

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.0000000000000260. 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