Revista Macambira

Sequenciamento de próxima geração: novas tecnologias para acessar às comunidades microbianas

Djalma Moreira Santana Filho [1]*

RESUMO

Os sequenciadores de próxima geração, do inglês *Next-Generation Sequencing* (NGS), são plataformas capazes de gerar milhões de dados de sequências de DNA ou RNA em uma só rodada através da clonagem *in vitro*, técnica utilizada para incrementar o número de moléculas e permitir a análise. Mas é necessário agrupar informações que deem ao pesquisador condições de decidir qual melhor metodologia adotar. Logo, o objetivo desse artigo foi realizar um estudo de revisão de literatura abordando tecnologias de NGS, compará-las entre si e a métodos mais comuns, mostrando sua importância em microbiologia agrícola e ambiental. O método de Sanger é o principal método de obtenção de sequências conhecido atualmente. Mas, a metodologia adotada pelas plataformas NGS utilizam abordagens que dinamizam o sequenciamento; e o avanço das técnicas vem reduzindo consideravelmente os custos do sequenciamento. Para realizar os estudos foram utilizados dados de artigos e das plataformas das empresas detentoras das tecnologias, realizando-se comparações e análises sobre o tema. Por isso, a pesquisa é de cunho bibliográfico e este artigo caracteriza-se como artigo de revisão de literatura. Várias plataformas de sequenciamento de próxima geração estão disponíveis no mercado, mas são de alto custo e estão restritas a alguns países. Ainda há dificuldade de encontrar algoritmos adequados para analisar a quantidade de dados (sequências) gerados, mas apesar disso os métodos de NGS vêm criando boas expectativas e prometem avanços importantes em estudos com a finalidade de acessar e identificar os grupos que compõem a estrutura da comunidade microbiana em pesquisas ambientais.

Palavras-chaves: Microrganismos. Diversidade. Genética. Dados. Análises.

Next-Generation Sequencing (NGS): new technologies to access microbial communities

ABSTRACT

Next-Generation Sequencing (NGS) are platforms capable of generating millions of DNA or RNA sequence data in a single round through *in vitro* cloning, a technique used to amplify the number of molecules for analysis. However, it is necessary to gather information that enables researchers to determine the best methodology for their work. Therefore, the objective of this article is to conduct a literature review on NGS technologies, comparing them with each other and with more traditional methods, highlighting their importance in agricultural and environmental microbiology. The Sanger method is currently the most well-known sequencing technique. However, NGS platforms employ methodologies that streamline sequencing and the advancement in these techniques have significantly reduced sequencing costs. To carry out the studies, data from articles and platforms developed by companies that own these technologies were used to perform comparisons and analyses on the topic. Therefore, this manuscript is a bibliographic research and characterized as review article. While several second-generation sequencing platforms are available on the market, most are expensive and restricted to a few countries. There is still difficulty in finding suitable algorithms to analyze the amount of data generated. Nevertheless, NGS methods have generated good expectations and promise important advancements in studies aimed at accessing and identifying the groups that make up the microbial community structure in environmental research.

Keywords: Microorganisms. Diversity. Genetic. Data. Analysis.

Secuenciación de Próxima Generación: nuevas tecnologías para acceso a las comunidades microbianas

[1]Graduado em Engenharia Agronômica (UFBA). Mestrado em Microbiologia Agrícola (UFRB). Doutor em Ciências Agrárias (UFRB). Servidora do Instituto Federal de Educação, Ciência e Tecnologia Baiano Campus Itapetinga, Bahia. ORCID: https://orcid.org/0000-0002-9577-5938. Lattes: https://lattes.cnpq.br/5740532819871527. *Autor correspondente: dj.msfilho@gmail.com.

ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

RESUMEN

Los secuenciadores de próxima generación, del inglés Next-Generation Sequencing (NGS), son plataformas capaces de generar millones de datos de secuencias de DNA o RNA en un solo circuito a través de la clonación in vitro, técnica utilizada para incrementar el número de moléculas y permitir el análisis. Pero hay que agrupar informaciones que ofrezcan al investigador condiciones de decidir cual mejor metodología a adoptarse. Luego, el objetivo de ese artículo fue realizar un estudio de revisión de literatura abordando tecnologías de NGS, compararlas entre si y a métodos más comunes, mostrando su importancia en microbiología agrícola y ambiental. El método de Sanger es el principal método de obtención de secuencias conocido actualmente. Sin embargo, la metodología adoptada por las plataformas NGS utilizan abordajes que dinamizan la secuenciación; y el avance de las técnicas vienen reduciendo considerablemente los costos de la secuenciación. Para realizar los estudios fueron utilizados datos de artículos y de las plataformas de las empresas detentoras de las tecnologías, realizándose comparaciones y análisis sobre el tema. Por eso, la investigación es de carácter bibliográfico y este artículo se caracteriza como artículo de revisión de literatura. Varias plataformas de secuenciación de próxima generación están disponibles en el mercado, pero son de alto costo y están restringidas a algunos países. Aún hay dificultad de encontrar algoritmos adecuados para analizar la cantidad de datos (secuencias) generados, pero, a pesar de eso, los métodos de NGS han creado buenas expectativas y prometen avances importantes en estudios con la finalidad de acceso e identificación de los grupos que componen la estructura de la comunidad microbiana en investigaciones ambientales.

Palabras-clave: Microorganismos. Diversidad. Genética. Datos. Análisis.

INTRODUÇÃO

O NGS se baseia na clonagem *in vitro* capaz de gerar milhões de pares de bases em uma única rodada. Essas tecnologias eliminam as etapas de produção de clones em laboratório, de montagem de placas de sequenciamento e de separação de fragmentos em géis. Consequentemente a quantidade de informações gerada através do NGS é, muitas vezes, maior que o sequenciamento de Sanger, com um menor custo por base e em menor espaço de tempo (Hu *et al*., 2021; Satam *et al*., 2023).

O NGS apresenta um sequenciamento massivo em paralelo, realizado através de equipamentos com diferentes técnicas de abordagens, tais como o pirossequenciamento com PCR em emulsão, sequenciamento por síntese com bridge PCR, sequenciamento por ligação, sequenciamento por detecção de H^+, e o sequenciamento de molécula única. Algumas tecnologias ainda estão em desenvolvimento ou em fase de testes, a exemplo do sequenciamento em nanoporos e de imagens da Polonator (Hu *et al*., 2021; Satam *et al.*, 2023; Athanasopoulou *et al*., 2021; Allen *et al*., 2022; Wang *et al.*, 2021; Suminda; Ghosh; Son, 2022; Zhang *et al.*, 2021).

Apesar dessa geração de sequenciadores massivos englobarem equipamentos mais potentes, que conseguem sequenciar um número maior de reads, com maior velocidade utilizando o processamento em paralelo de amostras; a metodologia de Sanger ainda é muito utilizada para obtenção de informações biológicas de indivíduos. Entretanto, a possibilidade de aperfeiçoamento das abordagens dos sistemas NGS tem proporcionado mais aplicações às técnicas, e diminuído muito o custo de trabalho. Mas, as vantagens dos sistemas NGS não os eximem de três problemas: geração de leituras de sequências curtas; erros de inserção; deleção

ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

e substituição de bases; inexistência de um algoritmo capaz de processar todos os dados gerados (Hu *et al*., 2021; Satam *et al*., 2023).

Os estudos de bioinformática que abordam métodos de montagem de genomas são importantes para que as plataformas tenham seu potencial totalmente explorado. A montagem e alinhamento das sequências em função dos fragmentos obtidos nas plataformas baseadas no pirossequenciamento é atualmente realizada através do *Overlap Layout Consensus* (OLC). Os dados gerados pelas plataformas de sequenciamento baseados nas abordagens de síntese com bridge PCR ou por ligação é realizada pelo *De Bruijn graph*. Ainda assim, as plataformas NGS são muito utilizadas e apresentam bons resultados para análises de comunidades microbianas (Wang *et al*., 2021).

O objetivo deste trabalho foi realizar uma revisão bibliográfica sobre as novas tecnologias de sequenciamento em plataformas de próxima geração, comparar com os métodos mais difundidos atualmente e mostrar a importância da aplicação dessas tecnologias em estudos de microbiologias agrícola e ambiental.

REFERÊNCIAL TEÓRICO

As técnicas de NGS tornaram-se ferramentas importantes em estudos que envolvem biologia molecular. Essas tecnologias são responsáveis por incrementos significativos na quantidade e na qualidade de dados, e consequentemente de informações sobre microrganismos, principalmente nos âmbitos ambiental e médico (Hu *et al*., 2021).

As ocorrências mais importantes na evolução dos métodos moleculares de estudo e identificação de organismos foram a descoberta da molécula de DNA na célula (Zhang *et al*., 2011), o seu estabelecimento como molécula responsável pela transferência de material genético celular, sua composição formada por quatro bases nitrogenadas, e sua conformação em dupla hélice (Liu *et al*., 2012; Lander *et al.*, 2001). Ainda no século XX, o desenvolvimento do método de Sanger de sequenciamento e da Reação da Cadeia da Polimerase (PCR), desenvolvida por Kary Mullis, permitiu obter a sequência completa de nucleotídeos do vírus (bacteriófago) chamado de phiX174 e de outras espécies biológicas (Sanger *et al*., 1977; Sanger *et al.*, 1978; Sanger *et al*., 1980).

O primeiro sequenciador automático de DNA foi o ABI 370 lançado em 1986. Esse equipamento foi aperfeiçoado e reapresentado como ABI 3700, com a adoção da eletroforese capilar, capaz de detectar 96 bases/h, obtendo 500Kb/dia, com comprimento de leitura de 600 bases (Liu *et al*., 2012). Os sequenciadores de próxima geração chegaram ao mercado em 2005, e se diferenciam do método de Sanger devido ao grande volume de fragmentos obtidos. Essas

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

plataformas desencadearam novos insights sobre variação funcional, associação genética e seleção natural nos seres vivos (Durbin *et al*., 2010). Com o lançamento do projeto genoma humano houve a necessidade de sequenciar uma grande quantidade de fragmentos para sua montagem. As possibilidades de automação transformaram as plataformas de NGS em ferramentas moleculares importantes em estudos genéticos (Olson, 1993).

Em suma, a primeira geração de sequenciadores envolve o método de terminação das cadeias, com a utilização de didesoxiribonucleotídeosfosfatos (ddNTPs) para interromper a montagem dos clones através da enzima Taq Polimerase. A segunda geração foi marcada pelo pirossequenciamento com PCR em emulsão, do sequenciamento por síntese com terminadores reversíveis e Bridge PCR e pelo sequenciamento por ligação utilizando uma enzima ligase ao invés de uma DNA polimerase. Tanto para a primeira como para a segunda geração, a reação de PCR ainda é necessária para o sequenciamento. A terceira geração de sequenciadores é marcada pelo surgimento do sequenciamento de molécula única, sem a necessidade de reações de PCR. O sequenciamento em nanoporos poderá ser a tecnologia de base para a quarta geração, mas como a tecnologia de sequenciamento ainda é considerada a mesma e está em fase de teste, ela tem sido considerada de terceira geração (Wang *et al*., 2021).

A primeira geração se baseia no método de Sanger, cujo princípio básico é a interrupção da montagem de fitas simples de DNA devido à ausência de um grupo ligante compatível a adição, que ocorre por ligação, de um novo nucleotídeo na extremidade 3'. O método de Sanger envolve algumas etapas de preparação da amostra conhecida como reação da cadeia da polimerase (PCR), utilizada para duplicar, amplificar ou sintetizar as moléculas de DNA de uma amostra purificada. É uma reação de aumento do número de fitas de DNA para um posterior sequenciamento (Zhang *et al.*, 2011).

A reação de PCR inicia-se com a desnaturação do DNA, que consiste na separação de suas fitas através do aumento da temperatura. Primers ou oligonucleotídeos são colocados juntos à amostra de DNA desnaturado e anelam-se às fitas simples do mesmo quando a temperatura do sistema é reduzida. Como os primers são oligonucleotídeos, eles conseguem rapidamente anelar nas extremidades das fitas simples, antes que elas se juntem e renaturem. Feito isso, inicia-se a fase de elongação da cadeia, realizada pela DNA polimerase ou Taq polimerase. Nesse processo, moléculas de desoxiribonucleotídeos são adicionadas ao sistema, e a DNA polimerase promove a ligação desses resíduos ao primer anelando-os à fita simples, e assim duplicando a quantidade inicial dessas fitas (Sanger *et al*., 1977; Sanger *et al.,* 1978; Sanger *et al.*, 1980).

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

Quando a amostra se encontra em quantidade suficiente de fitas moldes, são adicionados dideoxinucleotídeos ao sistema, os quais ao serem adicionados à cadeia pela Taq polimerase interrompem a elongação da cadeia por não apresentarem capacidade de realizar as ligações fosfodiéster com os próximos dNTPs (Sanger *et al.*, 1977). A interrupção da cadeia através desses terminadores justifica o nome dado à técnica, de método da terminação da cadeia (Zhang *et al.*, 2011).

Posteriormente, as amostras são submetidas a uma eletroforese em gel de poliacrilamida, e a sequência de nucleotídeos é determinada através da posição de cada ddNTP incorporado à fita molde. Quando o sequenciamento é automático, os resíduos são marcados na extremidade com substâncias fluorescentes (cada nucleotídeo é marcado com uma cor diferente) e incorporados às fitas complementares sintetizadas. As bases são excitadas e um laser as identifica e grava em eletroferogramas, nos quais cada pico representa uma base (Sanger *et al.*, 1977).

Na segunda geração, as plataformas se baseiam no processamento em paralelo de amostras e consequente obtenção massiva de dados, ou seja, as amostras de DNA podem ser sequenciadas todas ao mesmo tempo. Existe a necessidade de reações prévias de PCR para a preparação das amostras ou dos dados de entrada (fragmentação - adição de adaptadores - amplificação) e as máquinas de sequenciamento apresentam dois métodos de trabalho: PCR em emulsão – emPCR - e em pontes - Bridge PCR (Hu *et al*., 2021).

O pirossequenciamento foi a abordagem precursora dessas plataformas e baseia-se na emissão de sinais luminosos pela conversão de pirofosfato – PPi. Além dele, a adição de resíduos ao molde pode ocorrer pelo método de síntese através de terminadores reversíveis por PCR em pontes, onde os resíduos são marcados com substâncias fluorescentes e emitem sinais quando são adicionados por ligação ao molde. Nesse caso, dependendo da plataforma, ao invés da enzima Taq pol é utilizada uma ligase ou íons condutores H+ liberados da reação de adição do resíduo à fita, medidos por um mini medidor de pH, para adição da base ao molde (Allen *et al.*, 2022).

Outras diferenças entre as plataformas estão no volume de sequências de DNA, na precisão de bases, na tecnologia de operação, no tamanho do fragmento gerado por leitura, e principalmente no custo operacional. Desde os primeiros sequenciadores lançados com essas abordagens, plataformas que apresentam alto rendimento, e possibilidade de melhorias no desempenho com adequação a cada necessidade são as que mais se destacam (Andreoto, 2011).

A Roche 454 sequencing foi a primeira a lançar um sequenciador de próxima geração comercializado com sucesso (Loman *et al.*, 2012). Seus sequenciadores utilizam como método

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

o pirossequenciamento, com emPCR. Nessa abordagem, esferas de ligação de DNA de cadeia simples são encapsuladas por forte agitação em micelas aquosas contendo reações de PCR imersas em óleo para a amplificação dos fragmentos (Zhang *et al*., 2011; Margulies *et al.*, 2005; Allen *et al*., 2022).

Nessa reação de PCR, o DNA previamente fragmentado recebe a adição de adaptadores (sequências curtas conhecidas) nas extremidades. Esferas denominadas *beads*, ou *IonSpheres* no caso do *ION TORRENT*, recebem milhares de adaptadores complementares àqueles adicionados aos *templates*, os quais servirão de fixadores dos mesmos e ainda como *primer* para uma PCR. As esferas são colocadas em contato com os *templates* de forma tal que, em uma situação ideal, ao menos um deles se ligue ao adaptador complementar presente na mesma. Em seguida, as esferas são colocadas em uma emulsão contendo todos os reagentes necessários para a reação de cadeia da polimerase (dNTP, Taq pol, ATP sulfurilase, apirase) (Margulies *et al.*, 2005).

O sistema funciona como microrreatores cuja cada esfera é envolvida por uma gota da emulsão, e dentro dela a reação de amplificação do DNA é realizada de forma cíclica, com a fita acoplada a esfera sendo sintetizada, em seguida desnaturada e acoplada a outro adaptador da mesma esfera. No sequenciamento, as amostras são colocadas em placas de cultura de células dimensionadas para que cada poço receba uma única esfera. As sequências ligadas às esferas recebem primers e a DNA polimerase inicia o sequenciamento adicionando as bases complementares às sequências modelo (Liu *et al.*, 2012). Segue-se a isso o pirossequenciamento, técnica apresentada à comunidade científica em 1996, que se baseia na liberação do PPi durante a adição de deoxiribonucleotídeos (dNTP) à cadeia molde pela Taq polimerase. Essa adição de dNTP realizada pela Taq pol à fita molde de DNA acarreta a liberação de moléculas de PPi. O PPi é convertido em ATP pela ATP sulfurilase, e em seguida o ATP é convertido em ADP e fornece energia para a luciferase oxidar a luciferina à oxiluciferina e gerar luz (Ravin, 2010).

Os nucleotídeos são adicionados ciclicamente e os que sobram do processo são removidos pela enzima apirase antes de se iniciar um novo ciclo. As luzes são detectadas em tempo real por um chip e gravadas nos eletroferogramas, cada base emite um comprimento de onda de luz diferente, o que torna possível a sua identificação ao ser adicionada na fita molde. Picos duplos ou triplos significam duas ou três bases iguais adicionadas. Essa técnica pode apresentar erros na identificação das bases na presença de homopolímeros, ou seja, quando uma mesma base se repete várias vezes (Ronaghi *et al*., 1996; Alderborn; Kristofferson; Hammerling, 2000).

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

A extensão do primer utilizando a abordagem do pirossequenciamento pode ser conduzida em modelos de fase líquida ou sólida (Alderborn; Kristofferson; Hammerling, 2000). O de fase sólida oferece a vantagem de rapidez devido à redução no tempo do ciclo de adição de cada nucleotídeo. Não há acumulação de inibidores porque, após cada ciclo, é feita uma lavagem, há a possibilidade de usar baixas quantidades de enzimas para reduzir o custo, há integração entre a amplificação do PCR, preparo de moldes, e a análise do sequenciamento, com tudo sendo feito em um único sistema de fluxo (Ronaghi, 2001).

O sequenciamento por síntese com PCR em pontes é a base das plataformas da Illumina**.** O processo completo envolve a preparação das amostras, geração de clusters (agrupamentos), sequenciamento, e análise dos dados (Zhang *et al.*, 2021). Na preparação das amostras, a fita dupla de DNA extraído e purificado recebe a adição de adaptadores terminais através de transposomes. Em seguida, ela é desnaturada e os primers anelam às fitas simples e realizam um ciclo reduzido de amplificação e então os templates são fixados em um suporte sólido, geralmente de vidro. As superfícies são providas de oligonucleotídeos complementares aos adaptadores das fitas simples. Cada fita simples é então imobilizada no suporte por uma de suas extremidades e inicia-se a clusterização (Liu *et al*., 2012; Andreoto, 2011; Zhang *et al.*, 2021).

A clusterização ou agrupamento é o processo cujo cada fragmento é isotermicamente amplificado. As células de fluxo (*flow cells*) são lâminas de vidro formadas por linhas, cada uma correspondendo a canais compostos por dois tipos de oligonucleotídeos complementares, àqueles adaptadores utilizados para marcar as fitas simples de DNA. A fita simples de DNA é fixada através do acoplamento do adaptador ao oligonucleotídeo complementar da superfície da *flow cell* e, então, a *Taq pol* inicia a produção da fita simples complementar (reversa) e o *template* é lavado. A fita que permanece anelada à superfície inclina e anela sua outra extremidade (com adaptador, que encontra o seu oligonucleotídeo complementar) formando uma estrutura semelhante a uma ponte. A polimerase então inicia a formação de uma nova fita, complementando a estrutura de ponte dupla (dupla fita). Ocorre desnaturação das fitas duplas, gerando-se uma fita simples e outra fita reversa. O processo se repete várias vezes até haver a formação de milhares de *clusters* ou agrupamentos. Posteriormente, as fitas reversas são clivadas e novamente retiradas do sistema por uma lavagem, permanecendo apenas fitas acopladas a um único tipo de adaptador - árvore dos *clusters* (Liu *et al*., 2012; Andreoto, 2011; Zhang *et al*., 2021).

O sequenciamento inicia-se quando o primeiro *primer* anela ao *template* e produz a primeira leitura. Em cada ciclo, os quatro tipos de dNTPs marcados com substâncias fluorescentes identificadoras de cada base são adicionados ao sistema. A cada nucleotídeo que

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

a DNA pol incorpora à fita uma cor é emitida pelo seu respectivo marcador, isso é gravado em um chip que traduz o sinal na base correspondente. Após a incorporação da base e consequente emissão do sinal fluorescente, os marcadores são clivados e uma nova base será adicionada. Milhões de sequências são produzidas em um processo massivo e em paralelo. No sequenciamento por síntese, os nucleotídeos são marcados com corantes fluorescentes e capas terminais. A base emite um sinal fluorescente ao ser adicionada ao *primer* que é captado por um sistema que a identifica. O marcador e a capa terminadora são clivados do resíduo logo após a emissão de sinal luminoso (Liu *et al*., 2012; Andreoto, 2011; Zhang *et al*., 2021).

A plataforma *SOLiD*, cujo nome foi dado baseando-se nas iniciais de *Sequencing by Oligo Ligation Detection*, apresenta a abordagem de sequenciamento por ligação. Utiliza reações de emPCR em que os resíduos são ligados ao *template* pela enzima ligase ao invés da DNA polimerase. Ao invés de detectar a incorporação das bases pela *Taq pol*, a plataforma utiliza oligonucleotídeos de 8 bases marcadas com compostos fluorescentes nas últimas três bases, que são inosinas aneladas ao *template* (Satam *et al*., 2023; Liu *et al*., 2012; Zhang *et al*., 2011).

A biblioteca é construída adicionando-se um *primer* (A e B) a cada extremidade do fragmento. Em seguida, um conjunto de oligonucleotídeos de oito bases é adicionado e anelam-se ao *template*. Primers complementares (a) à extremidade (A) são adicionados e anelam-se as mesmas. Depois disso, uma ligase adicionada junto aos oligonucleotídeos de oito bases anela-os ao primer complementar de (A) - sentido 5’ - 3’. Uma lavagem remove os oligonucleotídeos não adicionados, a leitura do fluoróforo acoplado as inosinas (três últimas bases) é realizada antes de sua clivagem, e a extremidade é liberada para a ligação com o próximo oligonucleotídeos de oito bases (Ravin, 2010).

Terminado o ciclo de adição das sondas (no mínimo cinco ciclos, e o total depende do tamanho do template ou tipo de sequenciamento), a fita complementar ao *template* é desnaturada pelo aumento da temperatura do sistema e em seguida eliminada. Um novo *primer* (a’), que alinha uma base à esquerda de (A), é adicionado e o processo se repete. A cada ligação das sondas duas bases são lidas, e cada base é lida por duas sondas diferentes (a e a’). Esse sistema, denominado de *Two Bases Ecoding*, permite uma posterior correção dos erros para aumentar a precisão dos dados. Nesse sistema, ao saber a primeira base do par e a cor, uma tabela de cores e números indicará quem será a segunda base (Satam *et al.*, 2023; Liu *et al*., 2012; Zhang *et al*., 2011, Ravin, 2010).

Em 2011, Geeorge Watts desenvolveu uma técnica capaz de aperfeiçoar a região do *amplicon* sequenciada de genes 16S rRNA, para melhorar a discriminação em níveis de espécies

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

do perfil. Surge a técnica de sequenciamento pós-síntese baseada na detecção de íons da plataforma ION Torrent, que utiliza a reação de em PCR (Life Technologies Corporation, 2013). O sistema é formado por um chip de sílica composto por inúmeros poros nanoscópicos, com a largura equivalente a de um fio de cabelo, onde o DNA será decifrado. No sequenciamento, cada nanopoço recebe uma fita de DNA molde que será completada através da adição de um determinado nucleotídeo (dNTP) através da DNA pol adicionada junto do DNA molde. No momento que a *Taq pol* encontra a base compatível no poço e adiciona o nucleotídeo ao DNA template, são liberados íons H^+ que alteram a voltagem do chip. Essa alteração é detectada por um sensor de íons hipersensível, uma espécie de mini peagâmetro, que indica a ocorrência da reação. A série de impulsos elétricos gravados pelo chip é traduzida na leitura do DNA sequenciado (Shokralla *et al.*, 2012).

Nesse sistema, se o nucleotídeo introduzido não for complementar a base do molde não haverá mudança de tensão. No caso de homopolímeros na sequência, múltiplas bases serão incorporadas ao template em um único ciclo, e o sinal emitido pelo chip será proporcional ao número de H^+ liberados nas reações de adição (Loman *et al*., 2012; Miller; Koren; Sutton, 2010).

A plataforma *Polonator* da *Danaher/Dover/Azco* utiliza a tecnologia de sequenciamento por ligação com registro de imagens. É uma abordagem em desenvolvimento que se baseia no sequenciamento por síntese, com emPCR para amplificar fragmentos de DNA. Nessa plataforma de sequenciamento, há uma fase denominada de pré-ciclo, na qual ocorre a segmentação do material, em seguida, ocorrem 26 ciclos de registro de imagens, extração e identificação das bases para o sequenciamento. Os sistemas com esse tipo de abordagem de sequenciamento através de registro de imagens podem apresentar uma ou quatro câmeras, com capacidade de registrar imagens em intervalos de 60 milissegundos (Sumida; Ghosh; Son, 2022).

Os sequenciadores da terceira geração apresentam a tecnologia de sequenciamento de molécula única através de três abordagens: a *Single Molecule Real Time* (SMRT da *Pacific Bioscience*), a *True Single Molecule Sequencing* (tSMS da *Heliscope Sequencing*) e o sequenciamento de molécula única em nanoporos (Oxford Nanopore Technologies, 2016). Com a abordagem de sequenciamento de molécula única, as sequências são obtidas por síntese, sem a necessidade de ligação, amplificação via PCR, ou qualquer procedimento de preparação de bibliotecas. As moléculas são sequenciadas de forma independente. Pode realizar o sequenciamento completo de genomas pequenos, como os bacterianos, sem experimentos adicionais (Roberts; Carneiro; Schartz, 2013). Essas vantagens têm sugerido que a técnica seja

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

chamada de next-NGS (Zhang *et al.*, 2011). Pode ser utilizado em laboratórios clínicos, principalmente em pesquisas em microbiologia (Liu *et al*., 2012).

Essas técnicas podem dar uma descrição mais detalhada das comunidades microbianas, através de uma visão mais profunda do genoma dos microrganismos, possibilitando a realização de estudos mais detalhados da ecologia e evolução de comunidades microbianas complexas (Zhang *et al*., 2011; Roberts; Carneiro; Schartz, 2013). As técnicas também melhoram a compreensão sobre os microrganismos e sua atuação nas comunidades, com a possibilidade da obtenção de genomas completos, da caracterização de modificações epigenéticas, de investigar patogenicidade, evolução e seleção viral, e de entender um pouco mais sobre metagenômica e sobre essas comunidades microbianas, permitindo também a descoberta de novos genes e isoformas transcritos.

A síntese química da SMRT funciona com o auxílio de um chip formado por milhares de medidores de onda, conhecidos como *waveguides*, cada um apresentando milhares de pequenos poços denominados de *Zero-Mode Waveguides* (ZMWs), que funcionam como *containers* dotados de ferramentas de captura acopladas no fundo dos poços, que recebem uma iluminação de fundo de comprimento de onda suficiente para que a luz os atravesse. Uma luz branda da excitação das esferas penetra 20-30 cm de cada ZMW, criando um microscópio capaz de detectar volumes de 20 zeptolitros (10^{-21} litros). No sequenciamento, o complexo *template* - DNA polimerase é imobilizado no fundo do ZMW, cujos nucleotídeos marcados na terminação fosfato com fluoróforos de cores específicas para cada base (dNTPs) são introduzidos no *template*. Quando a base é detectada nos poços, um pulso de luz é gerado e a base é detectada pela cor do marcador, com a sequência sendo encontrada em função da fluorescência emitida por cada nucleotídeo. Após a incorporação da base, o fosfato marcado com o fluoróforo é clivado. Isso ocorre em paralelo nos milhares de poços ZMWs que formam as células SMRT (Rhoads; Au, 2015).

Na tecnologia de molécula única por tSMS, são utilizados “flow cells” com adaptadores, cujos fragmentos de DNA com cauda poli-A são adicionados. Os *primers*, a DNA polimerase, e os nucleotídeos são adicionados a um suporte de vidro. Os fragmentos de ácidos nucléicos são hibridizados com os primers e ancorados em posições preestabelecidas, em placas de vidro, numa “*flow cell*”. As leituras são curtas, com média de 35pb, e são realizadas por sequenciadores Heliscope (Pettersson; Lunderberg; Ahmadian, 2009). O sequenciamento engloba os passos de síntese dos *templates*, seguido de lavagem das bases não adicionadas à fita molde, logo depois, os marcadores fluorescentes emitem o sinal que é gravado e identificam a base que foi adicionada, e então a ligação entre o marcador e a base é quebrada e o processo

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

recomeça, de forma massiva e em paralelo. Nas transcrições raras, não há polarização, perda de representação ou sub-representações. A probabilidade de adesão das moléculas à superfície da *flow cell* é igual para todas (Lemos *et al*., 2011).

As tecnologias baseadas em sequenciamento de molécula única em nanoporos baseiam-se no trânsito de moléculas de DNA e suas bases por poros nanoscópicos e a consequente interrupção da corrente elétrica ou sinal óptico; revelam conformação, comprimento, e diâmetro do DNA através de modulação da corrente iônica que atravessa o poro. Nessa abordagem, os nucleotídeos do DNA obstruem a passagem de corrente elétrica nos nanoporos quando clivados individualmente da molécula e são identificados de acordo com as características dessa interrupção, pois cada base interfere na corrente de forma característica (Lemos *et al*., 2011; Schadt; Turner; Kasarskis, 2010).

Estudos estão sendo realizados no intuito de entender melhor as interações que ocorrem entre o DNA e os nanoporos e o processo de translocação no sistema, que contribuirá para o desenvolvimento de aparelhos (chips) mais adequados (Ramachandran *et al.*, 2009). Assim como para a maioria dos sequenciadores NGS, trabalhos futuros são requeridos para determinar o algoritmo apropriado para análises secundárias (Quick; Quinlan; Loman, 2014).

METODOLOGIA

Neste artigo, foi utilizada a metodologia de pesquisa bibliográfica. Ela tem o caráter qualitativo e descritivo de conhecimentos através da análise e interpretação dos dados disponíveis (Koche, 2015, p.122). Tem como objetivo informar, argumentar e discutir sobre as novas tecnologias que vêm revolucionando os estudos de biologia molecular.

A análise de dados de fonte primária foi a estratégia utilizada para obter as informações e discutir as formas com as quais o sequenciamento de DNA ou de RNA, utilizando ferramentas NGS, poderá contribuir para os estudos em microbiologia agrícola e ambiental. As fontes primárias de informações técnicas se resumiram a artigos publicados em bases indexadas, dentre elas *PubMed, Web of Science, Scopus, SciELO*, e em catálogos disponibilizados por fabricantes de sequenciadores de próxima geração, bem como dissertações e teses disponibilizadas em repositórios institucionais.

Na primeira seleção, baseando-se na leitura do título e resumo, foram selecionados cerca de 150 trabalhos acadêmicos. Os dados técnicos das plataformas foram obtidos em sites das empresas que desenvolveram as tecnologias. A metodologia exigiu a escolha adequada de materiais de referência, implicando diretamente na qualidade das análises realizadas, com interpretações e conclusões mais seguras para responder o problema da pesquisa.

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

A segunda seleção foi realizada através do levantamento e organização de dados da leitura e do fichamento de textos completos utilizando o software Mendeley, uma ferramenta de gerenciamento de referências bibliográficas interessante, pois apresenta ferramentas com funções análogas ao

> [...] ato de registrar em fichas, cartões, folhas avulsas, cadernos, arquivos de texto ou qualquer outro recurso físico ou eletrônico de armazenamento de dados às informações relevantes de uma fonte bibliográfica ou documental, tomada em parte ou integralmente, na forma de esquemas, resumos, paráfrases ou citações (Rauen, 2018, n.p.).

Essa técnica ajudou no registro das informações mais aderentes ao tema. Além disso, serviu de base de informações para consultas durante a escrita desse artigo científico.

Os artigos foram as principais fontes de informações deste trabalho e a seleção dos mesmos foi norteada pelo objeto de pesquisa proposto. A busca restringiu-se a manuscritos que tratam sobre a descrição de plataformas NGS e suas aplicações em estudos de microbiologia ambiental ou agrícola. Nesse contexto, foram utilizados artigos de revisão de literatura e artigos científicos de periódicos que utilizam o processo de revisão por pares conhecido como *peer review* ou *refereeing*. As informações obtidas serviram para confrontar os dados dos catálogos publicadas no site dos fabricantes de sequenciadores. O tempo desde o desenvolvimento das primeiras técnicas de estudos de biologia molecular até o surgimento das tecnologias NGS foi o espaço amostral adotado. Assim, as informações obtidas consideraram o contexto geral sobre o sequenciamento de DNA, comparando passado e futuro, bem como a acessibilidade para as áreas agrícola e ambiental.

ANÁLISES E RESULTADOS

A literatura mostra que as primeiras técnicas de estudos de biologia molecular, aliadas ao método de Sanger, foram utilizadas para amplificar fragmentos do gene 16S rRNA e sequenciar o genoma de novas espécies, a exemplo da bactéria *Facklamia tabacinasalis sp.* nov. (Collins *et al*., 1999). Assim como essa bactéria, outros microrganismos foram identificados graças ao método de Sanger (Tabela 1). Esse é o método mais acessível, atualmente, à microbiologia agrícola e ambiental para realizar o sequenciamento do DNA, principalmente no sequenciamento de indivíduos dependente de cultivo para investigações de metabolismo, função e atuação deles em diferentes meios. Além disso, o método tem boa resposta em pesquisas sobre a distribuição e papel da comunidade microbiana em diversos tipos de bioprocessos (Felske; Akkermans; Vos, 1998).

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

Tabela 1. Primeiros indivíduos com genomas completamente sequenciados e seus respectivos grupos.

Grupo	Espécie
Bactéria	*Haemophilus influenza* (Meningite)
Levedura	*Saccharomyces cerevisiae* (Fermentação)
Nematoide	*Caenorhabditis elegans*
Protozoário	*Plasmodium falciparum* (malária)

Fonte: Elaboração baseada em Liu *et al.* (2012).

O avanço tecnológico proporcionou ao método de Sanger a capacidade de obter leituras de sequências de grandes de DNA. Entretanto, para sequenciá-las é necessário cortar e remontar fragmentos. Isso é possível devido ao desenvolvimento do sequenciamento *Shotgum* durante o projeto genoma humano. Basicamente, o DNA é quebrado, enzimática ou mecanicamente, em pequenos fragmentos e clonado em vetores de sequenciamento em que fragmentos de DNA clonados podem ser individualmente sequenciados (Standen, 1979).

A evolução dos equipamentos de sequenciamento na obtenção de respostas de alto rendimento, aliada a redução dos custos das operações, permitiu o surgimento de equipamento capazes de realizar trabalhos mais rápidos e objetivos. Os instrumentos podem dinamizar os estudos de associação de genótipos, mutações, evolução e de perfil ambiental da comunidade microbiana (Hall, 2007; Ansorge, 2009).

Os sequenciadores NGS apresentam roupagens com sistemas de análise massivo em paralelo, e trazem vantagens como maior velocidade, precisão, maior rendimento, e custos compatíveis com o sequenciamento de Sanger. A desvantagem da obtenção de fragmentos de baixo comprimento e das exigências de infraestrutura informatizada complexa são minimizadas pela ampliação de possibilidades de aplicação e redução do custo por base a longo prazo. Entretanto, ainda existem dificuldades para encontrar softwares que analisem esses dados com a mesma velocidade em que eles são gerados (Ansorge, 2009).

Sequenciadores que utilizam as abordagens de pirossequenciamento com emPCR, de terminadores reversíveis com síntese por Bridge PCR, e de detecção de íons são capazes de gerar dados precisos para montar um rascunho de um genoma bacteriano em alguns dias. O sequenciamento por ligação apresenta os menores comprimentos de leitura, entretanto, apresenta boa precisão (Hu *et al.*, 2021).

As plataformas que utilizam o pirossequenciamento evoluíram ao longo dos anos, melhorando características como o comprimento de leituras, o rendimento e otimizando o tempo de trabalho (Tabela 2) (Loman *et al.*, 2012). As ferramentas desse sistema possibilitam uma amostragem mais ampla da diversidade de populações microbianas, gerando grandes

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

quantidades de dados em curto espaço de tempo. Dispensa o uso de clonagem convencional para a preparação de moldes, incrementando assim a eficiência do sequenciamento de DNA. Sua maior desvantagem é a alta taxa de erros de inserção, deleção e substituição de bases, mas as melhorias nas plataformas podem elevar a qualidade dos dados a 99,9% (Huse *et al.*, 2007).

Tabela 2. Comparativo entre as Plataformas de Sequenciamento (PS) da Roche 454 Sequencing, para o sequenciamento Shotgum (*) e de Amplicons (**). Comprimento de Leitura (CL) em pares de bases, Rendimento (R) em Megabases, Número de Leitura por Rodada (NLR), Acurácia (A) em porcentagem, tempo (t) em horas. Dados obtidos em http://454.com/index.asp.

Plataformas de sequenciamento Roche 454					
PS	**CL (pb)**	**R (Mb)**	**NLR**	**A (%)**	**T (h)**
GS Junior	400	35	100.000* 70.000**	99	10 h
GS Junior +	700	70	100.000* 70.000**	99	18 h
GS FLX LR 70	↑ 600	450	1.000.000* 700.000**	100	10 h
GS FLX XL +	↑ 1000	700	1.000.000*	99,997	23 h

Fonte: Elaboração baseada nos dados disponíveis em: https://web.archive.org/web/20120829172226/http://www.roche.com/med-cor-2007-03-29.

A metodologia de sequenciamento por síntese com Bridge PCR é muito utilizada por estar presente em várias plataformas. Mas só algumas são aplicáveis às pesquisas nas áreas de agrigenômica e genômica microbiana (Tabela 3). As escolhas das plataformas Illumina devem ser realizadas em função dos objetivos e alvos do sequenciamento (SEQUENCING PLATFORMS, 2014).

Tabela 3. Comparativo entre características (C) dos sistemas da Illumina. Tamanho dos dados (D) em Gigabytes, Tempo por rodada (T) em horas, Reads por rodada em milhões (R), Leitura máxima (L) em pares de bases.

Dados dos Sistemas Illumina				
C	**MiSeq**			**NextSeq**
	MiSeq	MiSeqDx	MiSeq FGx	NextSeq 500
D (Gb)	0,3 - 15	0,3 - 15	0,3 - 15	20 – 120
T (h)	5 - 55	4 - 55	4 - 55	11 – 29
R (milhões)	1 - 25	1 - 25	1 - 25	130 – 400
L (2 x bp)	300	300	300	150

Fonte: Elaboração baseada nos dados disponíveis em: http://www.illumina.com.

O pirossequenciamento apresenta sequências de maior comprimento de leitura (Liu *et al.*, 2012), confirmado em testes de rendimento no sequenciamento de *Escherichia coli*, no qual também foram obtidos os maiores conjuntos de montagem, no entanto o rendimento foi baixo. Por outro lado, os equipamentos com sequenciamento através de terminadores reversíveis por Bridge PCR apresentam maior saída de dados e menor custo de reagentes, e nos testes com *E.*

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

coli apresentaram o maior rendimento por corrida e a menor taxa de erro. O sequenciamento por ligação de octâmeros apresenta maior precisão que os dois anteriores (Tabela 4). Já o sequenciamento por detecção de íons apresenta maior rendimento rodando em um modelo com 100bp (Loman *et al*., 2012).

Tabela 4. Comparativo para as versões G.007 e H da plataforma de sequenciamento Polonator.

Características	G.007	H
Rendimento (Gb/h)	0,4	8,7 (1 câmera)
		34,9 (4 câmeras)
Tempo/Flow cell (h)	45	14 (1 câmera)
		3,5 (4 câmeras)
Tempo/Genoma Humano (h)	710	34,2 (1 câmera)
		8,5 (4 câmeras)

Fonte: Elaboração baseada em Loman *et al.* (2012) e Liu *et al*. (2012)

Outro problema, além do tamanho dos fragmentos obtidos por sequenciadores de segunda geração (Tabela 5), são os erros de inserção, deleção e substituição de bases, que podem inviabilizar a análise de dados obtidos. Apesar de apresentar abordagem de ligação e emPCR, a plataforma polonator diferencia-se das demais pela leitura das bases ser realizada através de imagens de câmeras, mas a tecnologia ainda está em desenvolvimento, portanto os dados ainda são escassos (Huse *et al*., 2007).

A quebra das moléculas de DNA em fragmentos nos sequenciadores gera um grande volume de dados que serão utilizados posteriormente para a montagem da molécula. Apesar da rapidez dos sequenciadores da segunda geração, os fragmentos obtidos por eles são menores que os obtidos por *Sanger*, o que pode acarretar uma série de erros de sequenciamento. Os sequenciadores da terceira geração apresentam uma nova abordagem, em que as plataformas utilizam a técnica de sequenciamento de molécula única. Também surgem, nessa geração, os sistemas em miniatura, utilizando a nanotecnologia. Eles dispensam amplificações iniciais (necessárias a todos os outros sistemas), e conservam as características iniciais desejadas dos sequenciadores das outras gerações, tais como rapidez e economia. No entanto, a representatividade dos dados e a redução dos erros experimentais continuam sendo os principais obstáculos a serem superados pelas plataformas dessa geração (Athanasopoulou *et al.*, 2021).

Tabela 5. Comparativo das principais plataformas de sequenciamento de próxima geração. As plataformas (P) são apresentadas, em seguida o tipo de sequenciamento (S), o tipo de amplificação (A), a geração (G) do sequenciador, o tamanho das leituras (L) em pares de base, o rendimento (R) em gigapares de bases, a precisão (P) em porcentagem, número médio de reads (Rd), a enzima (E) envolvida no processo (pol – polimerase; lig – ligase), o tipo de erro (Er) comum à plataforma (Ins – erro de inserção; Sub- erro de substituição; Del – erro de deleção de bases). Dados de leitura, rendimento e o número de reads são apresentados por rodada de sequenciamento.

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

P	S	A	G	L (pb)	R (Gpb)	P (%)	Rd	E	Er
Sanger	Terminação de cadeias	PCR	1ª	600 – 1000	-	100	96	Pol	-
454(Roche)	Pirossequencia mento	emPCR	2ª	400-700	0,7	99,9	$1x10^6$	Pol	Ins
Illumina	Terminadores reversíveis	Bridge PCR	2ª	100-300	600	99,9	$1\text{-}8x10^9$	Pol	Sub
SOLiD	Ligação de octâmeros	emPCR	2ª	75-85	80-360	99,9	$8x10^8$	Lig	Sub
Ion torrent	Detecção de H+	emPCR	2ª	100-400	0,1 – 64	99	$8,2x10^7$	Pol	-
Polonator	Nanomeros	emPCR	2ª	13	-	-	-	Lig	Sub
PacBio	Nt fluorescentes marcados	SMRT	3ª	4000-5000	0,2 – 1	95	$3,5\text{-}7,5x10^4$	-	-
Helicos	Terminadores reversíveis	tSMS	3ª	25-55	35	97	-	Pol	Del

Fonte: Elaboração baseada em informações de Rhoads e Au (2015); Barba, Czosnek e Hadidi. (2014); Shendure e Ji, (2008); Treffer e Decker (2010).

O maior obstáculo para a implementação dos sequenciadores de próxima geração continua sendo a montagem dos genomas utilizando o volume de dados obtidos. A montagem de genomas pode ser feita apenas com os *reads* obtidos dos sequenciadores (montagem de novo) ou por comparação com um genoma de referência. Entretanto, encontrar as ferramentas adequadas para a montagem de genomas com a grande quantidade de fragmentos de sequências de DNA gerados pelos NGS não é uma tarefa muito fácil de ser realizada. As duas principais tecnologias de montagem são o grafo de Bruijn e a estratégia de *Overlap-Layout-Consensus* (OLC) (Athanasopoulou *et al*., 2021).

Na Estratégia de *overlap-layout-consensus* (OLC) ou grafos de sobreposição, a sobreposição dos fragmentos determina a conexão entre eles. Os *reads* são comparados, e conectados por arestas. Isso pode ser feito por várias estratégias de cálculo de sobreposição. Pode ser aplicado a montagens que utilizam fragmentos maiores como os obtidos nos sequenciadores baseados no método de *Sanger* (Schatz; Delcher; Salzberg, 2010; Miller; Koren; Sutton, 2010; Conway; Bromage, 2011; Couto *et al.*, 2012). A OLC utiliza como ferramentas os montadores MIRA (Chevreux; Wetter; Suhai, 1999; Chevreux, 2007), Celera (Myers *et al.* 2000) e *Minimus* (Sommer *et al.*, 2007).

O Grafo de Bruijn ou grafo K-mer utiliza a estratégia de sobreposição de fragmentos, no entanto, os reads não são comparados par a par. É permitido apenas sobreposições exatas e faz-se necessária a realização de correção dos erros antes e durante a montagem para a obtenção das sequências consenso com qualidade. O grafo pode ser aplicado à montagem de genomas utilizando fragmentos curtos de alta precisão, como os obtidos nos sequenciadores NGS (Conway; Bromage, 2011; Nagarajan; Pop, 2013); também utiliza os montadores Euler

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

(Pevzner; Tang; Waterman, 2001), *Velvet* (Zerbino; Birney, 2008), *Allpaths* (Butler *et al*., 2008), *AbySS* (Simpson *et al*., 2009), SOAPdenovo (Li *et al*., 2010) e o Ray (Boisvert; Laviolette; Corbeil, *et al*., 2010).

A aplicação das tecnologias dos sistemas NGS melhora as ferramentas de análises das comunidades microbianas em determinados ambientes de ecossistemas aquáticos e terrestres (Shokralla *et al*., 2012). Ela tem proporcionado mudanças positivas nas áreas do sequenciamento de genomas, ecologia, descoberta de novas espécies, epidemiologia, transcriptomas, replicação, detecção e identificações microbianas (Barba; Czosnek; Hadidi, 2014). Estudos de ecologia microbiana, utilizando métodos moleculares de sequenciamento em massa, a exemplo do pirossequenciamento, tem mostrado maior adequação para a obtenção de sequências de maior comprimento, e em larga escala, entretanto, os estudos utilizando os terminadores reversos com Bridge PCR e outras tecnologias de NSG também têm apresentado crescimento (Nikolaki; Tsiamis, 2013).

As técnicas também permitem comparar a diversidade microbiana de ambientes muito diferentes, com poucos reads, ou ambientes muito parecidos, que podem precisar de milhares deles (Lemos *et al.*, 2011). Roesch *et al*. (2007) utilizaram a técnica para acessar a diversidade microbiana de solos, Bokulich *et al*. (2012) acessaram a diversidade bacteriana do vinho botritizado, Di Maiuta *et al*. (2013) estudaram a dinâmica populacional em fezes de peixes bagres para obtenção de enzimas celulolíticas e Yergeau *et al*. (2012) sequenciaram fragmentos das comunidades microbianas em áreas contaminadas por petróleo. Essa abordagem também foi utilizada para estudar a diversidade microbiana de amostras de bioaerosol causadores da síndrome da poeira orgânica tóxica, obtendo-se sequências de alta qualidade tanto para fungos como para bactérias (Madsen *et al.*, 2015).

Geralmente, as plataformas NGS podem ser utilizadas no sequenciamento de amplicons, na captura de sequências, e no sequenciamento do genoma completo (vírus, fungos e bactérias), em metagenômica, e no sequenciamento de transcriptomas. Também podem ser aplicadas em genotipagem do polimorfismo de nucleotídeos simples, tipagem microbiana, ressequenciamento, e sequenciamento de marcadores (tag sequencing) (Ronaghi, 2001).

O pirossequenciamento foi utilizado para estudar a diversidade microbiana e a dinâmica populacional associada à dieta de peixes bagres (Di Maiuta, 2013), e para estudar a comunidade microbiana em um biorreator anaeróbico dinâmico de membrana para tratamento municipal de águas subterrâneas (Ma *et al.*, 2013). O sequenciamento por terminadores reversos com Bridge PCR foi utilizado para acessar 24 comunidades microbianas, com indivíduos associados a

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

hospedeiros e de vida livre (Caporaso *et al*., 2012), e pode ser utilizado para sequenciar amplicons de bactérias (Loman *et al*., 2012).

As abordagens de pirossequenciamento com emPCR e de síntese por terminadores reversíveis com Bridge PCR foram utilizadas para estudar a dinâmica da comunidade microbiana em cultivos de uma espécie de camarão branco, *Litopenaeus vannamei*. Os dados obtidos acessando os genes 16S rRNA com os dois tipos de sequenciamento foram suficientes para detecção de 31 Filos, 66 classes, 90 ordens, 213 famílias e 697 gêneros de bactérias (Zhang *et al.*, 2016). Já Zimmer *et al*. (2014) utilizaram Bridge PCR para estudar a adaptação de isolados de *Saccharomyces cerevisiae* ao sulfito na fermentação do vinho.

O uso da abordagem de sequenciamento por ligação possibilitou a identificação de 55 espécies bacterianas. Foi possível observar a abundância de filos, classe e gêneros. Além disso, foram obtidas sequências de enzimas selecionadas para o estudo de análise evolutiva no DNA, dentre eles enzimas que provocam a hidrólise do amido, tais como a alfa-amilase, e a glucoamilase; e as enzimas relacionadas à desconstrução da celulose, tais como a alfa-glucosidase, beta-glucosidase e a endoglucanase (Lima, 2014).

As plataformas que utilizam o sequenciamento por detecção de íons se aplicam bem a trabalhos de pequena escala e rápida resposta; mostraram-se adequadas ao sequenciamento de vírus e bactérias, gastando apenas uma hora (Miller; Koren; Sutton, 2010); apresentam um amplo espectro de uso na caracterização de populações de bactérias, análises taxonômicas e identificação de espécies (Life Technologies Corporation, 2013), para o sequenciamento em estudos de microbiologia ambiental - metagenômica e para o sequenciamento de microrganismos em microbiologia médica (Yergeau *et al*., 2012).

A abordagem também foi utilizada para sequenciar amplicons da região ITS para verificar o efeito das queimadas controladas em comunidades de fungos de florestas. Seus resultados mostraram que há alterações na comunidade de fungos em locais cujo intervalo de queimadas é de três anos, entretanto, naquelas em que o intervalo foi de seis anos, a comunidade se mostrou similar às florestas onde não se aplica a técnica de queimada controlada (Brown *et al*., 2013). Os resultados de Pylro *et al*. (2014), com a abordagem de terminadores reversos com Bridge PCR e detecção de íons, foram satisfatórios na avaliação de perfis filogenéticos de genes 16S rRNA de comunidades microbianas naturais utilizando estudos de alfa e beta diversidade.

O NGS também pode ser aplicado em estudos de interação planta patógeno. A abordagem RNA sequencing, presente nos sequenciadores baseados em PCR e molécula única, foi eficiente em identificar os genes diferencialmente expressos durante a interação mamoeiro-*Papaya Meleira Virus* (PMeV) e indicou alguns genes chaves na indução à resistência ao

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

patógeno (Madroñero, 2014). Em outro trabalho, isolados de *Johnsongrass mosaic virus* (*JGMV*), encontrados infectando plantas do gênero *Brachiaria* e *Panicum* foram utilizados para a obtenção de primers específicos para detecção do vírus nas plantas estudadas (Silva, 2015). A diversidade de *Oomycetes*, como *Phytophthora infestans* e *Plasmopara viticola* também já foram estudos por essa abordagem. Anjos (2013) estudou a diversidade de *Plasmopara viticola*, e a estrutura genética de populações para obter marcadores microssatélite para estudos desse patógeno em vinhedos.

A técnica de True Single Molecule Sequencing (TSMS) foi utilizada para gerar dados que permitissem sequenciar o DNA de ossos de cavalos encontrados no permafrost (Orlando, 2016). O sequenciamento de molécula única Nanopore ainda está em teste, e mostrou-se capaz de cobrir todo o operon do RNA em uma única leitura, mas apresentou erro de leitura, o que indica a necessidade de melhorias (Roberts; Carneiro; Schartz, 2013).

Na indústria, a TSMS pode ser utilizada para conhecer melhor a atuação individual e coletiva dos microrganismos. Isso pode ser realizado devido a geração de dados sobre esses conjuntos de microrganismos industrialmente importantes, pela reconstrução de genes, e descoberta de seus agrupamentos; por meio de estudos sobre a ação da regulação epigenética na eficiência da manipulação genética dos micróbios, mediante as identificações e isolamento de membros críticos de comunidades microbianas complexas e do desenvolvimento de produtos microbianos comerciais; realizando estudos das populações virais para facilitar o desenvolvimento e avaliações de drogas e vacinas; obtendo sequências de transcriptomas e metatranscriptomas para definir diretamente os operons bacterianos completos; gerar produtos microbianos industrialmente importantes; e validar marcadores (targets) (Roberts; Carneiro; Schartz, 2013).

As vantagens são a obtenção de sequências de alta qualidade, com comprimento de leitura que reduzem os gastos com as análises, que são rápidas e duram cerca de 2 horas, não há necessidade de clonar o DNA. Entretanto, os equipamentos apresentam como desvantagens o alto custo de aquisição e manutenção, além de requerer massa alta de DNA (1 micrograma) (Zhang *et al*., 2011; Roberts; Carneiro; Schartz, 2013).

De maneira geral, os métodos de sequenciamento utilizados pelas plataformas NGS podem ser aplicados a estudos de polimorfismo de nucleotídeo único (SNP), identificação de bactérias, tipagem fúngica e viral, detectar mutações, análises de metilação do DNA, sequências múltiplas, verificações de clones, sequenciamento do genoma humano, sequenciamento completo do genoma de vírus e bactérias e mudanças epigenéticas. Em metagenômica para estudos de diversidade ambiental de microrganismos, essas plataformas de sequenciamento do

Revista Macambira, Serrinha (BA), v. 8, n. 1, e081017, Jan.-Dez., 2024
ISSN: 2594-4754 | **DOI:** https://doi.org/10.35642/rm.v8i1.1248

DNA, dentre muitas finalidades, podem ser usadas para caracterizar e identificar, através de análises taxonômicas, espécies da comunidade microbiana de determinado ambiente (Life Technologies Corporation, 2013).

CONSIDERAÇÕES FINAIS

As plataformas NGS apresentam infraestrutura computacional cara e requerem um centro de dados com instalações e pessoal qualificado, além de espaço de mídia suficiente para armazenagem de dados, redes rápidas de acesso e sistemas de bancada associados. Entretanto, há um dinamismo muito grande nas alterações envolvendo o campo do sequenciamento de genes por plataformas NGS. Além do surgimento de novas tecnologias, as remanescentes melhoram seu desempenho de forma geral, preocupando-se com a qualidade dos dados obtidos e com as explicações biológicas que possam validar esses achados.

Apesar de haver uma escassez de informações detalhadas e de ferramentas adequadas ao alto potencial dos sequenciadores dessa natureza, os avanços da NGS irão redefinir o campo da genômica. Principalmente para a identificação dos microrganismos de sistemas agrícolas e ambientais, pois algumas técnicas possibilitam a identificação de microrganismos não cultiváveis. Além disso, podem possibilitar estudo de toda a comunidade microbiana presente em amostras retiradas de diferentes biomas ou até envolvidos em determinados bioprocessos.

Os estudos do tema levaram a conclusão de que, desde a descoberta do DNA, cientistas desenvolvem tecnologias importantes para os estudos de biologia molecular. As novas ferramentas poderão ser utilizadas em diversas áreas e seráo de suma importância para verificação do papel e da composição das comunidades microbianas presentes no meio ambiente, nos estudos da composição de solos produtivos, da aplicação de microrganismos na agricultura, no controle de pragas e doenças, entre outras.

A aplicação tanto na agricultura como nas análises ambientais dessas ferramentas de nova geração poderão ser uma vantagem estratégica no combate ao uso indesejado dos recursos naturais, possibilitando aplicação de abordagens menos onerosas ao ambiente e adoção de práticas mais conservacionistas. O acesso a essas tecnologias poderá ajudar na aplicação mais eficiente dos conceitos de sustentabilidade, promovendo a melhoria das condições sociais, maior economia de recursos e respeito às características ambientais.

REFERÊNCIAS BIBLIOGRÁFICAS

ALDERBORN, A.; KRISTOFFERSON, A.; HAMMERLING, U. Determination of Single-Nucleotide Polymorphisms by Real-time Pyrophosphate DNA Sequencing. **Genome research**, v.10, p.1249–1258, 2000. DOI: https://doi.org/10.1101/gr.10.8.1249.

ALLEN, R. A.; WILLIAMS, C. L.; PENROD, Y.; MCCLOSKEY, C.; CARPENTER-AZEVEDO, K.; HUARD, R. C.; KING, E.; DUNN, S. T. A pyrosequencing protocol for rapid identification of SARS-CoV-2 variants. **Journal of Medical Virology**, v. 94, n. 8, p. 3661-3668, 2022. DOI: https://doi.org/10.1002/jmv.27770.

ANDREOTO, F.D. Análise genômica e transcriptômica de *Methylobacterium mesophilicum* SR1.6/6 em interação com a planta hospedeira. 2011. 80f. **Dissertação** (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba-SP, 2011.

ANJOS, L. M. Diversidade genética de *Plasmopara viticola* e mapeamento de QTLs de resistência ao míldio em videira (*Vitis spp*.). Brasília, Novembro 2013.277p. : il. **Tese** (Doutorado). Programa de Pós-graduação em Fitopatologia, Universidade de Brasília, Brasília. 227p. 2013.

ANSORGE, Wilhelm J. Next-generation DNA sequencing techniques. **New biotechnology,** v. 25, n. 4, p. 195-203, 2009. DOI: https://doi.org/10.1016/j.nbt.2008.12.009

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, 2021. DOI: https://doi.org/10.3390/life12010030.

BARBA, M.; CZOSNEK, H.; HADIDI, A. Historical perspective, development and applications of next-generation sequencing in plant virology. **Virus**, v.6, p.106-136, 2014. DOI: https://doi.org/10.3390/v6010106

BOISVERT, S.; LAVIOLETTE, F.; CORBEIL, J. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. **Journal of computational biology**, v.17, p.1519–1533, 2010. DOI: https://doi.org/10.1089/cmb.2009.0238

BOKULICH, N. A. *et al*. Next-Generation Sequencing Reveals Significant Bacterial Diversity of Botrytized Wine. **Plos one**, v.8, n. 10, p.1-10, 2012. DOI: https://doi.org/10.1371/journal.pone.0036357

BROWN, S. P.; CALLAHAM JR, M. C.; OLIVER, A. K.; JUMPPONEN, A. Deep Ion Torrent sequencing identifies soil fungal community shifts after frequent prescribed fires in a southeastern US forest ecosystem. **FEMS Microbiology Ecology**, v. 86, p.557–566, 2013. DOI: https://doi.org/10.1111/1574-6941.12181.

BUTLER, J.; MACCALLUM, I.; KLEBER, M.; SHLYAKHTER, I. A.; BELMONTE, M. K.; LANDER,E. S.; NUSBAUM, C. ; JAFFE, D. B. Allpaths: de novo assembly of wholegenome shotgun microreads. **Genome Research,** v.18, p. 810–820, 2008. DOI: https://doi.org/10.1101/gr.7337908

CAPORASO, J. G. *et al*. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. **The ISME journal**, v.6, n. 8, p.1621–1624, 2012. DOI: https://doi.org/10.1038/ismej.2012.8.

CHEVREUX, B. MIRA: an automated genome and EST assembler. 2007. Disponível em: https://archiv.ub.uni-heidelberg.de/volltextserver/7871/1/thesis_zusammenfassung.pdf. Acesso em: 01 jul. 2016.

CHEVREUX, B.; WETTER, T.; SUHAI, S. Genome sequence assembly using trace signals and additional sequence information. **Journal of Computer Science & Systems Biology**, v.99, p. 45–56, 1999. Disponível em: https://xueshu.baidu.com/usercenter/paper/show?paperid=871ec78542a2b2c43b1a3ce337066be1. Acesso em: 01 jul. 2016.

COLLINS, M. D.; HUTSON, R. A.; FALSEN. E.; SJÖDÉN, B. *Facklamia tabacinasalis* sp. nov., from powdered tobacco. **International journal of systematic bacteriology,** v.49, p.1247-1250, 1999. DOI: https://doi.org/10.1099/00207713-49-3-1247.

CONWAY, T. C.; BROMAGE, A. J. Succinct data structures for assembling large genomes. **Bioinformatics,** v.27, p.479–486, 2011. DOI: https://doi.org/10.1093/bioinformatics/btq697

COUTO, A. D.; CERQUEIRA, F. R.; GUERRA, R. L.; GONÇALVES, L. B.; GOULART, C. C.; SIQUEIRA-BATISTA, R.; FERREIRA, R. S.; OLIVEIRA, A. P. Theoretical basis of a new method for dna fragment assembly in k-mer graphs. **31st International Conference of the Chilean Computer Science Society,** 2012. DOI: https://doi.org/10.1109/SCCC.2012.16

DI MAIUTA, N. *et al.* Microbial population dynamics in the faeces of wood-eating loricariid catfishes. **Letters in applied microbiology**, v.56, n. 6, p.401-407, 2013. DOI: https://doi.org/10.1111/lam.12061.

DURBIN, R. M. *et al.* A map of human genome variation from population-scale sequencing. **Nature**, v. 467, p.1061-1073, 2010. DOI: https://doi.org/10.1038/nature09534.

FELSKE, A.; AKKERMANS, A. D. L.; VOS, W. M. In situ detection of an uncultured predominant bacillus in Dutch grassland soils. **Applied and environmental microbiology,** v.64, p.4588–4590, 1998. DOI: https://doi.org/10.1128/AEM.64.11.4588-4590.1998

HALL, N. Advantages sequencing technologies and their wider impact in microbiology. **The journal of experimental biology**, v. 209, p.1518-1525, 2007. DOI: https://doi.org/10.1242/jeb.001370

HU, T.; CHITNIS, N.; MONOS, D.; DINH, A. Next-generation sequencing technologies: An overview. **Human Immunology**, v. 82, n. 11, p. 801-811, 2021. DOI: https://doi.org/10.1016/j.humimm.2021.02.012.

HUSE, S. M.; HUBER, J. A.; MORRISON, H. G.; SOGIN, M. L.; WELCH, D. M. Accuracy and quality of massively parallel DNA pyrossequencing. **Genome biology**, v.8, p.1-9, 2007. DOI: https://doi.org/10.1186/gb-2007-8-7-r143

KOCHE, J. C. **Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa.** 34. ed. Petrópolis, RJ: Vozes, 2015.

LANDER, E.S. *et al.* Initial sequencing and analysing of the human genome. **Nature,** v. 409, p.860-921, 2001. DOI: https://doi.org/10.1038/35057062.

LEMOS, L. N.; FULTHORPE, R. R.; TRIPLETT, E. W.; ROESCH, L. F.. Rethinking microbial diversity analysis in the high throughput sequencing era. **Journal of microbiological methods**, v.86, p.42–51, 2011. DOI: https://doi.org/10.1016/j.mimet.2011.03.014

LI, R.; ZHU, H.; RUAN, J.; QIAN, W.; FANG, X.; SHI, Z.; LI, Y.; LI, S. et al. De novo assembly of human genomes with massively parallel short read sequencing. **Genome research**, v.20, p.265–272, 2010. DOI: https://doi.org/10.1101/gr.097261.109

LIFE TECHNOLOGIES CORPORATION. Longer read lengths improve bacterial identification using 16s rRNA Gene Sequencing on Theion PGM™ System. **Your Innovative Research**, p.1-6, 2013. Disponível em: https://assets.thermofisher.com/TFS-Assets/LSG/brochures/16S-rRNA-Gene-Sequencing-App-Note.pdf. Acesso em: 20 mar. 2016.

LIMA, Milena Tavares. Análise funcional de um consórcio microbiano de solo e prospecção de genes envolvidos na desconstrução da biomassa. 2014. iii, 45 p. **Dissertação** (mestrado) -

Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, 2014. Disponível em: http://hdl.handle.net/11449/121846. Acesso em: 20 mar. 2016.

LIU, L.; LI, Y.; LI, S.; HU, N.; HE, Y.; PONG, R.; LIN, D.; LU, L.; LAW, M. Comparison of Next-Generation Sequencing Systems. **Journal of biomedicine and biotechnology**, p.1-12, 2012. DOI: https://doi.org/10.1155/2012/251364.

LOMAN, N. J.; MISRA, R. V.; DALLMAN, T. J.; CONSTANTINIDOU, C.; GHARBIA, S. E.; WAIN, J.; PALLEN, M. J. Performance comparison of benchtop high-throughput sequencing platforms. **Nature biotechnology,** v.30, n.5, p.434-562, 2012. DOI: https://doi.org/10.1038/nbt.2198.

MA, J.; WANG, Z.; ZOU, X.; FENG, J.; WU, Z.Microbial communities in an anaerobic dynamic membrane bioreactor (AnDMBR) for municipal wastewater treatment: Comparison of bulk sludge and cake layer. **Process biochemistry**, v.48, n. 3, p.510–516, 2013. DOI: https://doi.org/10.1016/j.procbio.2013.02.003

MADROÑERO, L. J. Análise trancriptômica da interação mamoeiro-Papaya Meleira Virus. 2014. 76 f. : il. **Dissertação** (Mestrado em Biotecnologia) – Universidade Federal do Espírito Santo, Centro de Ciências da Saúde.

MADSEN, A.M.; ZERVAS, A.; TENDAL, K.; NIELSEN, J. L. Microbial diversity in bioaerosol samples causing ODTS compared to reference bioaerosol samples as measured using Illuminas equencing and MALDI TOF. **Environmental research,** v.140, p.255–267, 2015. DOI: https://doi.org/10.1016/j.envres.2015.03.027

MARGULIES, M. *et al.* Genome Sequencing in Open Microfabricated High Density Picoliter Reactors. **Nature,** v. 437, p. 376–380. 2005. DOI: https://doi.org/10.1038/nature03959.

MILLER, J. R.; KOREN, S.; SUTTON, G. Assembly algorithms for next-generation sequencing data. **Genomics,** v.95, p. 315-327, 2010. DOI: https://doi.org/10.1016/j.ygeno.2010.03.001.

MYERS, E. W. *et al.* A whole-genome assembly of drosophila. **Science,** v.287, p.2196–2204, 2000. DOI: https://doi.org/10.1126/science.287.5461.2196

NAGARAJAN, N.; POP, M. Sequence assembly demystified. **Nature Review Genetics**, v.14, p.157–167, 2013. DOI: https://doi.org/10.1038/nrg3367

NIKOLAKI, S.; TSIAMIS, G. Microbial diversity in the era of omic technologies. **BioMed Research International**, v.2013, p.1-15, 2013. DOI: http://dx.doi.org/10.1155/2013/958719

OLSON, M. V. The human genome project. **Proceedings of the National Academy of Sciences**, v. 90, n. 10, p. 4338-4344, 1993. DOI: https://doi.org/10.1073/pnas.90.10.4338.

ORLANDO, L. *et al*. True single-molecule DNA sequencing of a pleistocene horse bone. **Genome Research**, v. 21, p. 1705–1719. DOI: https://doi.org/10.1101/gr.122747.111

OXFORD NANOPORE TECHNOLOGIES. Disponível em: www.nanoporetech.com. Acesso em: 01 jul. 2016.

PETTERSSON, E.; LUNDEBERG, J.; AHMADIAN, A. Generations of sequencing technologies. **Genomics,** v. 93, n. 2, p. 105-111, 2009. DOI: https://doi.org/10.1016/j.ygeno.2008.10.003.

PYLRO, V. S.; ROESCH, L. F. W.; MORAIS, D. K.; CLARK, I. M.; HIRSCH, P. R.; TÓTOLA, M. R. Data analysis for 16S microbial profiling from different benchtop

sequencing platforms**. Journal of Microbiological Methods**, v.107, p.30–37, 2014. DOI: https://doi.org/10.1016/j.mimet.2014.08.018

PEVZNER, P.A.; TANG, H., WATERMAN, M.S. An Eulerian path approach to DNA fragment assembly. **Proceedings of the National Academy of Sciences**, vol. 98, no. 17, 9748-9753, 2001. DOI: https://doi.org/10.1073/pnas.171285098.

QUICK, J.; QUINLAN, A.R.; LOMAN, N.J. A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. **GigaScience**, v.3, n.22, p.1-6, 2014. DOI: https://doi.org/10.1186/2047-217X-3-22

RAMACHANDRAN, A.; LIU, Y.; ASQHAR, W.; IQBAL, S. M. Characterization of DNA-nanopore INteretions by Molecular Dynamics. **American journal of biomedical sciences,** v.1, p.344-351, 2009. DOI: https://doi.org/10.5099/aj090400344.

RAUEN, F. J. **Roteiros de investigação científica.** 2. ed. Tubarão: Unisul, 2018.

RAVIN, N. V. Modern Methods of Genome Sequencing and Their Application for Deciphering Genomes of Microorganisms. **Applied Biochemistry and Microbiology**, V.46, p.663–670, 2010. DOI: http://dx.doi.org/10.1134/S000368381007001X.

RHOADS, A.; AU, K. F. PacBio sequencing and its applications. **Genomics proteomics bioinformatics**, v.13, p.278–289, 2015. DOI: http://dx.doi.org/10.1016/j.gpb.2015.08.002.

ROBERTS, R.J.; CARNEIRO, M.O.; SCHARTZ, M.C. The advantages of SMRT sequencing. **Genome biology**, v.14, n.405, p.1-4, 2013. DOI: http://doi.org/10.1186/gb-2013-14-7-405.

ROESCH, L.F.W. et al. Pyrosequencing enumerates and contrasts soil microbial diversity. **The ISME journal**, v.1, n. 4, p.283–290, 2007. DOI: http://doi.org/10.1038/ismej.2007.53

RONAGHI, M. Pyrosequencing Sheds Light on DNA Sequencing. **Genome research,** v.11, p.3-11, 2001. DOI: http://doi.org/10.1101/gr.11.1.3.

RONAGHI, M.; KARAMOHAMED, S.; PETTERSON, B.; UHLÉN, M.; NYRÉN, P. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. **Analytical Biochemistry,** v.242, p.84-89, 1996. DOI: http://dx.doi.org/10.1006/abio.1996.0432.

SANGER, F.; COULSON, A. R.; BARREL, B. G.; SMITH, A. J. H.; ROE, B. A. Cloning in Single-stranded Bacteriophage as an Aid to Rapid DNA Sequencing. **J. Mol. Biol**. v.143, p.161-178, 1980. DOI: https://doi.org/10.1016/0022-2836(80)90196-5.

SANGER, F.; COULSON, A. R.; FRIEDMANN, T.; AIR, G. M.; BARRELL, B. G.; BROWN, N. L.; FIDDES, J. C.; HUTCHISON, C. A.; SLOCOMBE, P. M.; SMITH, M. The Nucleotide Sequence of Bacteriophage ØX174. **Journal of molecular biology**, v.125, p.225-246, 1978. DOI: https://doi.org/10.1016/0022-2836(78)90346-7.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Procedings of the national academy of sciences USA,** v.74, p.5463-5467, 1977. DOI: https://doi.org/10.1073/pnas.74.12.5463.

SATAM, H.; JOSHI, K.; MANGROLIA, U.; WAGHOO, S.; ZAIDI, G.; RAWOOL, S.; TRAKARE, R. P.; BANDAY, S.; MISHRA, A. K.; DAS, G; MALONIA, S. K. Next-generation sequencing technology: current trends and advancements. **Biology**, v. 12, n. 7, p. 997, 2023. DOI: https://doi.org/10.3390/biology12070997.

SCHADT, E. E.; TURNER, S.; KASARSKIS, A. A window into third-generation sequencing. **Human Molecular Genetics**, v.19, p.227–240, 2010. DOI: https://doi.org/10.1093/hmg/ddq416.

SCHATZ, M. C.; DELCHER, A. L.; SALZBERG, S. L. Assembly of large genomes using second-generation sequencing. **Genome research**, v.20, pp.1165–1173, 2010. DOI: https://doi.org/10.1101/gr.101360.109

SEQUENCING PLATFORMS. Disponível em: http://www.illumina.com. Acesso em: 20 dez. 2014.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature Biotecnologie**, v.26, p.1135-1145, 2008. DOI: https://doi.org/10.1038/nbt1486

SHOKRALLA, S.; SPALL,J. L.; GIBSON, J. F.; HAJIBABAEI, M. Next-generation sequencing technologies for environmental DNA research. **Molecular Ecology**, v.21, p.1794–1805, 2012. DOI: https://doi.org/10.1111/j.1365-294X.2012.05538.x.

SILVA, K. N. Caracterização molecular de Johnsongrass mosaic virus em plantas forrageiras dos gêneros Brachiaria, Panicum e Pennisetum. Brasília, 2015.p.111. **Dissertação** (Mestrado). Programa de Pós-graduação em Fitopatologia, Universidade de Brasília, Brasília.

SIMPSON, J. T.; WONG, K.; JACKMAN, S. D.; SCHEIN, J. E.; JONES, S. J.; BIROL,N. Abyss: A parallel assembler for short read sequence data. **Genome research,** v.19, p. 1117–1123, 2009. DOI: https://doi.org/10.1101/gr.089532.108

SOMMER, D.; DELCHER, A.; SALZBERG, S. ; POP, M. Minimus: a fast, lightweight genome assembler. **BMC bioinformatics**, v.8, p.1-11, 2007. DOI: https://doi.org/10.1186/1471-2105-8-64

STADEN, R. A Strategy of DNA sequencing employing computer programs. **Nucelic Acids Research,** v.6, n.7, p.1-10, 1979. DOI: https://doi.org/10.1093/nar/6.7.2601

SUMINDA, G. G. D.; GHOSH, M.; SON, Y. The innovative informatics approaches of high-throughput technologies in livestock: spearheading the sustainability and resiliency of agrigenomics research. **Life**, v. 12, n. 11, p. 1893, 2022. DOI: https://doi.org/10.3390/life12111893.

TREFFER, R.; DECKER, V. Recent advances in single-molecule sequencing. **Current Opinion in Biotechnology**, v.21, n. 1, p.4–11, 2010. DOI: https://doi.org/10.1016/j.copbio.2010.02.009

WANG, Y.; ZHAO, Y.; BOLLAS, A.; WANG, Y.; AU, K. F. Nanopore sequencing technology, bioinformatics and applications. **Nature biotechnology,** v. 39, n. 11, p. 1348-1365, 2021. DOI: https://doi.org/10.1038/s41587-021-01108-x.

YERGEAU, E.; LAWRENCE, J. R.; SANSCHAGRIN, S.; WAISER, M. J.; KORBER, D. R.; GREER, C. W. Next-Generation Sequencing of Microbial Communities in the Athabasca River and Its Tributaries in Relation to Oil Sands Mining Activities. **Applied and environmental microbiology**, v.78, n.21, p.7626–7637, 2012. DOI: https://doi.org/10.1128/AEM.02036-12

ZERBINO, D. R.; BIRNEY, E. Velvet: Algorithms for de novo short read assembly using de bruijn graphs. **Genome Research**, v.18, p. 821–829, 2008. DOI: https://doi.org/10.1101/gr.074492.107

ZHANG, J.; CHIODINI, R.; BADR, A.; ZHANG, G. The impact of next-generation sequencing on genomics. **Journal of genetics and genômics,** v. 38, n. 3, p.96-109, 2011. DOI: https://doi.org/10.1016/j.jgg.2011.02.003.

ZHANG, L. *et al*. Advances in metagenomics and its application in environmental microorganisms. **Frontiers in microbiology**, v. 12, p. 766364, 2021. DOI: https://doi.org/10.3389/fmicb.2021.766364.

ZHANG,H.; SUN, Z.; LIU, B.; XUAN, Y.; JIANG, M.; PAN, Y.; ZHANG, Y.; GONG, Y.; LU, X.; YU, D.; KUMAR, D.; HU, X.; CAO, G.; XUE, R.; GONG, C. Dynamic changes of microbial communities in *Litopenaeus vannamei* cultures and the effects of environmental factors. **Aquaculture,** v.455, p.97-108, 2016. DOI: https://doi.org/10.1016/j.aquaculture.2016.01.011.

ZIMMER, A.; DURAND, C.; LOIRA, N.; DURRENS, P.; SHERMAN, D. J.; MARULLO, P. QTL. Dissection of Lag Phase in Wine Fermentation Reveals a New Translocation Responsible for *Saccharomyces cerevisiae* adaptation to sulfite. **Plos One**, v.9, 2014. DOI: https://doi.org/10.1371/journal.pone.0086298 .

Revista Macambira

Informações do Artigo	Article Information
Recebido em: 13/02/2024 **Aceito em**: 09/07/2024 **Publicado em**: 23/09/2024	**Received on**: 02/13/2024 **Accepted in**: 07/09/2024 **Published on**: 23/09/2024
Conflitos de Interesse O autor declara não haver nenhum conflito de interesse de ordem pessoal, comercial, acadêmico, político e financeiro referente a este manuscrito.	**Interest conflicts** The author declare that there is no personal, commercial, academic, political or financial conflict of interest regarding this manuscript.
Como Citar este artigo - ABNT SANTANA FILHO, Djalma Moreira. Sequenciamento de próxima geração: novas tecnologias para acessar às comunidades microbianas. **Revista Macambir**a, Serrinha (BA), v. 8, n. 1, e081017, jan./dez., 2024. https://doi.org/10.35642/rm.v8i1.1248	**How to cite this article - ABNT** SANTANA FILHO, Djalma Moreira. Next-Generation Sequencing (NGS): new technologies to access microbial communities. **Revista Macambir**a, Serrinha (BA), v. 8, n. 1, e081017, jan./dez., 2024. https://doi.org/10.35642/rm.v8i1.1248.
Licença de Uso A Licença Creative Commons Atribuição-NãoComercial-CompartilhaIgual 4.0 Internacional (CC BY 4.0). Esta licença permite compartilhar, copiar, redistribuir o manuscrito em qualquer meio ou formato. Além disso, permite adaptar, remixar, transformar e construir sobre o material, mesmo que comercialmente, desde que seja atribuído o devido crédito de autoria e publicação inicial neste periódico.	**Use license** The Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY 4.0). This license allows sharing, copying, redistributing the manuscript in any médium or format. In addition, it allows adapting, remixing, transforming and building on the material, even commercially, as long as due credit for authorship and initial publication in this journal is attributed.