

Tallis Dixon

T00696719

Oct 29th 2024

Microbial Ecology

Methods Assignment

1. Microbial community composition & diversity via amplicon sequencing (e.g., 16S & 18S rRNA gene regions, ITS regions).

Method:

Amplicon sequencing largely remains an important tool for the identification of microbial organisms (Principles and Workflow, 2020). This process begins with the extraction of DNA from a sample, followed by using PCR to amplify a targeted sequence, typically the 16S rRNA gene in prokaryotes, ITS gene regions in fungi and the 18S rRNA gene in eukaryotes (Principles and Workflow, 2020). Three components are essential to initiate PCR: primers specific to the 16S rRNA gene, polymerases capable of withstanding high temperatures, and a supply of free nucleotides (Bottos, 2024). The primers bind to their complementary base pairs on the conserved regions of the 16S rRNA gene (Principles and Workflow, 2020). DNA polymerase can bind to the primer and attaches the first nucleoside to the 3' OH group and continue synthesis down the strand (Principles and Workflow, 2020). The PCR process cycles through denaturation, annealing, and elongation phases, during which DNA is unwound, primers attach to the template strand, and polymerase synthesizes complementary DNA strands (Principles and Workflow, 2020). Once the DNA region of interest is amplified, adapters or barcodes are added to the PCR products, allowing each sample to be tracked to its source during sequencing; these are called libraries (Principles and Workflow, 2020). The barcoded DNA is then sequenced with next-generation sequencing technologies, such as Illumina, which uses fluorescently labelled nucleotides that emit unique signals as they are incorporated into the growing DNA strand (Principles and Workflow, 2020). After sequencing, researchers process the raw data to assess the composition of microbial communities within the samples (Principles and Workflow, 2020). The conserved nature of the 16S rRNA gene, along with small mutations that accumulate over time, allows researchers to use it as a molecular clock to trace evolutionary relationships and lineage

divergence through mutations (Bottos 2024). The grouping of these sequences into related groups makes Operational Taxonomic Units (OTU).

Information obtained:

Sequencing generates millions of short, high accuracy reads, providing raw data on nucleotide variation and abundance within the target region (Principles and Workflow, 2020). This raw data undergoes filtering to remove low-quality reads, short reads, and chimeric sequences, which enhances data quality (Principles and Workflow, 2020). After this cleanup, sequences are grouped based on similarity and then classified taxonomically into operational taxonomic units (OTUs) (Metagenomics, 2024). Typically, a 97% sequence similarity threshold is used to assign organisms to the same species group (Metagenomics, 2024).

Example:

In the article by Pascoal et al. 2024, they used full-length 16S rRNA gene sequencing to classify Arctic marine prokaryotes. Seawater samples were collected from various depths, DNA extraction was followed by PCR and 16S rRNA gene amplicon sequencing using both Illumina for the V4-V5 region and PacBio for the full-length gene (Pascoal et al. 2024). After sequencing, the raw data was analyzed to analyze the diversity within microbial communities (Pascoal et al. 2024). The results of sequencing the two different lengths of genes showed that using Illumina from the V4-V5 region is faster and cheaper but less accurate for species and genus classification (Pascoal et al. 2024). Whereas full-length sequencing using PacBio was more time-consuming and costly but increased the microbial diversity observed in the sample (Pascoal et al. 2024).

Advantages:

Amplicon sequencing using methods like next-generation sequencing is advantageous because it can sequence DNA from multiple prokaryotes in one sample without the need to culture organisms (Principles and Workflow, 2020). It is also faster and more specific due to it being focused on a small region within the genome, not

requiring sequencing of the entire genome which is time-consuming and costly (Principles and Workflow, 2020).

Disadvantages:

Often limited to short sequences which can cause difficulty in species identification and limits the information gained (Principles and Workflow, 2020). As seen in the paper by Pascoal et al. 2024, they compared short reads versus long and found a greater diversity of microbial communities at the species and genus levels as compared to mostly only the phyla level with shorter reads. PCR amplification bias can also occur where there is a skewed distribution of PCR products due to unequal amplification (Acinas et al. 2005).

2. Metagenomics for microbial or viral communities (NOT 16S and 18S rRNA sequencing).

Method:

Metagenomics studies the entire microbial community within a sample, analyzing the DNA of all microorganisms in an environment (including bacteria, archaea, fungi, and viruses) (Yu et al. 2020). Samples from sources like soil, water, and tissues are collected, and DNA is extracted from all organisms present using extraction kits (Yu et al. 2020). To prepare the DNA for sequencing, the extracted genetic material is randomly fragmented using methods like sonication and then tagged with markers via enzymes (transposases or ligases) (Yu et al. 2020). PCR amplification can be used to increase the amount of microbial DNA, making high enough concentrations for sequencing (Yu et al. 2020). In shotgun metagenomics, PCR is generally not employed as it requires primers for polymerases to bind and replicate which means there would be a specific target and not the whole genome like in shotgun (Yu et al. 2020).

Next-generation sequencing technology then sequences the prepared DNA libraries using fluorescently labelled nucleotides that emit unique signals during synthesis, allowing real-time, high-accuracy read generation (Yu et al. 2020). These millions of short DNA reads are then assembled into longer sequences, or contigs, through overlapping regions

(Lee 2019). This assembly enables two primary methods for analyzing community composition: matching contigs to known reference genomes or grouping DNA fragments into metagenome-assembled genomes (MAGs), which are scaffolded based on shared characteristics like tetranucleotide frequencies and complementary marker genes (Xu et al. 2014) (Yang et al.). MAGs allow for taxonomic classification and help determine overall community composition, providing a detailed look at both the diversity and potential functional roles within the microbial community (Yang et al.).

Information obtained:

Sequencing data includes short reads of nucleotide sequences after examining the data quality and filtering out low-quality reads using bioinformatics these can be reconstructed into MAGs to make up genomic sequences (Setubal 2021). This combines overlapping reads into longer contiguous sequences using software (Setubal 2021). MAGs convert the short fragmented sequences into long continuous sequences which gives a picture of the entire genome, making it easier to find genes and identify microbes (Setubal 2021).

Example:

In the paper by Lan et al. 2024, the microbiome, including bacteria, archaea and viruses was conducted on drinking water using metagenomic sequencing. Following DNA extraction, samples were sheared with an interruption enzyme to a size of ~300bp and ligated with tags for Illumina sequencing with further PCR amplification (Lan et al. 2024). PCR products were then isolated, and Illumina NGS was used to extract high-quality clean reads (Lan et al. 2024). Metagenomics was used here to analyze not just microbes but also viral content, pathways, and gene abundances for the evaluation of safe drinking water (Lan et al. 2024). The results showed that metagenomics was able to identify a larger range of microorganisms in DW as compared to culture-dependent methods and reverse-transcription PCR (Lan et al. 2024).

Advantages:

Allows the profiling of all microorganisms in the sample, including viruses, bacteria, archaea, and fungi (Yu et al. 2020). Without selective amplification, researchers can look at the entire genome and identify genes associated with functions like metabolism, antibiotic resistance, and virulence (Yu et al. 2020).

Disadvantages:

Metagenomic analysis requires extensive bioinformatics and analysis due to the non-targeted nature of the analysis (Yu et al. 2020). This makes it more expensive, time-consuming, and prone to mistakes when very similar genomes are compared. Metagenomics also cannot tell which genes are active in an environment, therefore introducing possible bias as to whether present genes are currently doing anything in the environment (Yu et al. 2020).

3. Microbial community transcriptomics (whole-transcriptome shotgun sequencing; RNA-seq).

Method:

To analyze the transcriptome of a microbial community, all RNA types (mRNA, tRNA, rRNA, and non-coding RNA) are first isolated from the sample (Yang et al. 2015). This involves lysing cells, deactivating RNases (by heat), adding DNases (to remove DNA), and removing debris to ensure a clean RNA extract (Lowe et al. 2017). Since RNA is predominantly rRNA and tRNA, these molecules are removed to enrich for mRNA using terminator 5'-phosphate-dependent exonuclease, which degrades rRNA and tRNA by removing their 5'-caps, 5'-hydroxyl groups, or 5'-triphosphates (Ma et al. 2019). After mRNA isolation, polyadenylation adds a poly-A tail to the mRNA to prevent degradation and prepare for amplification and labelling (Ma et al. 2019). RNA is fragmented into smaller pieces, and reverse transcription is performed to convert fragmented RNA into complementary DNA (cDNA) using reverse transcriptase (Yang et al. 2015). Adapter sequences are ligated to the ends of each cDNA fragment to enable sample identification to the sequencing platform (Yang et al. 2015). Amplification through PCR may be used to increase the library's complexity and ensure enough

material for sequencing if targeting a particular gene (Yang et al. 2015). The cDNA library is sequenced on a next-generation sequencing platform (e.g., Illumina) (Yang et al. 2015). Transcriptomics allows the study of active genes in the environment; this allows researchers to track microbial responses.

Information obtained:

This generates millions of short reads of cDNA, which represent the fragments of all RNA molecules present in the original sample (Yang et al. 2015). These are represented in raw data files containing the sequence data of the string of nucleotides (Novak et al. 2024). Gene counts can be used to quantify the number of times a gene was detected, and identifying different variants and expression levels can enable detailed information about community dynamics (Novak et al. 2024). Understanding which genes are active in different environmental conditions allows researchers to understand how microbes adapt and survive in varied ecosystems (Novak et al. 2024).

Example:

This study by Novak et al. 2024, underscores the role of bacteria in decomposing fungal necromasses and how their metabolic activity changes based on the melanin content of the necromass. Samples were collected from high and low melanin necromasses along different time points to evaluate the different metabolic activity at different stages (exponential/stationary) (Novak et al. 2024). Transcriptomics was used to collect sequence information to determine the expression levels across the different samples. The results showed higher expression levels on low-melanin than on high-melanin (Novak et al. 2024).

Advantages:

Transcriptomics allows a complete analysis of all RNA in the sample; mRNA can be isolated to identify active biological and regulatory pathways (Yang et al. 2015). mRNA is transcribed from the genome to be translated into proteins, by isolating the mRNA researchers can reverse transcribe it back into cDNA (without any introns) and sequence the exact DNA sequence the cell is actively using (Yang et al. 2015). This is a

useful tool for monitoring microbial responses to changes in an environment (Yang et al. 2015).

Disadvantages:

Transcriptomics is much more extensive and time-consuming due to the amount of genetic information obtained (Yang et al. 2015). To deal with large datasets, bioinformatics and computer processing is a time-consuming process (Yang et al. 2015). RNA is also more susceptible to degradation requiring more specialized equipment to prevent degradation during isolating steps (Ma et al. 2019). Another, concern is when a low abundance of RNA is present in bacterial cells, this requires amplification and primers which can introduce PCR bias or require targeting of specific genes instead of the whole genome (Ma et al. 2019).

4. Microbial community proteomics.

Method:

Proteins are extracted and purified from microbial samples using organic solvents, detergents, or physical methods to lyse cells and release the proteins (Doerr et al.). Once extracted, proteins are commonly separated by gel electrophoresis or liquid chromatography (Tamang, 2023). For separation by gel electrophoresis, SDS is used to denature the proteins, binding along the backbone to create a consistent mass-to-charge ratio (Tamang, 2023). The negative charge from SDS enables proteins to migrate through the gel towards the positive end, with larger proteins moving more slowly than smaller ones (Tamang, 2023). After separation, proteins are excised from the gel and digested into peptides, typically using the enzyme trypsin, which cleaves at the carboxyl terminal of lysine and arginine residues (Karpievitch et al. 2010). The peptides can then be analyzed using liquid chromatography-mass spectrometry (LC-MS) (Introduction, 2023). Liquid chromatography separates the peptides based on their interaction with the mobile and stationary phases, allowing them to enter the mass spectrometer sequentially (Introduction, 2023). In mass spectrometry, the peptides are ionized, either by

electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI). In MALDI, for instance, peptides are mixed with a UV-absorbing matrix, which, when pulsed with a laser, transfers energy to the peptides and causes them to become ionized (Introduction, 2023). The ionized peptides are then directed into the mass spectrometer, where they are separated and analyzed based on their mass-to-charge ratio (m/z) (Introduction, 2023). Ions are accelerated down a tube, with lighter ions moving faster and heavier ions taking longer this is called time-of-flight, the time it takes the ions to exit the mass analyzer and hit the detector generates spectra (Introduction, 2023). The mass spectrum displays peaks that correspond to different peptides, where the relative abundance and types of peptides can be determined from the literature (Introduction, 2023). This proteomics approach provides a detailed profile of the proteins in a microbial community, giving insights into the functional roles of different microbes within the community (Introduction, 2023).

Information obtained:

Once peptides have been sequenced, bioinformatics programs can be used to search for the identity of a protein in a database of known proteins (Introduction, 2023). This generates a mass spectrum spanning the m/z values that were measured (Introduction, 2023). Spectra are compared against theoretical spectra in databases that contain possible peptide/protein sequences (Introduction, 2023).

Example:

In this study by Bielen et al. 2024, mass spectrometry (MALDI-TOF MS) and 16S rDNA sequencing were compared for their effectiveness in identifying bacteria associated with cave-dwelling mussels, *Congerja jalzici*. Bacterial colonies were cultured in the lab, and proteins were extracted by means of chemical extraction and centrifugation. MALDI was used to ionize the peptides via a matrix solution that helps transfer laser energy to the peptides. MS was used to identify the proteins in the sample. The resulting mass spectra, which represent protein "fingerprints" unique to each bacterial sample, were analyzed and matched against known databases.

Advantages:

Unlike genomics or transcriptomics, proteomics directly measures the proteins: the molecules functioning in the cell (Tamang, 2023). Proteomics can identify post-translational modifications which are critical for protein expression, but not seen in genomic methods (Tamang, 2023). Using this method, researchers can identify changes in post-transcriptional modifications in response to different environment stressors, temperature, pH, salinity, etc (Tamang, 2023).

Disadvantages:

Analyzing the peaks on MS can be difficult due to isomers (different molecules with the same mass) and isotopes (same charge with different masses caused by different numbers of neutrons) (Tamang, 2023). Other limitations include identification being reliant on current databases, therefore misidentifications are possible (Wang et al. 2016).

5. Microbial community metabolomics.

Method:

Metabolomics is the identification of metabolites that correspond to a biological phenotype within a given environment (Sudeshna et al. 2024). Samples of interest can be collected either from a growth culture in the lab or from the environment. Metabolites are highly sensitive to environmental changes like pH, temperature and pressure so once collected samples are frozen in liquid nitrogen and stored at -80°C (Kamal et al. 2022). Metabolite extraction steps include quenching to stop the enzymatic activity and stabilize the metabolites by treatment with cold methanol (Kamal et al. 2022). This prevents ongoing reactions and stops metabolic activity (Kamal et al. 2022). Centrifugation can then further separate metabolites from proteins and other debris, this is usually done for 8-10 minutes until a pellet is formed. The result is a pellet contains cell material and intracellular metabolites, whereas the supernatant contains extracellular material (including metabolites) (Kamal et al. 2022). Using a combination of chemical solvents and precipitates the metabolites can be separated from the other proteins and debris

(Sudeshna et al. 2024). Environmental samples often have much lower concentrations of metabolites than samples grown in the lab under optimal conditions (Sudeshna et al. 2024). Isolated metabolites can be separated by either liquid or gas chromatography and prepared by MALDI (ionizes the metabolites to identify mass-to-charge ratio) to be identified and detected by MS (Sudeshna et al. 2024). Once isolated into types of metabolites, the structure of each metabolite of interest can be examined by nuclear magnetic resonance (NMR) to provide structural information (Sudeshna et al. 2024).

Information obtained:

MS generates a mass spectra graph of the mass-to-charge ratio that can be compared to known libraries of metabolites to identify the type of metabolite present (Tamang, 2023). The graph contains peaks, with each peak corresponding to a specific mass-to-charge ratio (Tamang, 2023). If metabolite structure was needed, NMR results in spectra showing peaks at different chemical shifts for the different hydrogens on the molecules (Sudeshna et al. 2024). This allows the mapping of the chemical structure based on the different atoms present (Sudeshna et al. 2024).

Example:

In this study, the authors Kim et al. 2024, examined metabolites to look at the effects of different dietary treatments on pigs infected with enterotoxigenic *Escherichia coli*. The study integrated metabolomics findings with microbial diversity metrics to observe how metabolites correlate with microbial composition shifts (Kim et al. 2024). Samples from the gut were taken, and an untargeted metabolites approach was used, in which metabolites were isolated and profiled using mass-spectrometry (Kim et al. 2024). 16S rRNA gene sequencing was also done to assess microbial diversity and composition and compared to the diversity and composition of metabolites (Kim et al. 2024).

Advantages:

Metabolomics allows the detection of metabolites produced by organisms in the environment showing the microbial activity in the ecosystem and how they might change

in response to changes (Kim et al. 2024). Also determining what organisms are alive and dead (Kim et al. 2024). This shows what metabolic pathway is active in an ecosystem which is very useful in community ecology for understanding what resources are being consumed and produced (Bottos 2024).

Disadvantages:

The main challenge of metabolomics is their small size, low concentrations especially in environmental samples and chemical diversity, making isolation difficult (Kim et al. 2024). Monitoring metabolites in environmental samples is difficult unless the organism producing the metabolite is in high concentrations and making lots of metabolites (Kim et al. 2024).

6. Stable Isotope probing for tracking microbial metabolism through biomarker analysis.

Method:

Stable isotope probing (SIP) is a technique used to trace microbial activity by labelling nucleic acids (rRNA and rDNA) with stable isotopes, such as ^{13}C , ^{15}N , or ^{18}O , incorporated from substrate molecules (Kashyap et al. 2014). To label the microbial DNA or RNA, isotopes like ^{13}C are added glucose in environmental samples (e.g., soil, water, or biofilm) and allowed to incubate under controlled or natural conditions (Kashyap et al. 2014). Microorganisms that metabolize the labelled substrate incorporate the isotope into their cellular components, including DNA and RNA, where it can be detected (SD 2023). After incubation, DNA or RNA is extracted using standard molecular techniques, and isotope-labeled nucleic acids are separated from unlabeled ones using density gradient ultracentrifugation (Kashyap et al. 2014). In this process, the higher-density ^{13}C -labeled nucleic acids form a distinct layer in a CsCl gradient, allowing physical separation from unlabeled ^{12}C DNA, which has a lower density (Uhlik et al. 2009). Controls are run to identify the typical position of unlabeled ^{12}C DNA relative to G-C content (Uhlik et al. 2009). After centrifugation, ^{13}C -labeled DNA is recovered from the lower part of the centrifuge tube and can be subjected to PCR for

amplification, tagging, and sequencing to analyze functional genes of interest (Kashyap et al. 2014). Other ways to examine isotope probes for tracking microbial metabolism aside from 16S rRNA sequencing and RT-PCR are, IRMS (Isotope ratio mass spectrometry), and FISH (fluorescent in situ hybridization) (Kashyap et al. 2014).

Information obtained:

Sequencing of 16S rRNA genes from labelled nucleic acids gives a table of base pairs that allows for the taxonomic identification of active microbes in that sample (Uhlik et al. 2009). Isotope Ratio Mass Spectrometry (IRMS) provides quantitative data on isotope enrichment in the biomarker (Uhlik et al. 2009). Data are often shown as isotope ratios (e.g., $^{13}\text{C}/^{12}\text{C}$ or $^{15}\text{N}/^{14}\text{N}$) in the labelled nucleic acid fraction compared to natural background ratios (Uhlik et al. 2009). If fluorescent in situ hybridization (FISH) is used, fluorescent microscopy images provide spatial localization of active microbes (Kashyap et al. 2014). Images show labelled microbes within the sample matrix (e.g., biofilms or soil aggregates), often with fluorescent probes that highlight specific taxa (Kashyap et al. 2014).

Example:

In the paper by Ana et al. 2023, stable isotope probing was used to identify and track active microbial populations and their metabolic functions in their natural environment. The researchers used deuterium (^2H), a non-radioactive isotope of hydrogen, they enriched the water ($^2\text{H}_2\text{O}$) and added this to the soil (Ana et al. 2023). They were able to track which microbes incorporated ^2H into their lipids (Ana et al. 2023). Phospholipids are extracted from cells at different time points, isolated and identified and quantified using gas chromatography (Ana et al. 2023). The results showed that microbial growth rates were generally slow the alpine tundra soil showed the slowest growth rate due to its cold environment, with an average generation time of about 45 days, compared to faster growth in conifer forest and grassland soils (Ana et al. 2023).

Advantages:

Identifies microbes that are actively metabolizing a labelled substrate enabling the distinguishing between active and dormant microbial populations (Ana et al. 2023). This method involves minimal interference, allowing the examination of microbial metabolism in a natural environment (Ana et al. 2023).

Disadvantages:

Difficult in environments where metabolic activity is low/slow, this can make detection difficult as there is a slow turnover of biomolecules therefore low amounts of isotopes are being metabolized (Ana et al. 2023). Making this technique best for faster-growing organisms (Ana et al. 2023).

7. Reverse-transcriptase quantitative PCR for monitoring microbial gene expression.

Method:

Reverse transcription polymerase chain reaction (RT-PCR) is a technique that uses RNA as a substrate for the detection of DNA sequences (Preedy & Watson, 2004). RNA can be isolated and purified using specific buffers, columns, or kits designed to remove contaminants like DNA, proteins, and lipids (Preedy & Watson, 2004). Treatment with DNase can degrade any remaining DNA, ensuring only RNA is present (Preedy & Watson, 2004). Primers initiate reverse transcription, using short six-nucleotide hexamers, which may be random or sequence-specific (Preedy & Watson, 2004). The primer binds to the 3' end of the mRNA strand, allowing reverse transcriptase to synthesize a hybrid strand that is half cDNA and half mRNA (SD 2023). The mRNA portion is then cleaved by RNase H, resulting in 3' hydroxyl (OH) ends (SD 2023). DNA polymerase recognizes these 3' OH ends and synthesizes a complementary DNA strand, yielding a double-stranded cDNA molecule (SD 2023). Once the cDNA is synthesized, it can be amplified through PCR. Specific primers target the gene region of interest, while dNTPs and polymerase facilitate amplification via thermocycling (Preedy & Watson, 2004). The amplified cDNA, containing only the exon sequences (as introns have been spliced out in the mRNA), can then be analyzed via gel electrophoresis and sequenced to confirm the presence and sequence of the targeted gene (Preedy & Watson, 2004). What

makes this method quantitative is the actual amount of PCR product in a given sample is known (Sigma Aldrich, 2024). SYBR Green I dye is included in the reaction mixture, the dye binds all newly synthesized double-stranded DNA complexes and fluoresces them (Sigma Aldrich, 2024). After each cycle of PCR the intensity of the fluorescence generated is measured to quantify the amount of new DNA (Sigma Aldrich, 2024). This makes RT-qPCR a valuable tool for examining gene expression, as it quantifies the amount of DNA, which enables the detection of expression levels in genes without the interference of intronic regions (Preedy & Watson, 2004).

Information obtained:

RT-PCR is often used to quantify gene expression by measuring the abundance of specific mRNA transcripts (SD 2023). The raw data obtained contains nucleotide sequences, these can be expressed graphically to illustrate the concentration of genes expressed in that environment and confirm the presence or absence of the gene in that community (SD 2023).

Example:

In this article by Zeng et al. 2023, the researchers used RT-PCR to examine mRNA expression of inflammatory markers. This allowed them to track the gene expression changes in the gut microbiome in response to antibiotic stress and if dysbiosis was resolved by the addition of *Bacillus licheniformis* (potential therapeutic probiotic) (Zeng et al. 2023). RT-PCR was an essential tool for determining the genes currently expressed using mRNA sequences (Zeng et al. 2023). The results of this study found increases in pathogenic bacteria a decrease in pro-inflammatory and an increase in anti-inflammatory factors (Zeng et al. 2023).

Advantages:

RT-PCR provides quantitative data on gene expression by measuring transcribed mRNA and reverse transcribing it back into cDNA to determine what gene it is, what it expresses, and its location on the genome (Preedy & Watson, 2004).

Disadvantages:

RNA is susceptible to degradation, making collection and storage more difficult. To preserve RNA samples, need to be kept at low temperatures (-80°C) and reverse transcribed quickly (Preedy & Watson, 2004). This is due to the nature of the RNA molecule. RNA not required to stick around in the cell for long durations like DNA and proteins are (Preedy & Watson, 2004).

8. Fluorescence *in situ* hybridization (FISH) microscopy.**Method:**

Fluorescence in situ hybridization (FISH) microscopy is a powerful technique used to fluorescently label chromosomes and detect specific DNA sequences or genes within a cell (Moter, 2000). The process begins with preparing bacterial samples on a microscope slide and fixing the bacterial cells to prevent RNA degradation and enable probe penetration through the cell membrane (Moter, 2000). The DNA sample and the fluorescently labeled DNA probes, which are complementary to the target sequence, are then denatured by heat or chemicals to separate the DNA strands (O'Connor, 2008). This allows the probes, which are short oligonucleotides (15-30 base pairs) labeled with fluorescent dyes at both ends, to bind to their specific complementary sequences on the chromosome through hybridization (Moter, 2000). There are different types of probes used in FISH microscopy depending on the application: whole chromosome painting probes for visualizing entire chromosomes, repetitive sequence probes for identifying repeat-rich regions, and locus-specific probes for targeting specific gene sites (Shakoori et al., 2017). After denaturation, hybridization occurs by placing the samples in a controlled environment, such as a dark, humid chamber at 37°C to 50°C, for about 30 minutes to a few hours, allowing the probe to bind effectively to the target DNA sequence (Moter, 2000). Following hybridization, the samples are rinsed to remove any unbound probes (Dutra, 2023). Fluorescent signals on the chromosomes indicate the locations of the probes on the DNA sequences, and researchers can use this to construct gene maps (Dutra, 2023).

Information obtained:

Chromosomes appear as grey or blue structures in the image background, while the specific DNA sequences labeled with fluorescent probes appear as bright, colored spots or bands (Huang et al. 2023). The number of fluorescent spots can be counted to determine if specific sequences are present in normal copy numbers or if there are duplications or deletions (O'Connor, 2008). The position of the fluorescent spots within the chromosomal structure provides information on the specific location of the gene or DNA sequence (O'Connor, 2008). This can be useful for looking at spatial organization and interactions between microbes and host cells, FISH can show specific microbes clustered in certain areas like animal guts (O'Connor, 2008).

Example:

In this study, the authors Huang et al. 2023, used fluorescence in situ hybridization (FISH) to study the spatial distribution of symbionts, *Candidatus Sulcia muelleri* and yeast-like fungal symbionts (YLS), within the organs of two cicada species. Tissue samples were collected and fixed to slides, and fluorescently labelled oligonucleotide probes were used to target rRNA sequences in both species (Huang et al. 2023). After hybridization, tissues were visualized with a fluorescence microscope to observe the spatial distribution (Huang et al. 2023). Fluorescent signals enabled the mapping of *Sulcia* in bacteriomes and YLS in fat bodies, aiding in determining symbiont localization within host tissues (Huang et al. 2023). FISH allowed visualization of "symbiont balls" within the ovaries, showing that both symbionts are maternally transmitted to offspring (Huang et al. 2023).

Advantages:

FISH allows the visualization of physical locations on the chromosome within intact cells without DNA isolation and fragmenting (O'Connor 2008). When applied to a sample, the fluorescently labelled probes bind only to their target organisms (O'Connor 2008). This specificity enables the direct visualization of multiple microbial types in a single sample, each glowing in a unique colour under a fluorescence microscope

(O'Connor 2008). This is useful in identifying competition zones or symbiotic relationships between microbes (O'Connor 2008). If two or more microbes are found in the same space but not intermixed, this can indicate competition for resources, where each species is constrained to its territory (O'Connor 2008). Or microbes clustered to one host cell (O'Connor 2008).

Disadvantages:

The fluorescence signal intensity can vary based on probe accessibility, cell permeability, and target concentrations (O'Connor 2008). FISH provides spatial information and relative abundance; it doesn't offer precise quantification (O'Connor 2008). FISH is also disadvantaged in low abundances of sequences or genes this limitation is due to FISH being reliant on direct visualization, which can be hard to see in small quantities making it more useful in dense populations (Moter & Gobel 2000).

References:

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF. 2005. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Applied and Environmental Microbiology*. 71(12):8966–8969. doi: <https://doi.org/10.1128/AEM.71.12.8966-8969.2005>.

Ana Isabel Tristán, McFarlin J, Jech S, Fierer N, Kopf S. 2023. Hydrogen stable isotope probing of lipids demonstrates slow rates of microbial growth in soil. *Proceedings of the National Academy of Sciences of the United States of America*. 120(16). doi: <https://doi.org/10.1073/pnas.2211625120>

Bielen A, Babić I, Marija Vuk Surjan, Snježana Kazazić, Šimatović A, Jasna Lajtner, Nikolina Udiković-Kolić, Zrinka Mesić, Hudina S. 2024. Comparison of MALDI-TOF mass spectrometry and 16S rDNA sequencing for identification of environmental bacteria: a case study of cave mussel-associated culturable microorganisms. *Environmental Science and*

Pollution Research International. 31(14):21752–21764. doi: <https://doi.org/10.1007/s11356-024-32537-1>

Bottos E. 2024. Microbial Ecology. Thompson Rivers University.

Preedy Victor, and Watson Ross. Comprehensive Handbook of Alcohol Related Pathology. 2004. ScienceDirect. <https://www.sciencedirect.com/book/9780125643702/comprehensive-handbook-of-alcohol-related-pathology>

Doerr SH, Wei F, Jansson C. Metaproteomics. ScienceDirect. [accessed 2023 Nov. 4]. [https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/metaproteomics#:~:text=Metaproteomics%20\(also%20community%20proteomics%2C%20environmental](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/metaproteomics#:~:text=Metaproteomics%20(also%20community%20proteomics%2C%20environmental)

Huang Z, Zhou J, Zhang Z, He H, Wang C. 2023. A Study on Symbiotic Systems of Cicadas Provides New Insights into Distribution of Microbial Symbionts and Improves Fluorescence In Situ Hybridization Technique. International Journal of Molecular Sciences. 24(3):2434–2434. doi: <https://doi.org/10.3390/ijms24032434>

Introduction to mass spectrometry data analysis. 2023. Aspect Analytics. <https://www.aspect-analytics.com/media-blog/introduction-to-mass-spectrometry-data-analysis>

Kamal KM, Maifiah MHM, Rahin NA, et al. 2022. Bacterial Metabolomics: Sample Preparation Methods. NIH. doi: 10.1155/2022/9186536. <https://pubmed.ncbi.nlm.nih.gov/35465444/>

Karpievitch Y, Polpitiya A, Anderson G, Smith RD, Dabney AR. 2010. Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological

Aspects. NIH. 4(4):1797-1823. doi: 10.1214/10-AOAS341.

<https://pubmed.ncbi.nlm.nih.gov/21593992/>

Kashyap SK, Maherchandani S, Kumar N. 2014 Jan 1. Chapter 18 - Ribotyping: A Tool for Molecular Taxonomy. Verma AS, Singh A, editors. ScienceDirect. 327–

344. <https://www.sciencedirect.com/science/article/abs/pii/B9780124160026000183>

Kim K, Jinno C, Li X, Bravo D, Cox E, Ji P, Liu Y. 2024. Impact of an oligosaccharide-based polymer on the metabolic profiles and microbial ecology of weanling pigs experimentally infected with a pathogenic *E. coli*. Journal of Animal Science and Biotechnology. 15(1). doi: <https://doi.org/10.1186/s40104-023-00956-8>

Lan W, Liu H, Weng R, Zeng Y, Lou J, Xu H, Yu Y, Jiang Y. 2024. Microbial community of municipal drinking water in Hangzhou using metagenomic sequencing. Environmental Pollution. 342:123066–123066. doi: <https://doi.org/10.1016/j.envpol.2023.123066>

Lee R. 2019. Metagenomic Next Generation Sequencing: How Does It Work and Is It Coming to Your Clinical Microbiology Lab? ASM.

[https://asm.org/articles/2019/november/metagenomic-next-generation-sequencing-how-does-it#:~:text=Metagenomic%20NGS%20\(mNGS\)%20is%20simply.present%20and%20in%20what%20proportions](https://asm.org/articles/2019/november/metagenomic-next-generation-sequencing-how-does-it#:~:text=Metagenomic%20NGS%20(mNGS)%20is%20simply.present%20and%20in%20what%20proportions)

Lowe R, Shirley N, Bleakley M, Dolan S, Shafee T. 2017. Transcriptomics technologies. NIH. 13(5). doi: 10.1371/journal.pcbi.1005457.

<https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005457>

Metagenomics - Operational taxonomic unit (OTU). 2024.

Metagenomicswiki. <https://www.metagenomics.wiki/pdf/taxonomy/16s/operational-taxonomic-unit-otu#>

Ma C, Lo KP, Xu J, et al. 2020. Molecular mechanisms underlying lignocellulose degradation and antibiotics resistant genes removal via metagenomics analysis during different agricultural wastes composting. ScienceDirect.

314. <https://doi.org/10.1016/j.biortech.2020.123731>

Moter A, Gobel UB. 2000. Fluorescence in situ hybridization (FISH) for direct visualization of microorganisms. ScienceDirect. 41(2):85-112. doi: [https://doi.org/10.1016/S0167-7012\(00\)00152-4](https://doi.org/10.1016/S0167-7012(00)00152-4)

Novak JK, Kennedy PG, Gardner JG. 2024. Transcriptomic analyses of bacterial growth on fungal necromass reveal different microbial community niches during degradation. Nickel PI, editor. Applied and Environmental Microbiology. 90(10).

doi: <https://doi.org/10.1128/aem.01062-24>

O'Connor C. 2008. Fluorescence In Situ Hybridization (FISH) | Learn Science at Scitable. Naturecom. <https://www.nature.com/scitable/topicpage/fluorescence-in-situ-hybridization-fish-327/>

Pascoal F, Duarte P, Assmy P, Costa R, Magalhães C. 2024. Full-length 16S rRNA gene sequencing combined with adequate database selection improves the description of Arctic marine prokaryotic communities. Annals of Microbiology. 74(1).

doi: <https://doi.org/10.1186/s13213-024-01767-6>

Principles and Workflow of 16S/18S/ITS Amplicon Sequencing - CD Genomics. 2020. Cd-genomicscom. <https://www.cd-genomics.com/resource-principels-and-workflow-of-16s-18s-its-amplicon-sequencing.html#:~:text=The%20principal%20steps%20of%2016S>

Setubal JC. 2021. Metagenome-assembled genomes: concepts, analogies, and challenges. *Biophysical Reviews*. 13(6):905–909. doi: <https://doi.org/10.1007/s12551-021-00865-y>

Shakoori AR, et al. 2017. Fluorescence In Situ Hybridization (FISH) and Its Applications. *NIH*. 10:343-367. doi: 10.1007/978-81-322-3673-3_16.
https://link.springer.com/chapter/10.1007/978-81-322-3673-3_16

Sigma Aldrich. 2024. Blue-White Screening & Protocols for Colony Selection. Merck. 1(1).
<https://www.sigmaaldrich.com/MX/en/technical-documents/technical-article/genomics/cloning-and-expression/blue-white-screening>

Sudeshna Bhattacharjya, Ghosh A, Sahu A, Agnihotri R, Pal N, Sharma P, Manna MC, Sharma MP, Singh AB. 2024. Utilizing soil metabolomics to investigate the untapped metabolic potential of soil microbial communities and their role in driving soil ecosystem processes: A review. *Applied Soil Ecology*. 195:105238–105238.
doi: <https://doi.org/10.1016/j.apsoil.2023.105238>

Tamang S. 2023. Proteomics: Types, Methods, Steps, Applications. *microbenotes.com*. <https://microbenotes.com/proteomics/>

Uhlik O, Jecna K, Leigh MB, Mackova Martinaand Macek Tomas. 2009. DNA-based stable isotope probing: a link between community structure and function. *ScienceDirect*. 407(4).
<https://doi.org/10.1016/j.scitotenv.2008.05.012>

