biocompound extraction based on microalgae. A robust bioeconomy built on a technology platform based on innovative microalgal processes requires an interdisciplinary approach involving biologists, biotechnologists, biochemists, engineers, chemists, microbiologists, bioreactor manufacturers, aquaculturists, and modelers.

7. References

- AHMAD, A.; BANAT, F.; ALSAFAR, H.; HASAN, S. W. Algae biotechnology for industrial wastewater treatment, bioenergy production, and high-value bioproducts. *Science of The Total Environment*, v. 806, 2022, 150585.
- ALABI, A.O. *Microalgal Technologies and Processes for Biofuels/Bioenergy Production in British Columbia*. Vancouver: The British Columbia Innovation Council, 2009. p. 1–74.

- BEHERA, B. et al. Integrated biomolecular and bioprocess engineering strategies for enhancing the lipid yield from microalgae. *Renewable and Sustainable Energy Reviews*, v. 148, 2021, 111270.
- BEN-AMOTZ, A. Bioactive compounds: Glycerol production, carotenoid production, fatty acids production. In: *The Alga Dunaliella Biodiversity, Physiology, Genomics and Biotechnology*. (Eds. Ben-Amotz, A., Polle, J.E.W., and Subba Rao, D.V.). Enfield, NH: Science Publishers, 2009. p. 189–207.
- BOSSA, R.; DI COLANDREA, M.; SALBITANI, G.; CARFAGNA, S. Phosphorous utilization in microalgae: physiological aspects and applied implications. *Plants*, v. 13, n. 15, 2024, 2127.
- CHAKRABORTY, B.; GAYEN, K.; BHOWMICK, T. K. Transition from synthetic to alternative media for microalgae cultivation: A critical review. *Science of The Total Environment*, 2023, 165412.
- CHIA, S. R. et al. CO₂ mitigation and phycoremediation of industrial flue gas and wastewater via microalgae-bacteria consortium: possibilities and challenges. *Chemical Engineering Journal*, v. 425, 2021, 131436.
- DE PAUW, N.; MORALES, J.; PERSOONE, G. Mass culture of microalgae in aquaculture systems: Progress and constraint. *Hydrobiologia*, v. 116/117, 1984, p. 121–134.
- ELADL, S. N.; ELNABAWY, A. M.; ELTANAHY, E. G. Recent biotechnological applications of value-added bioactive compounds from microalgae and seaweeds. *Botanical Studies*, v. 65, n. 1, 2024, 28.
- FUAD, N. et al. Harvesting marine microalgae *Nannochloropsis* sp. using Dissolved Air Flotation (DAF) technique. *Sains Malaysiana*, v. 50, n. 1, 2021, p. 73–83.
- GRIMA, E.M.; BELARBI, E.H.; FERNÁNDEZ, F.G.A.; MEDINA, A.R.; CHISTI, Y. Recovery of microalgal biomass and metabolites: process options and economics. *Biotechnology Advances*, v. 20, n. 7–8, 2003, p. 491–515.
- HONG, J. S. et al. Sedimentation and rheological study of microalgal cell (*Chlorella* sp. HS2) suspension. *Biotechnology and Bioprocess Engineering*, v. 27, n. 3, 2022, p. 451–460.
- JACOB-LOPES, E.; FRANCO, T. Teixeira. From oil refinery to microalgal biorefinery. *Journal of CO₂ Utilization*, v. 2, Sep. 2013, p. 1–7.
- JACOB-LOPES, E.; RAMÍREZ-MÉRIDA, L.G.; QUEIROZ, M.I.; ZEPKA, L.Q. Microalgal Biorefineries. In: JACOB-LOPES, E.; ZEPKA, L. (Ed.). *Biomass*. Rijeka: InTech, 2015. p. 81.
- KHOZIN-GOLDBERG, I.; COHEN, Z. The effect of phosphate starvation on the lipid and fatty acid composition of the freshwater eustigmatophyte *Monodus subterraneus*. *Phytochemistry*, v. 67, 2006, p. 696–701.

- LI, S.; TSAI, H. Transgenic microalgae as a non-antibiotic bactericide producer to defend against bacterial pathogen infection in the fish digestive tract. *Fish & Shellfish Immunology*, v. 26, 2008, p. 316–325.
- MANDAL, S.; MALLICK, N. Microalga *Scenedesmus obliquus* as a potential source for biodiesel production. *Applied Microbiology and Biotechnology*, v. 84, 2009, p. 281–291.
- MARKOU, G.; DEPRAETERE, O.; MUYLEAERT, K. Effect of ammonia on the photosynthetic activity of *Arthrospira* and *Chlorella*: A study on chlorophyll fluorescence and electron transport. *Algal Research*, v. 16, June 2016, p. 449–457.
- MASOJÍDEK, J. et al. Photosynthesis monitoring in microalgae cultures grown on municipal wastewater as a nutrient source in large-scale outdoor bioreactors. *Biology*, v. 11, n. 10, 2022, 1380.
- MINHAS, A. K.; GAUR, S.; ADHOLEYA, A. Influence of light intensity and photoperiod on the pigment and lipid production of *Dunaliella tertiolecta* and *Nannochloropsis oculata* under three different culture medium. *Heliyon*, v. 9, n. 2, 2023.
- MOSEY, M.; DOUCHI, D.; KNOSHAUG, E. P.; LAURENS, L. M. Methodological review of genetic engineering approaches for non-model algae. *Algal Research*, v. 54, 2021, 102221.
- MURADYAN, E.A.; KLYACHKO-GURVICH, G.L.; TSOGLIN, L.N.; SERGEYENKO, T.V.; PRONINA, N.A. Changes in lipid metabolism during adaptation of the *Dunaliella salina* photosynthetic apparatus to high CO₂ concentration. *Russian Journal of Plant Physiology*, v. 51, 2004, p. 53–62.
- PADRÓN, R. A. R.; MÉRIDA, L. G. R. Hybrid and nonconventional photobioreactors applied to microalgae production. In: *Algal Bioreactors*. [S. l.], 2025. p. 483–491.
- PALADINO, O.; NEVIANI, M. Airlift photo-bioreactors for *Chlorella vulgaris* cultivation in closed-loop zero waste biorefineries. *Biomass and Bioenergy*, v. 144, 2021, 105926.
- RAMÍREZ MÉRIDA, L. G.; RODRÍGUEZ PADRÓN, R. A. Application of microalgae in wastewater: opportunity for sustainable development. *Frontiers in Environmental Science*, v. 11, 2023, 1238640.
- RAMÍREZ-MÉRIDA, L.; ZEPKA, L.Q.; JACOB-LOPES, E. Current Status, Future Developments and Recent Patents on Photobioreactor Technology. *Recent Patents on Engineering*, v. 9, n. 2, 2015, p. 80–90.
- RAMÍREZ-MÉRIDA, L.G.; ZEPKA, L.Q.; JACOB-LOPES, E. Fotobiorreactor: Herramienta para cultivo de cianobacterias. *Ciencia y Tecnología*, v. 6, n. 2, 2013, p. 9–19.
- REARTE, T. A.; RODRÍGUEZ, N.; SABATTE, F.; DE IORIO, A. F. Unicellular

- microalgae vs. filamentous algae for wastewater treatment and nutrient recovery. *Algal Research*, v. 59, 2021, 102442.
- RESHMA, R. et al. Enhancement of pigments production in the green microalga *Dunaliella salina* (PSBDU05) under optimized culture condition. *Bioresource Technology Reports*, v. 14, 2021, 100672.
- ROJAS-VILLALTA, D. et al. Exploring Extremotolerant and Extremophilic Microalgae: New Frontiers in Sustainable Biotechnological Applications. *Biology*, v. 13, n. 9, 2024, 712.
- SARAVANA, P. S.; UMMAT, V.; BOURKE, P.; TIWARI, B. K. Emerging green cell disruption techniques to obtain valuable compounds from macro and microalgae: a review. *Critical Reviews in Biotechnology*, v. 43, n. 6, 2023, p. 904–919.
- SARKER, N. K.; KAPARAJU, P. A critical review on the status and progress of microalgae cultivation in outdoor photobioreactors conducted over 35 years (1986–2021). *Energies*, v. 16, n. 7, 2023, 3105.
- SATIRO, J.; GOMES, A.; FLORENCIO, L.; SIMÕES, R.; ALBUQUERQUE, A. Effect of microalgae and bacteria inoculation on the startup of bioreactors for paper pulp wastewater and biofuel production. *Journal of Environmental Management*, v. 362, 2024, 121305.
- SEGURA-MORALES, F. J. et al. Unveiling potential of promising filamentous microalga *Klebsormidium* cf. *nitens*: Shear stress resilience and carotenoid-fatty acid dynamics in tubular photobioreactor. *Bioresource Technology*, v. 407, 2024, 131147.
- SEYFABADI, J.; RAMEZANPOU, Z.; KHOEYI, Z.A. Protein, fatty acid, and pigment content of *Chlorella vulgaris* under different light regimes. *Journal of Applied Phycology*, v. 23, 2011, p. 721–726.
- SEYFABADI, J.; RAMEZANPOU, Z.; KHOEYI, Z.A. Protein, fatty acid, and pigment content of *Chlorella vulgaris* under different light regimes. *Journal of Applied Phycology*, v. 23, 2011, p. 721–726.
- SHIVAKUMAR, S. et al. Cell Walls of Lipid-Rich Microalgae: A Comprehensive Review on Characterisation, Ultrastructure, and Enzymatic Disruption. *Fermentation*, v. 10, n. 12, 2024, 608.
- SHOW, P. L. Global market and economic analysis of microalgae technology: Status and perspectives. *Bioresource Technology*, v. 357, 2022, 127329.
- STEPHENS, E. et al. Future prospects of microalgal biofuel production systems. *Trends in Plant Science*, v. 15, 2010, p. 554–564.
- STREIT, N.; RAMÍREZ-MÉRIDA, L. G.; ZEPKA, L.Q.; JACOB-LOPES, E.; QUEIROZ, M.I. Producción de pigmentos naturales (clorofila-a) en biorrefinerías agroindustriales. *Ciencia y Tecnología*, v. 8, n. 2, 2015, p. 27–34.

- TAKAGI, M.; KARSENIO, S.; YOSHIDA, Y. Effect of salt concentration on intracellular accumulation of lipids and triglyceride in marine microalgae *Dunaliella* cells. *Journal of Bioscience and Bioengineering*, v. 101, 2006, p. 223–226.
- TAMBAT, V. S. et al. Sustainable mixotrophic microalgae refinery of astaxanthin and lipid from *Chlorella zofingiensis*. *Bioresource Technology*, v. 387, 2023, 129635.
- WANG, Z. et al. CO₂ gradient domestication produces gene mutation centered on cellular light response for efficient growth of microalgae in 15% CO₂ from flue gas. *Chemical Engineering Journal*, v. 429, 2022, 131968.
- WU, J. et al. Enhancement in lipid productivity of the marine microalga *Nannochloropsis* sp. SCSIO-45217 through phosphate adjustment strategies. *Journal of Applied Phycology*, v. 35, n. 3, 2023, p. 1023–1035.
- YAACOB, N. S.; ABDULLAH, H.; AHMAD, M. F.; MANIYAM, M. N.; SJAHIR, F. Microalgae biotechnology: Emerging biomedical applications. In: *Algal Biotechnology*. [S. l.]: Elsevier, 2022. p. 335–346.
- YAP, J. K. et al. Advancement of green technologies: A comprehensive review on the potential application of microalgae biomass. *Chemosphere*, v. 281, 2021, 130886.
- ZHANG, B. et al. Comprehensive analysis of the combined flocculation and filtration process for microalgae harvesting at various operating parameters. *Science of The Total Environment*, v. 857, 2023, 159658.

Autores

Luis G. Ramírez Mérida

Universidad de Carabobo, Facultad de Ciencia y Tecnología (FACYT),
Departamento de Biología, Bárbula, Venezuela, AP 2005.



www.meridapublishers.com

