

Jaguariúna, SP / Maio, 2025

Explorando o Sequenciamento Genético do Microbioma do Solo na Agricultura

OBJETIVOS DE DESENVOLVIMENTO SUSTENTÁVEL



**Empresa Brasileira de Pesquisa Agropecuária
Embrapa Meio Ambiente
Ministério da Agricultura e Pecuária**

e-ISSN 2965-7342

Documentos 140

Maio, 2025

Explorando o Sequenciamento Genético do Microbioma do Solo na Agricultura

*Rodrigo Mendes
Lucas William Mendes
Fernando Dini Andreote*

**Embrapa Meio Ambiente
Jagariúna, SP
2025**

Embrapa Meio Ambiente
Rodovia SP 34, Km 127,5
Tanquinho Velho
www.embrapa.br/meio-ambiente
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações

Presidente

Janaina Paula Marques Tanure

Secretário-executivo

Anderson Soares Pereira

Membros

*Janaina Paula Marques Tanure,
Robson Rolland Monticelli Barizon,
Alfredo José Barreto Luiz, Fagoni Fayer
Calegario, Marcos Eliseu Losekann, Vera
Lúcia Ferracini, Victor Paulo Marques
Simão, Julio Ferraz de Queiroz, Márcia
Regina Assalin, Maria de Cléofas Faggion
Alencar e Sonia Cláudia do Nascimento
de Queiroz*

Edição executiva
Anderson Soares Pereira

Revisão de texto
Reinaldo Rodrigues

Normalização bibliográfica
Victor Paulo Marques Simão (CRB-8/5139)

Projeto gráfico
Leandro Sousa Fazio

Diagramação
Silvana Cristina Teixeira

Capa
Rodrigo Mendes

Publicação digital: PDF

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte,
constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Meio Ambiente

Mendes, Rodrigo.

Explorando o sequenciamento genético do microbioma do solo na agricultura /
Rodrigues Mendes, Lucas William Mendes, Fernando Dini Andreote. – Jaguariúna :
Embrapa Meio Ambiente, 2025.

PDF (15 p.) : il. color. – (Documentos / Embrapa Meio Ambiente, e-ISSN 2965-
7342 ; 140).

Texto publicado em inglês.

Título em inglês: Exploiting the genetic sequencing of the soil microbiome in
agriculture.

1. Microbioma. 2. Solo. 3. Sequência genética. I. Mendes, Lucas William. II.
Andreote, Fernando Dini. III. Série.

CDD (21. ed.) 574.88

Maria de Cléofas Faggion Alencar (CRB-8/1658)

© 2025 Embrapa

Autores

Rodrigo Mendes

Engenheiro-agrônomo, doutor em Genética e Melhoramento de Plantas, pesquisador da Embrapa Meio Ambiente, Jaguariúna, SP

Lucas William Mendes

Biólogo, doutor em Ciências, professor doutor do Centro de Energia Nuclear na Agricultura da Universidade de São Paulo, Piracicaba, SP

Fernando Dini Andreote

Engenheiro-agrônomo, doutor em Genética e Melhoramento de Plantas, professor titular da Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo, Piracicaba, SP

Apresentação

A saúde dos solos agrícolas resulta da interação entre suas propriedades físicas, químicas e biológicas. Enquanto as dimensões química e física têm sido tradicionalmente manejadas nos sistemas agrícolas, o componente biológico, incluindo os microrganismos do solo, tem recebido atenção crescente nos últimos anos.

O conceito de Saúde Única, instituído em 2008 pela Organização Mundial da Saúde (OMS), pela Organização Mundial da Saúde Animal (OIE) e pela Organização das Nações Unidas para Alimentação e Agricultura (FAO), destaca a interconexão entre a saúde humana, animal, vegetal e ambiental. Nesse contexto, o microbioma se destaca como um elo essencial, conectando e influenciando a saúde dos seres vivos e do ambiente.

Como o solo é um importante reservatório de diversidade microbiana, pode-se afirmar que solos saudáveis são a base para a saúde das plantas, dos animais e das pessoas. Assim, o conhecimento e o manejo do microbioma do solo são essenciais

para promover uma agricultura sustentável e, conseqüentemente, avançar a Saúde Única. Em consonância com a missão da Embrapa Meio Ambiente, que busca integrar a produção agropecuária à sustentabilidade em todas as suas dimensões, o manejo do microbioma do solo contribui diretamente para alcançar o Objetivo de Desenvolvimento Sustentável (ODS) 2: Fome zero - erradicar a fome, alcançar a segurança alimentar, melhorar a nutrição e promover a agricultura sustentável.

Neste texto, os autores oferecem um panorama geral da rápida evolução de técnicas avançadas de biologia molecular, com ênfase no sequenciamento genético, e discutem como essa ferramenta pode ser aplicada em sistemas de produção. Este texto também pode ser visto como um guia que aponta como analisar e explorar ao máximo os dados de sequenciamento genético do microbioma do solo na agricultura. Surgimento de técnicas avançadas de sequenciamento genético revolucionou várias áreas da ciência.

Ana Paula Contador Packer
Chefe-Geral da Embrapa Meio Ambiente

Sumário

Introdução	7
Abordagens e objetivos do sequenciamento de DNA ou RNA do solo	8
A Extração de DNA e RNA do solo	8
Metataxonômica (sequenciamento de amplicons)	10
Metagenômica (sequenciamento de DNA) e Metatranscriptômica (sequenciamento de RNA)	10
Roteiro analítico de dados de sequenciamento genético	11
Oportunidades e desafios do sequenciamento genético do microbioma em sistemas agrícolas	12
Considerações finais	14
Referências	14

Introdução

O surgimento de técnicas avançadas de sequenciamento genético revolucionou várias áreas da ciência no final do século XX e início do século XXI. O grande marco histórico foi a conclusão do sequenciamento do genoma humano em 2003 (Schmutz et al., 2004). O Projeto Genoma Humano (Lander et al., 2001) surgiu no início da década de 1980 a partir de duas percepções fundamentais. Primeiramente, a de que a capacidade de se conhecer o genoma poderia acelerar significativamente a pesquisa biomédica, permitindo que os pesquisadores abordassem problemas de maneira abrangente e imparcial. Em segundo lugar, o desenvolvimento deste projeto exigiria um enorme esforço da comunidade científica na construção de infraestrutura de pesquisa, como nunca observado anteriormente nas ciências biológicas. Curiosamente, o primeiro organismo multicelular a ter o genoma completamente sequenciado não foi o homem, mas foi um pequeno nematoide de vida livre, o *Caenorhabditis elegans*, o qual compõe o microbioma do solo (Genome..., 1998). No Brasil, inspirado nas duas percepções fundamentais que impulsionaram o Projeto Genoma Humano, neste mesmo período, foi colocada em curso uma iniciativa coordenada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundecitrus, para o primeiro sequenciamento de um patógeno de plantas. O sequenciamento completo do genoma da bactéria patogênica *Xylella fastidiosa*, causadora da Clorose Variegada dos Citros (CVC), foi publicado em 2000 (Simpson et al., 2000), o que então permitiu uma comparação detalhada de genomas completos de patógenos de animais e plantas.

Depois dos primeiros genomas completos serem publicados, houve a estruturação de diversos grupos de pesquisa pelo mundo e avanços importantes na evolução e disponibilidade das tecnologias de sequenciamento. Um indicador importante que reflete a popularização da aplicação do sequenciamento genético em diversas áreas da ciência, incluindo a

agricultura, é o custo do sequenciamento. Em 2001, por ocasião da conclusão do sequenciamento do genoma humano, o custo estimado para o sequenciamento completo foi de US\$ 95,3 milhões de dólares. Cerca de 20 anos mais tarde, em 2022, o custo para o sequenciamento de um genoma humano completo foi de US\$ 525 dólares (DNA..., 2024). Essa impressionante redução de custos é justificada pelo surgimento do “sequenciamento de DNA de segunda geração”, que engloba diversas tecnologias de sequenciamento paralelo massivo com ultra rendimento, escalabilidade e velocidade.

Na agricultura, as dimensões física, química e biológica do solo estão intrinsecamente ligadas e determinam o desenvolvimento e a saúde das plantas cultivadas. Em relação ao componente biológico dos solos, nossa visão sobre sua composição de funcionalidade permaneceu, por muito tempo, restrita a grupos de organismos acessíveis por métodos de cultivo, que acessam menos 1% da biodiversidade microbiológica dos solos (Pham; Kim, 2012; Anthony et al., 2023). Neste contexto, a aplicação das técnicas de sequenciamento tem sido fundamental para elucidar as complexas interações entre o microbioma do solo e as plantas. Entretanto, a maior parte dos estudos do microbioma do solo usando sequenciamento de segunda geração ainda é descritiva. A tradução do conhecimento teórico para a aplicação prática na agricultura sustentável exige o desenvolvimento de abordagens e conceitos que proporcionem uma compreensão mais aprofundada dos mecanismos envolvidos nessas interações microbiológicas. Neste texto, vamos colocar em perspectiva as potencialidades e limitações da exploração dos dados de sequenciamento genético do microbioma do solo na agricultura. Ao examinarmos criticamente essas questões, buscamos contribuir para a transição de um entendimento meramente teórico para estratégias práticas que impulsionam a agricultura em direção a padrões mais sustentáveis e eficazes.

O sequenciamento genético do DNA (ou RNA) obtido a partir de solos de sistemas de produção agrícola permite a investigação das interações microbianas no solo, na rizosfera e em outros

ambientes associados às plantas. Os pesquisadores podem estudar os efeitos de práticas agrícolas, incluindo rotação de culturas e uso de fertilizantes e bioinsumos nas comunidades microbianas do solo. Podem também compreender a interação entre o microbioma do solo e plantas cultivadas e direcionar práticas agrícolas sustentáveis que promovem a biodiversidade e/ou a atividade microbiana, resultando em melhorias na produtividade agrícola, suscitada por processos como a ciclagem de nutrientes e outros correlatos à saúde das plantas.

Abordagens e objetivos do sequenciamento de DNA ou RNA do solo

Existem duas abordagens principais que permitem avaliar de forma global a composição e funcionalidade de comunidades microbianas do solo a partir do isolamento e da análise do seu material genético. A primeira abordagem acessa genes marcadores conservados do DNA ribossomal, os quais são usados primariamente para determinar a identidade (taxonomia) dos microrganismos de uma comunidade, dando origem a uma técnica chamada sequenciamento de *amplicons* ou metataxonômica. A segunda, acessa virtualmente todos os segmentos de DNA (ou RNA) obtidos a partir de uma amostra de solo, possibilitando a análise tanto de genes marcadores taxonômicos quanto de genes relacionados a funções específicas presentes no microbioma. Esta abordagem é chamada de metagenômica quando é realizada a partir do DNA, ou metatranscriptômica, quando realizada a partir do RNA. Frequentemente, o termo metagenômica é usado de forma ampla, porém, tecnicamente errônea, quando inclui a abordagem de sequenciamento de *amplicons*. Estes métodos permitem estudar a diversidade genética e funcional dos microrganismos presentes no solo sem a necessidade de cultivá-los em meio de cultura no laboratório, aumentando exponencialmente nossa eficácia na análise dessas comunidades. As estratégias analíticas de cada abordagem estão ilustradas na Figura 1 e descritas nas seções seguintes.

Extração de DNA e RNA do solo

A primeira etapa para o estudo do microbioma do solo é a extração de DNA e/ou RNA das amostras, permitindo assim o acesso ao material genético dos

microrganismos presentes e, subsequentemente, a análise detalhada da diversidade e função microbiana presentes no ambiente a ser estudado. Diferentes métodos de extração são utilizados, cada um com suas vantagens e desvantagens. De maneira geral, o processo envolve a ruptura das células microbianas presentes na amostra de solo, seguida pela purificação do DNA ou RNA. A escolha do método de extração pode afetar significativamente os resultados obtidos, influenciando a eficiência de recuperação do material genético e a representatividade do microbioma estudado.

Além disso, é importante considerar as características específicas de cada solo, como a presença de materiais orgânicos e inorgânicos que podem interferir na extração e na qualidade do material genético recuperado. Métodos que levam em conta a complexidade e a variabilidade do solo, como a extração com agentes desestabilizantes físicos e químicos seguida por purificação cuidadosa, são frequentemente empregados para obter resultados mais precisos. A extração eficaz de DNA e RNA do solo não apenas permite a análise detalhada da composição e função microbiana, mas também contribui para uma melhor compreensão dos processos ecológicos e biogeoquímicos que ocorrem nesse importante compartimento ambiental.

Vale ressaltar que existem diferenças técnicas entre a extração de DNA e RNA de amostras de solo, baseadas principalmente nas características moleculares e nas técnicas de extração específicas necessárias para cada tipo de ácido nucleico. Para o DNA, o desafio principal reside na remoção eficiente de compostos que podem inibir processos posteriores, como por exemplo a reação de PCR, como polifenóis e compostos húmicos presentes no solo. Já a extração de RNA do solo apresenta desafios adicionais devido à sua natureza mais instável e à presença de enzimas ribonucleases (RNases), que podem degradar rapidamente o RNA. Métodos de extração de RNA frequentemente incluem a adição de agentes estabilizadores, como fenol ou guanidina tiocianato, para inativar as RNases e preservar o RNA. Além disso, a extração de RNA muitas vezes requer etapas adicionais, como a remoção de DNA residual por tratamento com DNase e a utilização de colunas de sílica ou resinas específicas para purificar o RNA. Atualmente, diversos kits comerciais estão disponíveis no mercado, oferecendo estratégias rápidas e eficientes para a extração de DNA e RNA a partir de amostras de solo.

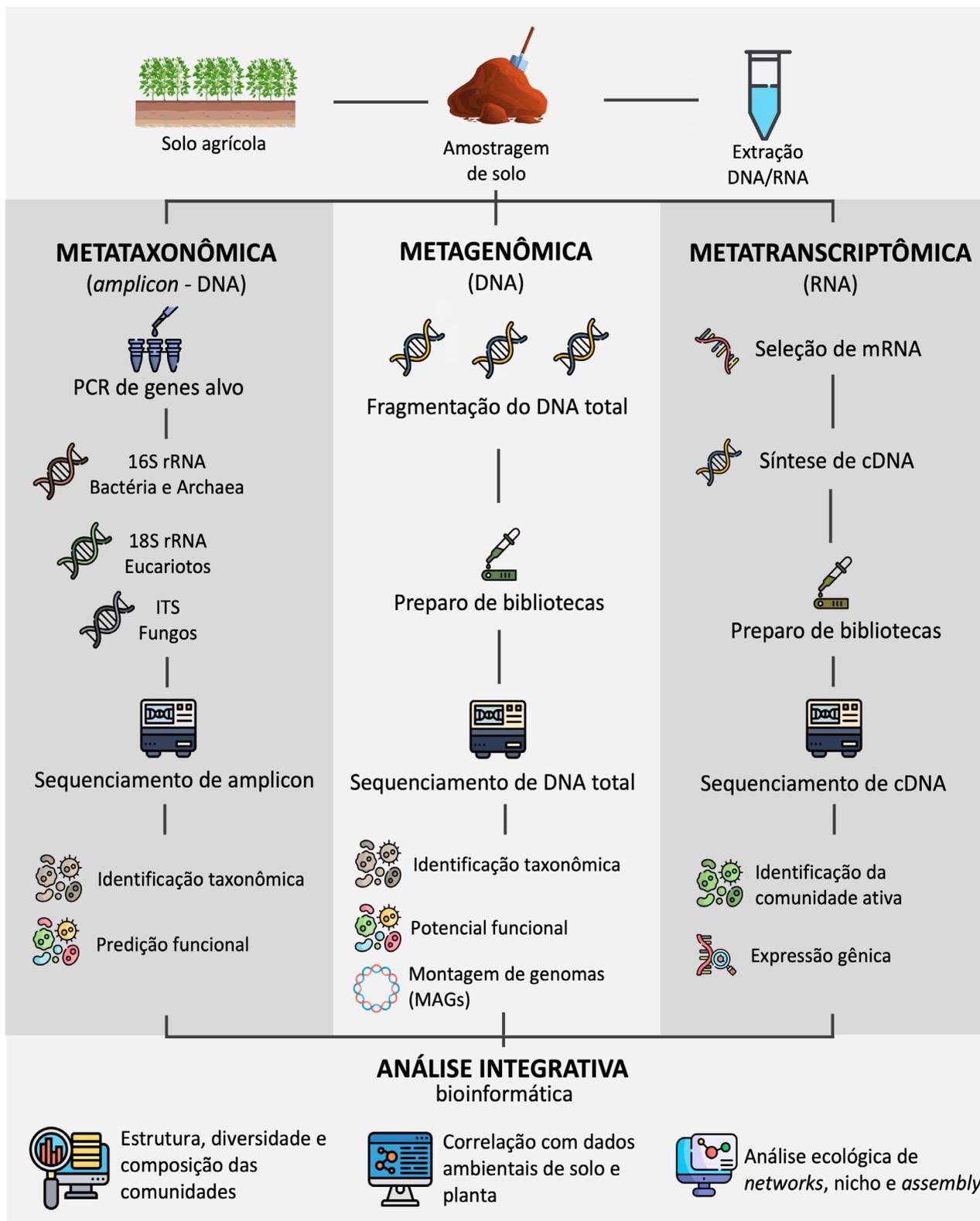


Figura 1. Esquema geral das abordagens analíticas para estudar o microbioma do solo em sistemas agrícolas.

Metataxonômica (sequenciamento de amplicons)

Nesta abordagem, utilizam-se como alvos os genes ribossomais marcadores, que estão presentes em todos os organismos do grupo alvo de estudo. Esses genes, conhecidos como genes essenciais ou *housekeeping*, desempenham atividades fundamentais para a sobrevivência celular. As regiões genômicas ribossomais são comumente escolhidas para este tipo de estudo por apresentarem regiões conservadas e variadas, as quais permitem a identificação precisa de grupos de organismos específicos, como bactérias, fungos e outros eucariotos, e a discriminação de membros destes grupos com resolução taxonômica mais detalhada. As porções variáveis do gene podem ser usadas para estudos de identificação e filogenia. O sequenciamento de amplicons envolve uma amplificação por *Polymerase Chain Reaction* (PCR) usando primers específicos de acordo com o gene alvo, 16S rRNA (bactérias e arqueias), ITS (fungos) e 18S rRNA (outros eucariotos). O produto da PCR é então usado para sequenciamento massivo, gerando milhares de sequências de DNA, as quais são submetidas à análise com ferramentas de bioinformática. Após o processamento da qualidade das sequências obtidas, aquelas de baixa qualidade são removidas, e os agrupamentos de sequências de alta qualidade são empregados para realizar análises de diversidade, identificação e comparação de comunidades. Os bancos de dados mais populares para a atribuição taxonômica são Silva, GreenGenes, Ribosomal Database Project (RDP) para bactérias e arqueias, e alguns específicos como UNITE para fungos e PR2 para protistas. É importante destacar que, devido ao tamanho limitado do fragmento de DNA sequenciado, a classificação das sequências raramente permite a descrição até o nível de espécie dos microrganismos presentes nos solos. A determinação taxonômica das sequências obtidas pode variar desde o nível de domínio até o nível de gênero. Além da afiliação taxonômica, as sequências são também analisadas por similaridade, o que gera classificações de grupos taxonômicos empíricos, a partir dos quais é possível estimar estatisticamente o número de espécies presentes em uma amostra, ou outros parâmetros de diversidade inerentes a esta. Inovações nesta aplicação buscam prever funcionalidades a partir de dados de sequenciamento, bem como traçar correlações negativas ou positivas entre grupos

microbianos presentes numa mesma amostra e determinar grupos microbianos essenciais para a organização da comunidade microbiana em estudo. Todas estas funcionalidades têm sua eficiência relacionada com o conjunto de dados empregado, onde o desenvolvimento da bioinformática e inteligência artificial, como aprendizado de máquina, tem papel fundamental.

Metagenômica (sequenciamento de DNA) e Metatranscriptômica (sequenciamento de RNA)

Nesta abordagem, o DNA ou RNA total é isolado e purificado a partir de amostras de solo, sendo posteriormente processado, fragmentado e utilizado para o sequenciamento. Isso permite acessar todo o material genético recuperado a partir do solo, e não apenas genes marcadores, como na abordagem metataxonômica, permitindo a identificação de bactérias, fungos e outros organismos. Além disso, fornece informações sobre as funções biológicas e o potencial metabólico destes microrganismos. Quando a análise é feita a partir do DNA, a informação dos genes funcionais representa o potencial funcional do microbioma na amostra de solo avaliada. Quando a análise é feita a partir do RNA, os dados do sequenciamento resultam dos genes efetivamente expressos pelo microbioma do solo no momento da amostragem. Essa abordagem fornece uma visão mais abrangente da diversidade microbiana em comparação com métodos tradicionais de cultura, permitindo a compreensão da ecologia microbiana, funções metabólicas e potenciais impactos ambientais no solo. No entanto, estas análises são mais laboriosas, de maior custo e de maior complexidade em sua interpretação. Os bancos de dados comumente usados para a atribuição funcional aos dados metagenômicos são MG-RAST, KEGG, IMG/M, SEED, Uniprot, pFAM. Outros bancos de dados são dedicados para a anotação de via metabólicas específicas, como NcycDB, dbCAN e Resfam, respectivamente para o ciclo do nitrogênio, carboidratos e genes de resistência a antibióticos. Ainda a partir dos dados gerados de metagenômica, há a possibilidade de montagem de genomas de organismos presentes nas amostras, do inglês metagenome-assembled genomes (MAG). Esta técnica computacional tem avançado muito e permite que tenhamos acesso a informações genômicas de organismos não cultivados, o que nos traz assertividade em

conectar funções microbianas no solo com a classificação taxonômica de grupos microbianos pouco estudados.

Roteiro analítico de dados de sequenciamento genético

Para a análise de dados de sequenciamento genético do microbioma, um roteiro abrangente envolve diversas etapas e diferentes ferramentas de bioinformática. Inicialmente, é essencial realizar o pré-processamento das sequências, incluindo verificação da qualidade, remoção de adaptadores e bases de baixa qualidade. Após a filtragem das sequências, é importante conduzir análises de diversidade, estrutura e composição das comunidades microbianas, atribuindo sequências a categorias taxonômicas, calculando métricas de diversidade e comparando a estrutura da comunidade entre diferentes amostras e tratamentos. As visualizações dessas inferências são bastante diversas, variando desde números até dispersões das amostras por similaridade na composição das comunidades. A análise funcional poder ser realizada tanto por predição, a partir de genes ribossomais, quanto por inferência por sequências obtidas a partir das abordagens metagenômica e metatranscriptômica, prevendo funções metabólicas e identificando genes específicos de interesse. Para entender diferenças entre grupos de amostras, testes estatísticos são aplicados, com correção para múltiplas comparações, se necessário. A integração com metadados é crucial para associar dados do microbioma com informações ambientais relevantes. A visualização dos dados por meio de gráficos e figuras ajuda na representação visual dos resultados. Após a interpretação dos resultados à luz dos objetivos do estudo e da literatura existente, é possível discutir as implicações biológicas, além de sugerir estudos futuros. Por fim, é essencial documentar detalhadamente todos os métodos, parâmetros e resultados obtidos para preparação de relatórios ou artigos científicos. Um roteiro analítico de dados de sequenciamento está representado na Figura 2 e descrito a seguir:

1) Pré-processamento dos dados de sequenciamento: verificação da qualidade dos dados brutos do sequenciamento, por exemplo, usando programas como FastQC; remoção de adaptadores e bases sequenciadas de baixa qualidade, por exemplo, utilizando ferramentas

como Trimmomatic ou Cutadapt; e montagem de sequências caso aplicável.

2) Análise de diversidade, estrutura e composição do microbioma: realização de análises para atribuir sequências a categorias taxonômicas utilizando bancos de dados específicos, por meio de softwares como QIIME e mothur; cálculo de métricas de diversidade, como riqueza e diversidade alfa; comparação da estrutura das comunidades microbianas entre os diferentes tratamentos por meio de gráficos de dispersão, por exemplo, utilizando análises multivariadas como PCA, PCoA, NMDS, RDA e CCA; identificação de espécies presentes e sua abundância relativa, por exemplo, utilizando bancos de dados como RDP, Silva, Greengenes; e montagem de genomas (MAGs) a partir de dados do metagenoma.

3) Análise funcional do microbioma: realização de predição funcional a partir de dados de *amplicons*, por exemplo, utilizando banco de dados como FAPROTAX, Tax4Fun2, PICRUSt2 e FUNGuild, e realização de inferência funcional identificando genes de interesse a partir de dados metagenômicos, por exemplo, utilizando bancos de dados como KEGG, SEED e COG.

4) Análise diferencial para comparação de grupos: realização de testes estatísticos para identificar diferenças significativas na composição taxonômica ou funcional de microrganismos entre grupos de amostras utilizando testes paramétricos e não paramétricos, por exemplo, utilizando testes t, ANOVA, DESeq2, edgeR, ALDEx2, ANCOM, LEfSe e Random Forest.

5) Integração de metadados: associação dos dados de sequenciamento do microbioma com informações ambientais ou outras metainformações relevantes para investigar correlações ou padrões, por exemplo, com dados de análise química do solo, ocorrência de doenças e produtividade de culturas. As análises de correlação podem ser realizadas usando métodos estatísticos, como regressão linear, correlação de Spearman, correlação de Pearson, entre outros.

6) Visualização de dados e interpretação: criação de gráficos e figuras para representar visualmente os resultados das análises, como gráficos de barras, diagramas de Venn, heatmaps, análises de network, entre outras; interpretação dos resultados obtidos à luz dos objetivos do estudo e da literatura existente; discussão das implicações biológicas dos resultados observados; e recomendações de manejo com base nas descobertas realizadas.

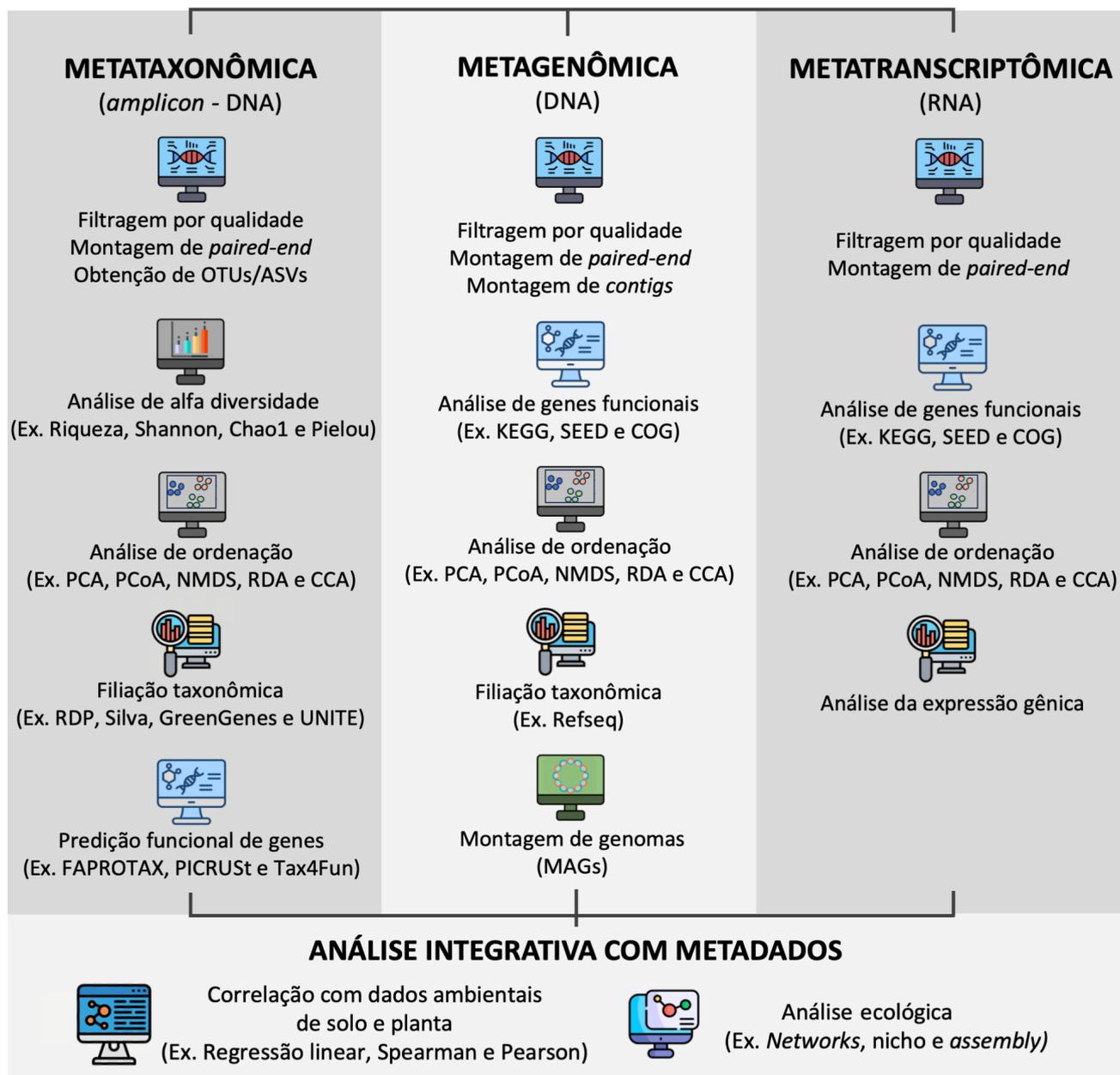


Figura 2. Roteiro analítico para processamento de dados de sequenciamento de material genético obtido de solos.

Oportunidades e desafios do sequenciamento genético do microbioma em sistemas agrícolas

A atividade do microbioma do solo desempenha um papel determinante na produtividade das plantas em sistemas agrícolas (Tkacz; Poole, 2015). Em um estudo de sequenciamento metagenômico, comparando áreas de baixa e alta produtividade agrícola, foi demonstrado que as diferenças na produtividade estão associadas à composição

do microbioma (Chang et al., 2017). O estudo de associação ampla com dados metagenômicos, chamado de *Metagenome-Wide Association Study* (MWAS), combinado com predição de aprendizado de máquina, além de identificar grupos de microrganismos específicos associados à alta produtividade, revelou que a abundância de membros do microbioma associados às transformações do nitrogênio são determinantes para a alta produtividade (Chang et al., 2017). Em um estudo de metataxonômica da comunidade

de bactérias, foi comparado o efeito da diversidade microbiana na produtividade de soja e na taxa de infecção pelo nematoide das galhas (*Meloidogyne* spp.). Os resultados mostraram que solos com maior diversidade bacteriana apresentaram menor taxa de infecção pelo patógeno e maior produtividade (Barros et al., 2022), evidenciando a importância da diversidade microbiana do solo para a saúde e o crescimento de plantas. Assim, considerando a essencialidade do microbioma para as plantas, a análise de sequenciamento genético do microbioma do solo em sistemas agrícolas tem o potencial de prever a saúde e a produtividade dos sistemas agrícolas.

Há um esforço da comunidade científica no sentido de organizar um sistema robusto de indicadores para monitorar a saúde do solo. Considerando que o microbioma contém informações sobre o estado biológico, químico e físico do solo, ele pode servir como uma medida integrada da saúde do solo. Wilhelm et al. (2022) demonstraram o potencial de utilizar dados do microbioma do solo para prever a saúde do solo de forma precisa e eficiente. Ao analisar um conjunto de dados de sequenciamento de *amplicons* do gene 16S rRNA, os pesquisadores desenvolveram modelos de aprendizado de máquina capazes de prever diversas métricas de saúde do solo, incluindo propriedades biológicas, químicas e físicas. Os resultados indicam que a composição do microbioma do solo pode servir como um indicador confiável da qualidade do solo, permitindo a criação de ferramentas de diagnóstico mais precisas e acessíveis para monitorar a saúde dos agroecossistemas (Wilhelm et al., 2022). Em um outro estudo realizado para prever com alta acurácia a ocorrência de doença em seis culturas, os autores investigaram a relação entre a distribuição dos microrganismos no solo e a ocorrência da murcha de *Fusarium*. Utilizando técnicas de análise de dados complexas, eles identificaram que a presença de determinados tipos de fungos no solo estava fortemente correlacionada com o aparecimento da doença. O método identificou 45 OTUs (*operational taxonomic units*) bacterianas e 40 OTUs fúngicas que categorizaram o estado de saúde do solo com uma precisão superior a 80%. A conclusão do estudo sugere que a análise da composição da comunidade microbiana do solo pode ser uma ferramenta útil para prever e controlar essa doença em cultivos agrícolas (Yuan et al., 2020). O sequenciamento, usando *primers* para a região ITS e primers que têm como alvo fungos micorrízicos arbusculares, foi usado para prever o sucesso da inoculação do fungo em 54 campos de milho distribuídos na Suíça.

Os indicadores do microbioma do solo, juntamente com alguns parâmetros do solo, conseguiram prever 86% da variação no crescimento do milho em resposta à inoculação. A abundância de fungos patogênicos, e não a disponibilidade de nutrientes, foi o melhor preditor do sucesso do inoculante. Esta capacidade de previsão amplia o potencial do uso do microbioma como uma ferramenta para o manejo agrícola sustentável (Lutz et al., 2023).

Quanto à detecção de patógenos e suas implicações, existem diversas particularidades que devem ser consideradas quando empregamos as técnicas de sequenciamento genético para prever ocorrência de doenças. É inegável o potencial do sequenciamento genético em identificar grupos microbianos, dentre os quais estão os patógenos de plantas. No entanto, quando resgatamos materiais genéticos de amostras complexas, os fragmentos sequenciados, tanto por metataxonômica como por metagenômica, normalmente não trazem uma resolução para inferir sobre o potencial patogênico dos organismos detectados. Isso ocorre porque, para muitos grupos com este potencial, a capacidade infectiva é bastante específica, seja a uma espécie ou variação dentro da espécie (por exemplo, forma *specialis*), ou relacionada a um ou alguns genes presentes em linhagens com maior potencial de infecção. Nesse caso, análises moleculares complementares, como o uso de PCR quantitativo direcionado a patógenos específicos, podem ser úteis para a identificação e quantificação precisas do patógeno em amostras de solo.

Os desafios que ainda permanecem para usar informações do microbioma para prever saúde do solo ou produtividade incluem aumentar a profundidade do sequenciamento, i.e. obter um maior número de sequências por amostra de solo; aumentar a resolução na classificação taxonômica; e aumentar o número de amostras que representem a variação e diversidade de solos com diferentes níveis de saúde (Wilhelm et al., 2022). Adicionalmente, os modelos de predição baseados no sequenciamento do microbioma irão ter um melhor desempenho à medida que eles sejam personalizados e adaptados considerando o tipo de solo, a biogeografia com as diferenças regionais do microbioma do solo (Gschwend et al., 2021) e o sistema de produção, devido ao efeito do legado das plantas cultivadas no sistema (Schmid et al., 2019).

Considerações finais

O sequenciamento genético do microbioma do solo representa uma ferramenta promissora para o manejo agrícola sustentável, ajudando a prever a saúde do solo, a ocorrência de doenças e a produtividade. Contudo, ainda há desafios na padronização e otimização dessas técnicas para que elas possam ser amplamente aplicadas. A combinação de dados de sequenciamento com inteligência artificial e aprendizado de máquina pode acelerar essa transição, permitindo uma abordagem mais eficiente e sustentável na agricultura. A evolução das tecnologias de sequenciamento, aliada ao avanço de ferramentas analíticas baseadas em inteligência artificial, tem o potencial de transformar radicalmente a gestão agrícola. No futuro, espera-se que essas tecnologias sejam cada vez mais acessíveis e personalizadas, adaptando-se a diferentes tipos de solo e sistemas de cultivo. Além disso, o desenvolvimento de bancos de dados mais robustos e modelos preditivos regionais contribuirá para um uso mais eficaz do microbioma do solo, garantindo uma agricultura mais sustentável e produtiva.

Referências

- ANTHONY, M. A.; BENDER, S. F.; VAN DER HEIJDEN, M. G. A. Enumerating soil biodiversity. **Proceedings of the National Academy of Sciences of the United States of America**, v. 120, n. 33, p. e2304663120, Aug. 2023. DOI: <https://doi.org/10.1073/pnas.2304663120>.
- BARROS, F.; PEDRINHO, A.; MENDES, L. W.; FREITAS, C. C. G.; ANDREOTE, F. D. Interactions between soil bacterial diversity and plant-parasitic nematodes in soybean plants. **Applied and Environmental Microbiology**, v. 88, n. 17, p. e0096322, Sept. 2022. DOI: <https://doi.org/10.1128/aem.00963-22>.
- CHANG, H. X.; HAUDENSHIELD, J. S.; BOWEN, C. R.; HARTMAN, G. L. Metagenome-Wide Association Study and Machine Learning Prediction of Bulk Soil Microbiome and Crop Productivity. **Frontiers in Microbiology**, v. 8, p. 519, Apr. 2017. DOI: <https://doi.org/10.3389/fmicb.2017.00519>.
- DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). **National Human Genome Research Institute**. Disponível em: <http://www.genome.gov/sequencingcostsdata>. Acesso em: 18 nov. 2024.
- GENOME sequence of the nematode *C. elegans*: a platform for investigating biology. **Science**, v. 282, n. 5396, p. 2012-2018, Dec. 1998. DOI: <https://doi.org/10.1126/science.282.5396.2012>.
- GSCHWEND, F.; HARTMANN, M.; HUG, A. S.; ENKERLI, J.; GUBLER, A.; FREY, B.; MEULI, R. G.; WIDMER, F. Long-term stability of soil bacterial and fungal community structures revealed in their abundant and rare fractions. **Molecular Ecology**, v. 30, n. 17, p. 4305-4320, Sept. 2021. DOI: <https://doi.org/10.1111/mec.16036>.
- LANDER, E. S.; LINTON, L. M.; BIRREN, B.; NUSBAUM, C.; ZODY, M. C.; BALDWIN, J.; DEVON, K.; DEWAR, K.; DOYLE, M.; FITZHUGH, W.; FUNKE, R.; GAGE, D.; HARRIS, K.; HEAFORD, A.; HOWLAND, J.; KANN, L.; LEHOCZKY, J.; LEVINE, R.; MCEWAN, P.; MCKERNAN, K.; MELDRIM, J.; MESIROV, J. P.; MIRANDA, C.; MORRIS, W.; NAYLOR, J.; RAYMOND, C.; ROSETTI, M.; SANTOS, R.; SHERIDAN, A.; SOUGNEZ, C.; STANGETHOMANN, Y.; STOJANOVIC, N.; SUBRAMANIAN, A.; WYMAN, D.; ROGERS, J.; SULSTON, J.; AINSCOUGH, R.; BECK, S.; BENTLEY, D.; BURTON, J.; CLEE, C.; CARTER, N.; COULSON, A.; DEADMAN, R.; DELOUKAS, P.; DUNHAM, A.; DUNHAM, I.; DURBIN, R.; FRENCH, L.; GRAFHAM, D.; GREGORY, S.; HUBBARD, T.; HUMPHRAY, S.; HUNT, A.; JONES, M.; LLOYD, C.; MCMURRAY, A.; MATTHEWS, L.; MERCER, S.; MILNE, S.; MULLIKIN, J.C.; MUNGALL, A.; PLUMB, R.; ROSS, M.; SHOWNKEEN, R.; SIMS, S.; WATERSTON, R. H.; WILSON, R. K.; HILLIER, L. W.; MCPHERSON, J. D.; MARRA, M. A.; MARDIS, E. R.; FULTON, L. A.; CHINWALLA, A. T.; PEPIN, K. H.; GISH, W. R.; CHISSOE, S. L.; WENDL, M. C.; DELEHAUNTY, K. D.; MINER, T. L.; DELEHAUNTY, A.; KRAMER, J. B.; COOK, L. L.; FULTON, R. S.; JOHNSON, D. L.; MINX, P. J.; CLIFTON, S. W.; HAWKINS, T.; BRANSCOMB, E.; PREDKI, P.; RICHARDSON, P.; WENNING, S.; SLEZAK, T.; DOGGETT, N.; CHENG, J. F.; OLSEN, A.; LUCAS, S.; ELKIN, C.; UBERBACHER, E.; FRAZIER, M.; GIBBS, R. A.; MUZNY, D. M.; SCHERER, S. E.; BOUCK, J. B.; SODERGREN, E. J.; WORLEY, K. C.; RIVES, C. M.; GORRELL, J. H.; METZKER, M. L.; NAYLOR, S. L.; KUCHERLAPATI, R. S.; NELSON, D. L.; WEINSTOCK, G. M.; SAKAKI, Y.; FUJIYAMA, A.; HATTORI, M.; YADA, T.; TOYODA, A.; ITOH, T.; KAWAGOE, C.; WATANABE, H.; TOTOKI, Y.; TAYLOR, T.; WEISSENBACH, J.; HEILIG, R.; SAURIN, W.; ARTIGUENAVE, F.; BROTTIER, P.; BRULS, T.; PELLETIER, E.; ROBERT, C.; WINCKER, P.; SMITH, D. R.; DOUCETTE-STAMM, L.; RUBENFIELD, M.; WEINSTOCK, K.; LEE, H. M.; DUBOIS, J.; ROSENTHAL, A.; PLATZER, M.; NYAKATURA, G.; TAUDIEN, S.; RUMP, A.; YANG, H.; YU, J.; WANG, J.; HUANG, G.; GU, J.; HOOD, L.; ROWEN, L.; MADAN, A.; QIN, S.; DAVIS, R. W.; FEDERSPIEL, N. A.; ABOLA, A. P.; PROCTOR, M. J.; MYERS, R. M.; SCHMUTZ, J.; DICKSON, M.; GRIMWOOD, J.; COX, D. R.; OLSON, M. V.; KAUL, R.; RAYMOND, C.; SHIMIZU, N.; KAWASAKI, K.; MINOSHIMA, S.; EVANS, G. A.; ATHANASIOU, M.;

- SCHULTZ, R.; ROE, B. A.; CHEN, F.; PAN, H.; RAMSER, J.; LEHRACH, H.; REINHARDT, R.; MCCOMBIE, W. R.; DE LA BASTIDE, M.; DEDHIA, N.; BLÖCKER, H.; HORNISCHER, K.; NORDSIEK, G.; AGARWALA, R.; ARAVIND, L.; BAILEY, J. A.; BATEMAN, A.; BATZOGLOU, S.; BIRNEY, E.; BORK, P.; BROWN, D. G.; BURGE, C. B.; CERUTTI, L.; CHEN, H. C.; CHURCH, D.; CLAMP, M.; COPLEY, R. R.; DOERKS, T.; EDDY, S. R.; EICHLER, E. E.; FUREY, T. S.; GALAGAN, J.; GILBERT, J. G.; HARMON, C.; HAYASHIZAKI, Y.; HAUSSLER, D.; HERMJAKOB, H.; HOKAMP, K.; JANG, W.; JOHNSON, L. S.; JONES, T. A.; KASIF, S.; KASPRYZK, A.; KENNEDY, S.; KENT, W. J.; KITTS, P.; KOONIN, E. V.; KORF, I.; KULP, D.; LANCET, D.; LOWE, T. M.; MCLYSAGHT, A.; MIKKELSEN, T.; MORAN, J. V.; MULDER, N.; POLLARA, V. J.; PONTING, C. P.; SCHULER, G.; SCHULTZ, J.; SLATER, G.; SMIT, A. F.; STUPKA, E.; SZUSTAKOWKI, J.; THIERRY-MIEG, D.; THIERRY-MIEG, J.; WAGNER, L.; WALLIS, J.; WHEELER, R.; WILLIAMS, A.; WOLF, Y. I.; WOLFE, K. H.; YANG, S. P.; YEH, R. F.; COLLINS, F.; GUYER, M. S.; PETERSON, J.; FELSENFELD, A.; WETTERSTRAND, K. A.; PATRINOS, A.; MORGAN, M. J.; DE JONG, P.; CATANESE, J. J.; OSOEGAWA, K.; SHIZUYA, H.; CHOI, S.; CHEN, Y. J.; SZUSTAKOWKI, J.; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. **Nature**, v. 409, n. 6822, p. 860-921, Feb. 2001. DOI: <https://doi.org/10.1126/aem.00963-22>.
- LUTZ, S.; BODENHAUSEN, N.; HESS, J.; VALZANO-HELD, A.; WAELCHLI, J.; DESLANDES-HÉROLD, G.; SCHLAEPPI, K.; VAN DER HEIJDEN, M. G. A. Soil microbiome indicators can predict crop growth response to large-scale inoculation with arbuscular mycorrhizal fungi. **Nature Microbiology**, v. 8, n. 12, p. 2277-2289, Dec. 2023. DOI: <https://doi.org/10.1038/s41564-023-01520-w>.
- PHAM, V. H.; KIM, J. Cultivation of unculturable soil bacteria. **Trends in Biotechnology**, v. 30, n. 9, p. 475-484, Sept. 2012. DOI: <https://doi.org/10.1016/j.tibtech.2012.05.007>.
- SCHMID, M. W.; HAHL, T.; VAN MOORSEL, S. J.; WAGG, C.; DENY, G. B. de; SCHMID, B. Feedbacks of plant identity and diversity on the diversity and community composition of rhizosphere microbiomes from a long-term biodiversity experiment. **Molecular Ecology**, v. 28, n. 4, p. 863-878, Feb. 2019. DOI: <https://doi.org/10.1111/mec.14987>.
- SCHMUTZ, J.; WHEELER, J.; GRIMWOOD, J.; DICKSON, M.; YANG, J.; CAOILE, C.; BAJOREK, E.; BLACK, S.; CHAN, Y. M.; DENYS, M.; ESCOBAR, J.; FLOWERS, D.; FOTOPULOS, D.; GARCIA, C.; GOMEZ, M.; GONZALES, E.; HAYDU, L.; LOPEZ, F.; RAMIREZ, L.; RETTERER, J.; RODRIGUEZ, A.; ROGERS, S.; SALAZAR, A.; TSAI, M.; MYERS, R. M. Quality assessment of the human genome sequence. **Nature**, v. 429, n. 6990, p. 365-368, May 2004. DOI: <https://doi.org/10.1038/nature02390>.
- SIMPSON, A. J.; REINACH, F. C.; ARRUDA, P.; ABREU, F. A.; ACENCIO, M.; ALVARENGA, R.; ALVES, L. M.; ARAYA, J. E.; BAIA, G. S.; BAPTISTA, C. S.; BARROS, M. H.; BONACCORSI, ED.; BORDIN, S.; BOVÉ, J. M.; BRIONES, M. R.; BUENO, M. R.; CAMARGO, A. A.; CAMARGO, L. E.; CARRARO, D. M.; CARRER, H.; COLAUTO, N. B.; COLOMBO, C.; COSTA, F. F.; COSTA, M. C.; COSTA-NETO, C. M.; COUTINHO, L. L.; CRISTOFANI, M.; DIAS-NETO, E.; DOCENA, C.; EL-DORRY, H.; FACINCANI, A. P.; FERREIRA, A. J.; FERREIRA, V. C.; FERRO, J. A.; FRAGA, J. S.; FRANÇA, S. C.; FRANCO, M. C.; FROHME, M.; FURLAN, L. R.; GARNIER, M.; GOLDMAN, G. H.; GOLDMAN, M. H.; GOMES, S. L.; GRUBER, A.; HO, P. L.; HOHEISEL, J. D.; JUNQUEIRA, M. L.; KEMPER, E. L.; KITAJIMA, J. P.; KRIEGER, J. E.; KURAMAE, E. E.; LAIGRET, F.; LAMBAIS, M. R.; LEITE, L. C.; LEMOS, E. G.; LEMOS, M. V.; LOPES, A. S.; LOPES, C. R.; MACHADO, J. A.; MACHADO, M. A.; MADEIRA, A. M.; MADEIRA, H. M.; MARINO, C. L.; MARQUES, M. V.; MARTINS, E. A.; MARTINS, E. M.; MATSUKUMA, A. Y.; MENCK, C. F.; MIRACCA, E. C.; MIYAKI, C. Y.; MONTERIRO-VITORELLO, C. B.; MOON, D. H.; NAGAI, M. A.; NASCIMENTO, A. L.; NETTO, L. E.; NHANI JUNIOR, A.; NOBREGA, F. G.; NUNES, L. R.; OLIVEIRA, M. A.; OLIVEIRA, M. C. de; OLIVEIRA, R. C. de; PALMIERI, D. A.; PARIS, A.; PEIXOTO, B. R.; PEREIRA, G. A.; PEREIRA JUNIOR, H. A.; PESQUERO, J. B.; QUAGGIO, R. B.; ROBERTO, P. G.; RODRIGUES, V.; ROSA, A. J. de M.; ROSA JUNIOR, V. E. de; SÁ, R. G. de; SANTELLI, R. V.; SAWASAKI, H. E.; SILVA, A. C. da; SILVA, A. M. da; SILVA, F. R. da; SILVA, W. A.; SILVEIRA, J. F. da; SILVESTRI, M. L.; SIQUEIRA, W. J.; SOUZA, A. A. de; SOUZA, A. P. de; TERENCE, M. F.; TRUFFI, D.; TSAI, S. M.; TSUHAKO, M. H.; VALLADA, H.; VAN SLUYS, M. A.; VERJOVSKI-ALMEIDA, S.; VETTORE, A. L.; ZAGO, M. A.; ZATZ, M.; MEIDANIS, J.; SETUBAL, J. C. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. **Nature**, v. 406, n. 6792, p. 151-159, July. 2000. DOI: <https://doi.org/10.1038/35018003>.
- TKACZ, A.; POOLE, P. Role of root microbiota in plant productivity. **Journal of Experimental Botany**, v. 66, n. 8, p. 2167-2175, Apr. 2015. DOI: <https://doi.org/10.1093/jxb/erv157>.
- WILHELM, R. C.; VAN ES, H. M.; BUCKLEY, D. H. Predicting measures of soil health using the microbiome and supervised machine learning. **Soil Biology and Biochemistry**, v. 164, 108472, Jan. 2022. DOI: <https://doi.org/10.1016/j.soilbio.2021.108472>.
- YUAN, J.; WEN, T.; ZHANG, H.; ZHAO, M.; PENTON, C. R.; THOMASHOW, L. S.; SHEN, Q. Predicting disease occurrence with high accuracy based on soil macroecological patterns of *Fusarium* wilt. **The ISME Journal**, v. 14, n. 12, p. 2936-2950, Dec. 2020. DOI: <https://doi.org/10.1038/s41396-020-0720-5>.

