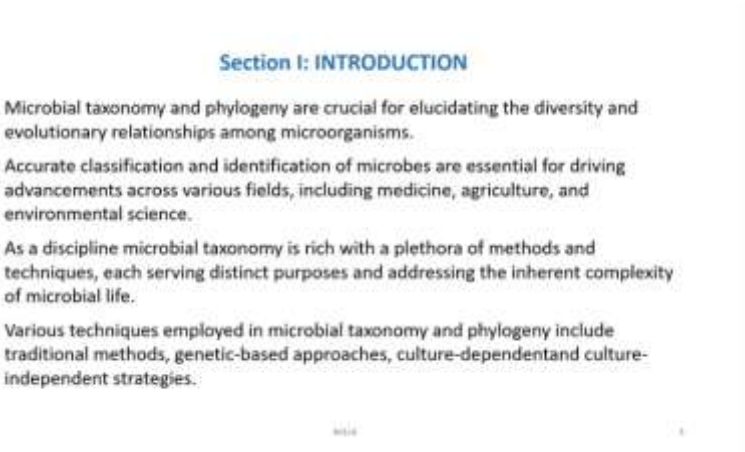
MICROBIAL BIOTECHNOLOGY

Prof. Utpal Bora
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati

Lecture04
Lec 4: Techniques for microbial Classification and Identification

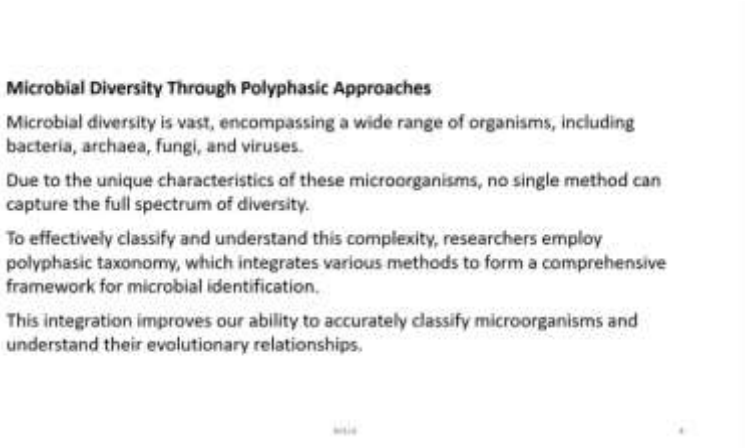
Welcome to my course on microbial biotechnology. We are in module 1, where we are discussing the principles of microbial biotechnology, classification and taxonomy of microbes. So, in continuation of the last lecture, today we are going to discuss about techniques for microbial classification and identification. Briefly, it has four sections, starting with introduction, then we go on to discuss traditional methods of classification, then genotyping methods of classification, and identification methods for bacteria and fungi. So, let us start with a small introduction to taxonomy and phylogeny and why they are crucial for elucidating the diversity and evolutionary relationship among microorganisms.



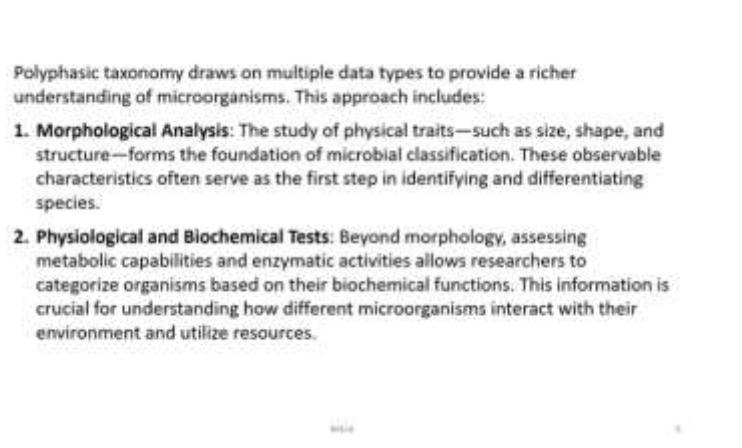
Accurate classification and identification of microbes are essential for driving advancements across various fields like medicine, agriculture and environmental sciences. As a discipline, microbial taxonomy is rich with a plethora of methods and techniques, each method or technique serving distinct purposes in addressing the inherent complexity of microbial life. Various techniques employed in microbial taxonomy and phylogeny include traditional methods, genetic based approaches, culture dependent and culture independent strategies. Microbial diversity through polyphysic approaches is something very very important as microbial diversity is vast and encompass a wide range of organisms

as you already know like bacteria, archaea, fungi and viruses and due to the unique characteristics of each and every one of these microorganisms, no single method can capture the full spectrum of biodiversity we have in the microbial world.

To effectively classify and understand this complexity, researchers employ polyphasic taxonomy which integrates various methods to form a comprehensive framework for microbial identification. This integration improves our ability to accurately classify microorganisms and understand their evolutionary relationship. Polyphysic taxonomy draws on multiple data types to provide a richer understanding of microorganisms. This approach includes number one, morphological analysis. This involves the study of physical traits such as size, shape and structure.



Some of these are already known to you and form the foundation of microbial classification. These observable characteristics often serve as the first step in identifying and differentiating species. The second important approach is the physiological and biochemical test, which goes beyond morphology and cannot be resolved by it alone. It assesses metabolic capabilities and enzymatic activities, allowing researchers to categorize organisms based on their biochemical functions. This information is crucial for understanding how different microorganisms interact with the environment and utilize resources.

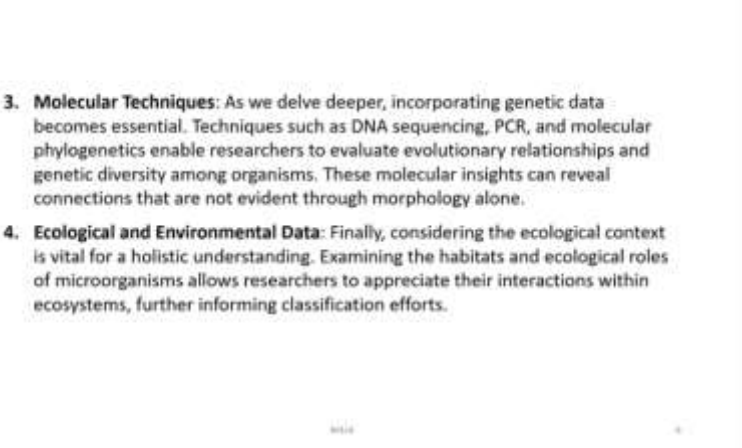


Polyphasic taxonomy draws on multiple data types to provide a richer understanding of microorganisms. This approach includes:

1. **Morphological Analysis**: The study of physical traits—such as size, shape, and structure—forms the foundation of microbial classification. These observable characteristics often serve as the first step in identifying and differentiating species.
2. **Physiological and Biochemical Tests**: Beyond morphology, assessing metabolic capabilities and enzymatic activities allows researchers to categorize organisms based on their biochemical functions. This information is crucial for understanding how different microorganisms interact with their environment and utilize resources.

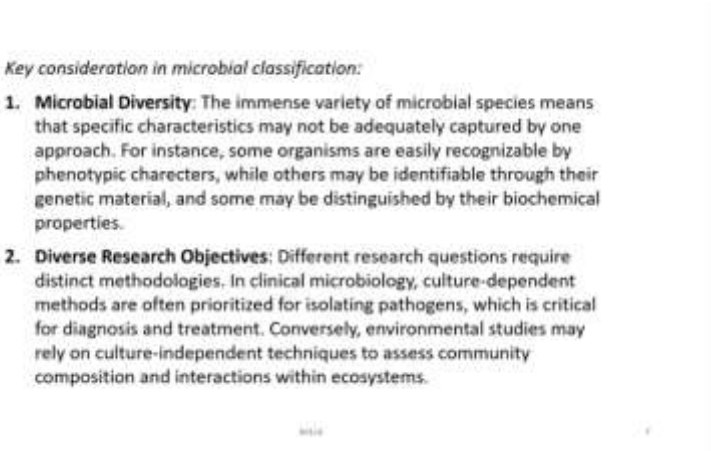
The third important approach is molecular techniques. As we delve deeper, incorporating genetic data becomes essential. Techniques such as DNA sequencing, PCR, and molecular phylogenetics enable researchers to evaluate evolutionary relationships and genetic diversity among organisms. These molecular insights can reveal connections that are not evident through morphology alone. The fourth important approach is ecological and environmental data.

These consider the ecological context, which is vital for a holistic understanding. Examining the habitats and ecological roles of microorganisms allows researchers to appreciate their interactions within ecosystems, further informing classification efforts. Now let us look into some of the key considerations in microbial classification. Why so many different kinds of methods are required has already been discussed in the earlier slides. That's mostly due to the vastness of the diversity we have and the uniqueness of each and every kind of microbe in this world.



3. **Molecular Techniques**: As we delve deeper, incorporating genetic data becomes essential. Techniques such as DNA sequencing, PCR, and molecular phylogenetics enable researchers to evaluate evolutionary relationships and genetic diversity among organisms. These molecular insights can reveal connections that are not evident through morphology alone.
4. **Ecological and Environmental Data**: Finally, considering the ecological context is vital for a holistic understanding. Examining the habitats and ecological roles of microorganisms allows researchers to appreciate their interactions within ecosystems, further informing classification efforts.

So, this microbial diversity is actually one of the key considerations. The immense variety of microbial species means that specific characteristics may not be adequately captured by only one of these approaches. For instance, some organisms are easily recognizable by phenotypic characteristics, while others may be identifiable through their genetic material, and some may be distinguished by their biochemical properties. Again, another important consideration is the diverse research objectives each researcher or explorer may have.



Key consideration in microbial classification:

1. **Microbial Diversity**: The immense variety of microbial species means that specific characteristics may not be adequately captured by one approach. For instance, some organisms are easily recognizable by phenotypic charecters, while others may be identifiable through their genetic material, and some may be distinguished by their biochemical properties.
2. **Diverse Research Objectives**: Different research questions require distinct methodologies. In clinical microbiology, culture-dependent methods are often prioritized for isolating pathogens, which is critical for diagnosis and treatment. Conversely, environmental studies may rely on culture-independent techniques to assess community composition and interactions within ecosystems.

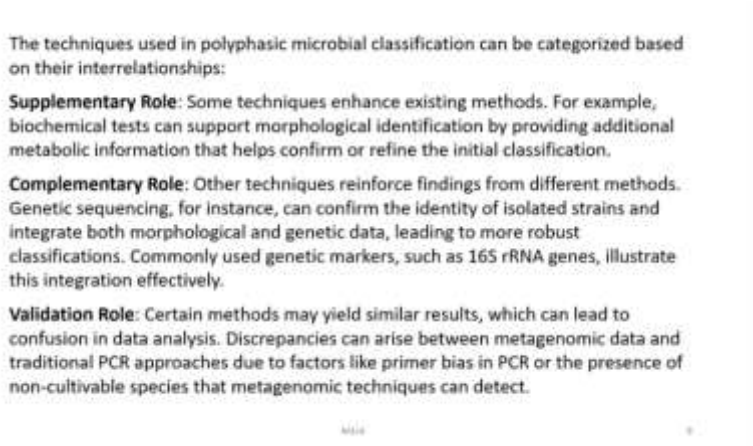
Different research questions require different methodologies. In clinical microbiology, culture-dependent methods are often prioritized for isolating pathogens, which is critical for diagnosis and treatment. Conversely, environmental studies may rely on culture-independent techniques to assess community composition and interactions within ecosystems. The third important key point is the limitations of individual techniques, which we have just discussed briefly. Each method has inherent limitations that can hinder our understanding.

Traditional culture techniques may miss unculturable organisms, which in fact they do. And unculturable organisms are far more numerous than the culturable ones, as reported as of now. They represent a significant portion of microbial diversity. On the other hand, genetic methods may not fully reveal the functional traits or interactions within microbial communities, leaving gaps in our knowledge. The third important point is technological advancements.

3. **Limitations of Individual Techniques**: Each method has inherent limitations that can hinder our understanding. Traditional culture techniques may miss unculturable organisms, which represent a significant portion of microbial diversity. On the other hand, genetic methods may not fully reveal the functional traits or interactions within microbial communities, leaving gaps in our knowledge.
4. **Technological Advancements**: Continuous innovations in technology have paved the way for new methods, particularly genetic and high-throughput techniques. Approaches like next-generation sequencing (NGS) have revolutionized microbial analysis, offering enhanced sensitivity, resolution, and the ability to explore complex microbial communities in unprecedented detail.

Continuous innovations in technology have paved the way for new methods. particularly genetic and high throughput techniques approaches like next-generation sequencing have revolutionized microbial analysis offering enhanced sensitivity resolution and the ability to explore microbial communities in unprecedented detail The techniques used in polyphysic microbial classification can be categorized based on their interrelationships. So many different methods, but how they are related to one another or how they are not related to one another, what is common and what is unique amongst them is important to understand. So, these techniques or methods sometimes are supplementary to one another, so it plays a supplementary role.

Some techniques would enhance the existing methods. For example, biochemical tests can support morphological identification by providing additional metabolic information that helps confirm or refine the initial classification. So it allows us in improving the microbial taxonomy. Sometimes, these techniques, methods play complementary role. Other techniques reinforce findings from different methods.



The techniques used in polyphasic microbial classification can be categorized based on their interrelationships:

Supplementary Role: Some techniques enhance existing methods. For example, biochemical tests can support morphological identification by providing additional metabolic information that helps confirm or refine the initial classification.

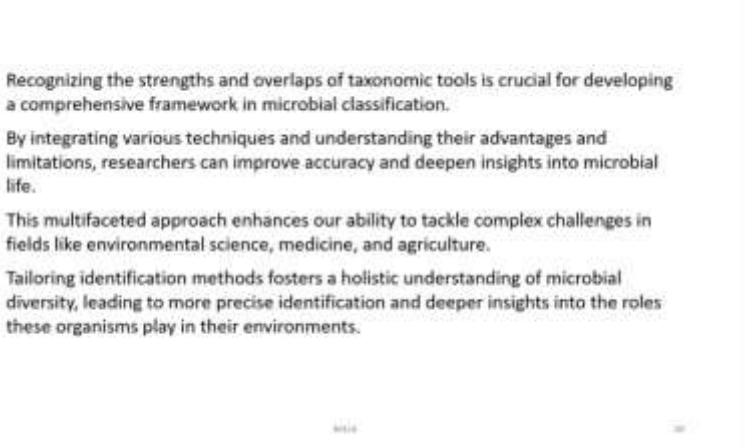
Complementary Role: Other techniques reinforce findings from different methods. Genetic sequencing, for instance, can confirm the identity of isolated strains and integrate both morphological and genetic data, leading to more robust classifications. Commonly used genetic markers, such as 16S rRNA genes, illustrate this integration effectively.

Validation Role: Certain methods may yield similar results, which can lead to confusion in data analysis. Discrepancies can arise between metagenomic data and traditional PCR approaches due to factors like primer bias in PCR or the presence of non-cultivable species that metagenomic techniques can detect.

Genetic sequencing, for instance, can confirm the identity of isolated strains and integrate both morphological and genetic data, leading to more robust classifications. Commonly used genetic markers such as 16S rRNA genes illustrate this integration effectively. We will be discussing some of these methods in detail later. Many times these alternative methods play the role of validation. So, certain methods may yield similar results which can lead to confusion in data analysis.

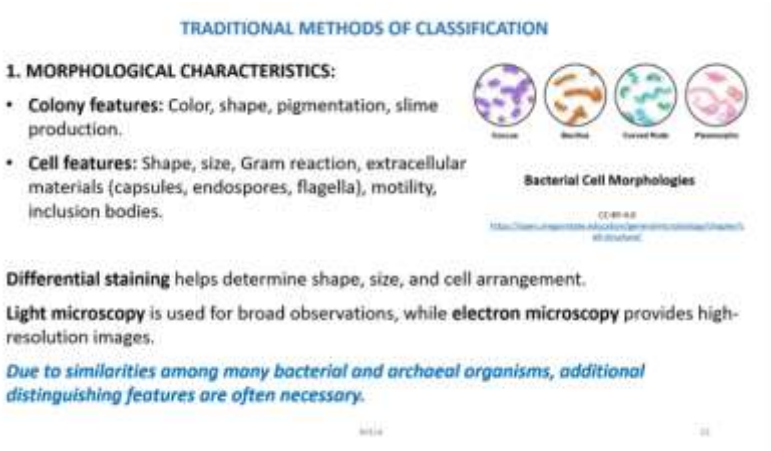
Discrepancies can arise between metagenomic data and traditional PCR approaches due to factors like primer bias in PCR or the presence of non-cultivable species that metagenomic techniques can detect. So, recognizing the strengths and overlaps of taxonomic tools and methods is crucial for developing a comprehensive framework in microbial classification. By integrating various techniques and understanding their advantages and limitations, researchers can improve accuracy and deepen insights into microbial life. This multifaceted approach enhances our ability to tackle complex challenges in fields like environmental science, medicine, and agriculture.

Tailoring identification methods fosters a holistic understanding of microbial diversity, leading to more precise identification and deeper insights into the roles these organisms play in their environments. Let us now move on to section 2, which deals with the traditional methods of classification. So, we begin with the morphological characters; some of these have already been discussed in earlier lectures. Still, we will just mention them here to show how they are useful as classification tools. So, colony features—mostly the color, shape, pigmentation, or slime production—are important morphological characteristics.

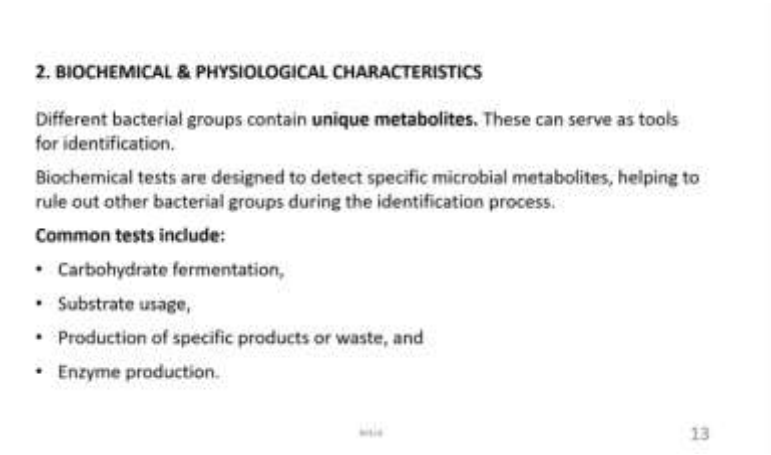


Then, cell features like shape, size, and Gram reaction; extracellular materials like capsules; and those for flagella, motility, and inclusion bodies—these are already explained

in detail in earlier lectures along with examples. Differential staining helps determine the shape, size, and cell arrangement, and these are some of the different kinds of clusters based on the cell morphologies, as you can see over here. Light microscopy is used for broad observations, while electron microscopy provides high-resolution images. Due to similarities among many bacterial and archaeal organisms, additional distinguishing features are often necessary for resolution. Biochemical and physiological characteristics.

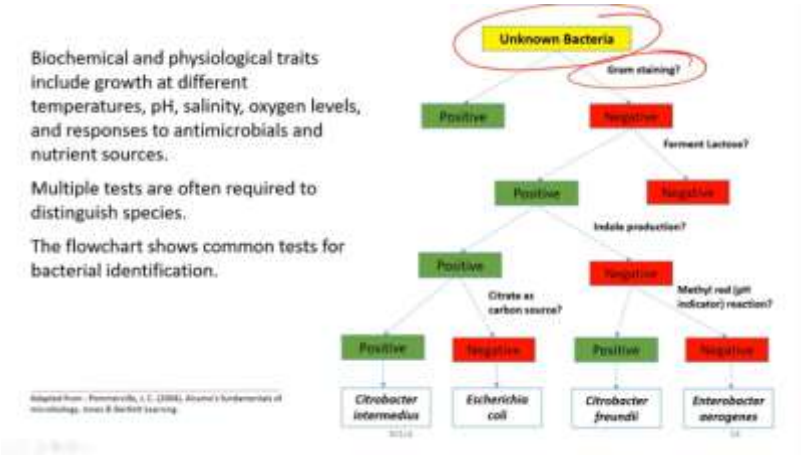


Different bacterial groups contain unique metabolites. These unique metabolites can serve as tools for identification and also classification. Biochemical tests are designed to detect specific microbial metabolites helping to rule out other bacterial groups during the identification process. Common tests are carbohydrate fermentation, substrate uses, production of specific products or wastes and enzyme production to name a few. So here we can see one example of an unknown bacteria which is given to us as a sample.



So, we go for gram staining. So, if it is positive, we classify it as gram positive. If it is negative, we try to find out does it for example ferment lactose. If the answer is yes, we

will go for further characterizations. If the answer is negative, we will stop there and take some kind of a decision.



If the answer is yes, it can ferment lactose, we may go for testing another metabolite. Does it produce indole? If the answer is no, then we will try to see whether it gives any kind of a methyl rate pH indication reaction and if it is positive, we can tell that it is a cytobacter or if it is negative we may tell that it is enterobacter. So this is just an example and there are many such flowcharts where you can see many other such case studies.

And then if in case the lactose indole production is positive and we can look if the citrate is being used as a carbon source and finally if the result is positive we can tell this is citrobacter intermedius and if it is negative we may tell that this is Escherichia coli. So, biochemical and physiological tests include growth at different temperatures, pH, salinity, oxygen levels and responses to antimicrobials and nutrient sources. For example, here citrate is taken as a carbon source. And then here we are looking into lactose fermentation. Multiple tests are often required to distinguish the species as you can see over here in this diagram.

The flowchart shows these common tests for bacterial identification which has already been discussed. Another way of identification and classification is using entero tube which is a kind of available commercially. This kind of entero tube too is a sterile tube with 12 media compartments and an inoculating wire. And overall it allows 15 biochemical tests for a single bacterial colony. For example, here we have to go for doing the biochemical test separately.

Enterotube II

The Enterotube II is a sterile tube with 12 media compartments and an inoculating wire.

It enables 15 biochemical tests for a single bacterial colony:

Glucose	Lactose	Arabinose	Sorbitol	Ornithine Decarboxylase
Citrate	Adonitol	Indole	Phenylalanine Deaminase	Lysine Decarboxylase
Urea	H2s	Dulcitol	Gas Production	Voges-proskauer

Media change color based on metabolic reactions, and after 24 hours, positive results are marked. The circled numbers form a unique ID, which is used to identify the bacterium via a reference book or software.

But here we can do the entire panel of these 15 biochemical tests in one single colony. For example, this provides us with the capability to test for glucose, lactose, arabinose, sorbitol, and then also enzymes like ornithine decarboxylase, phenylalanine deaminase, lysine decarboxylase. We can even look into the production of urea, then other gas productions, and then also dulcetol, etc. So, in simple terms you can see here the various media compartments and there are certain colour coding I mean provided in the leaflet. So, we can always match these tubes with this colour coding leaflets over there.

The media would change color based on metabolic reactions and after 24 hours positive results are marked. The circle numbers form a unique ID which is used to identify the bacterium by a reference book or software. The third is the cultural characteristics, microbial growth on solid media like nutrient agar displays distinct characteristics which can be used for identification. Some of the key features include the colony diameter, shape, elevation, transparency, whether it is clear, translational, opaque and the color. Changes in the medium, such as hemolysis on blood agar.

3. CULTURAL CHARACTERISTICS

Microbial growth on solid media, like nutrient agar, displays distinct characteristics useful for identification.

Key features include:

- Colony diameter, shape, elevation, translucency (clear, translucent, opaque), and color.
- Changes in the medium, such as haemolysis on blood agar.

Unique growth patterns reflect the organism's preferred conditions.

Its ability to grow with or without oxygen (aerobe, anaerobe),

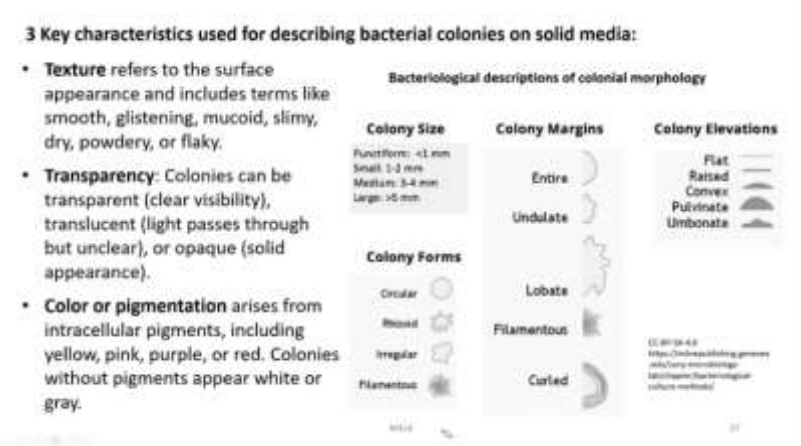
in reduced oxygen (microaerophile), or

in the presence of carbon dioxide, as well as

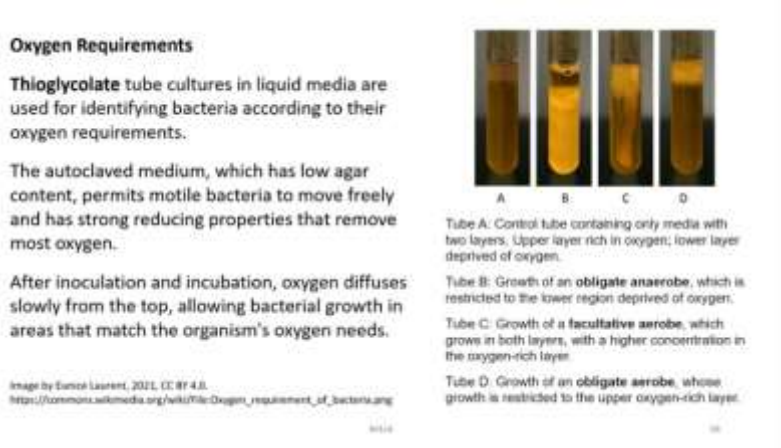
responses to selective inhibitors (e.g., bile salts, antimicrobial agents, pH), aid in identification.

Unit growth patterns reflect the organism's preferred conditions. Its ability to grow wheat or without oxygen, whether it is aerobic or anaerobic. In reduced oxygen, microaerophile or in the presence of carbon dioxide, as well as responses to selective inhibitors like bile salts, antimicrobial agents, pH, and aid in the identification process. So let us see some of the vectorological descriptions of the various morphologies of a colony briefly. So you have this form which is circular, rhizoid, irregular and filamentous.

Then you have colony margins which can be entire or which can be undulated or which can be lobbed and filamentous and curled then the elevations that is height from the media base it could be flat or it could be raised it could be convex or it can be pulvinate or it could be ammonate. Now let us study the oxygen requirements by a microbial culture. Here, thioglycolate tube cultures in liquid media are used for identifying bacteria according to their oxygen requirements. The medium is autoclaved, has low agar content.

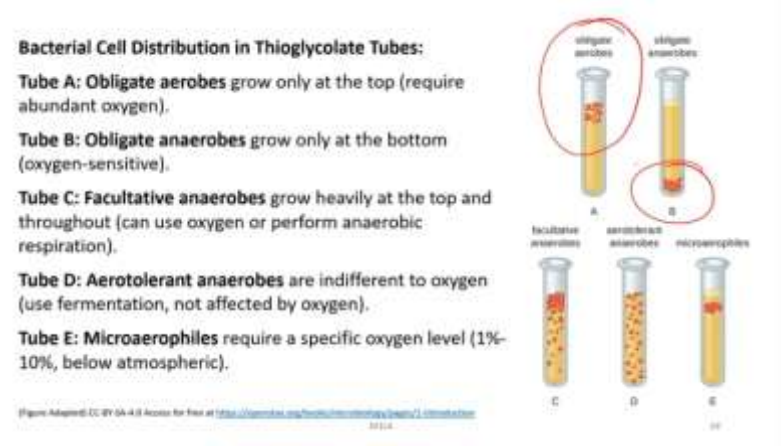


It permits motile bacteria to move freely and has strong reducing properties that remove most oxygen. After inoculation and incubation, oxygen diffuses slowly from the top. Allowing bacterial growth in areas that match the organism's oxygen needs. So here we have different tubes: A, B, C, D, and they have different oxygen conditions. Number A is a control tube which contains only media with two layers.



The upper layer is rich in oxygen, as you can see by the difference in transparency. And the lower layer is deprived of oxygen. Tube B shows the growth of an obligate anaerobe, which is restricted to the lower region deprived of oxygen. Tube C shows the growth of a facultative aerobe, which grows in both layers with a higher concentration in the oxygen-rich layer. Tube D shows the growth of an obligate aerobe, whose growth is restricted to the upper oxygen layer.

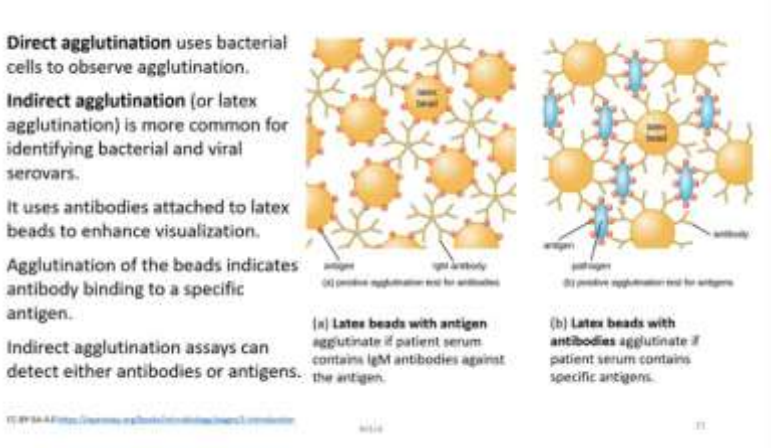
We will see this in simpler diagrams. As you can see here, tube A, the first tube, shows the obligate aerobe on the upper part, then the obligate anaerobe. On the lower part, you have facultative anaerobes, and number D shows aerotolerant anaerobes, while number E shows the microaerophiles. These microaerophiles require oxygen levels around 1% to 10%, and they are concentrated slightly below the top layer, as you can see here. Another important method or tool is serology.



Serology basically studies blood serum which contains antibodies that identify and distinguish microorganisms and strains. Using serology, we can diagnose infections like

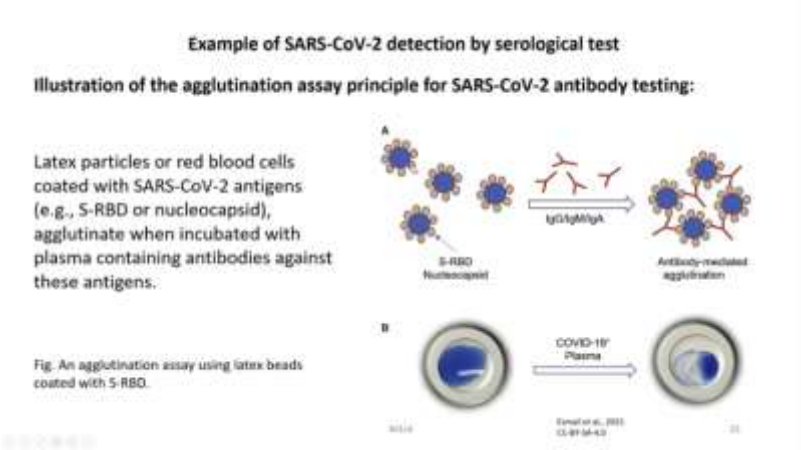
HIV, hepatitis, etc., We can determine immune status, I mean the vaccine history of a particular patient. Then blood tapping for transfusions and transplants and diagnosing autoimmune diseases like lupus. Then also testing for allergy like IgE, hypersensitivity.

And the common methods that are adopted for serology include ELISA, which is enzyme-linked immunosorbent assay, then agglutination assays and immunofluorescence. Here an example is given, Salmonella enticidum can cause Salmonella cells to clump and helps identify the pathogen in foodborne illnesses. So, if we see what is direct agglutination, it uses bacterial cells to observe agglutination or coming together. Then direct agglutination or latex agglutination is more common for identifying bacterial and viral cerebrars. It uses antibodies attached to latex beads to enhance visualization.



Agglutination of the beads indicates antibody binding to a specific antigen and indirect agglutination assays can detect either antibodies or antigen. There are two cases, the first one is a latex bead with antigen and second one is latex beads with antibodies and you can see this is the antigen and these bigger structure is the latex bead and here these are the Y-shaped antibodies as you all know and this big one is the latex bead you can see here already labeled. Now whether it is antigen or antibody in the first case these latex beads will bind to one another if patient serum contains IgM antibodies. So, these are the IgM antibodies.

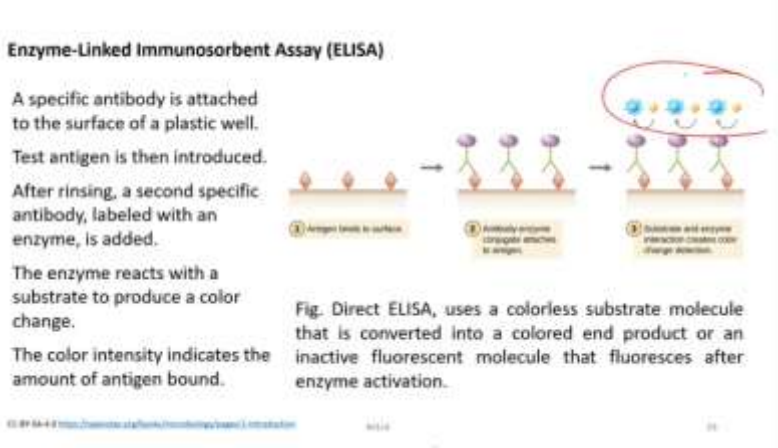
And these IgM antibodies will bind to different antigens on different latex particles and thereby agglutinate them. And in the case of latex beads with antibodies, you can see here that they will agglutinate if the patient serum contains specific antigens. So, in the past two years, since 2020, the world has been suffering from the SARS-CoV-2 pandemic. And here, the serological test was used for detection. And you can see the agglutination assay principle for SARS-CoV-2 antibody testing here.



So, you have latex particles or red blood cells coated with SARS-CoV-2 antigens. They agglutinate when incubated with plasma taken from the patient, which contains antibodies against these antigens. Next is the enzyme-linked immunosorbent assay. So, you can see here in the first step, we bind the antigens to the surface of polystyrene wells. So, a specific antibody is attached to the surface of the plastic well.

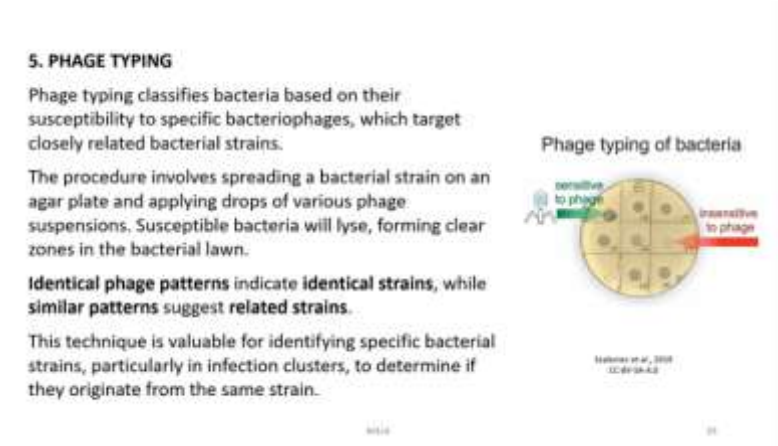
That is also another mode. We can either bind the antigen or we can bind the antibody. If the antibody is bound first to the well, we will introduce a test antigen. And if we bind antigens to the well, we will introduce antibodies or the sera. Anyway, in between there will be a blocking step to block these empty spaces with molecules such as bovine serum albumin so that there are no false-negative results from this analysis.

After rinsing a second specific antibody labeled with an enzyme here, the enzyme reacts with a substrate to produce a color change. As you can see over here, the color intensity indicates the amount of antigen that is bound. So this is direct analyzer. This uses a colorless substrate molecule that is converted into a colored end product or an inactive fluorescent molecule that fluoresces after the enzyme activation. In reverse ELISA purified antigen not specific antibody is attached to a well and test serum is added.



An enzyme linked system detects the antigen antibody reaction often utilizing labeled anti-human globulin for identification. Imlogrobilin M-antibody capsa-aliza or MAC-aliza is used to detect IgM in virological diagnosis. An anti-human mu-chain antibody typically derived from goat is bound to the well allowing IgM from the test serum to attach. After washing a purified antigen such as rubella is added and detection is performed using a suitable antibody. Another method is phase typing.

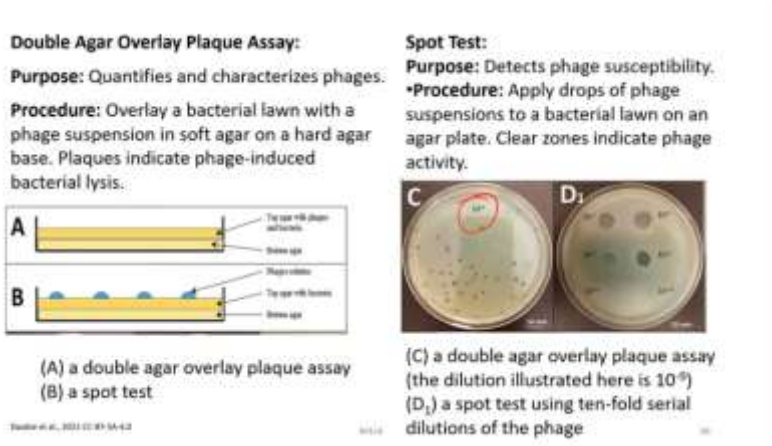
Phase typing classifies bacteria based on their susceptibility to specific bacteriophages which target closely related bacterial strains. The procedure involves spreading a bacterial strain on an agar plate and applying drops of various phage substances Susceptible bacteria will lies forming clear zones in the bacterial lawn as you can see here which is sensitive to the phage and here we see that they are not sensitive to the phage or insensitive to the phage. Identical phase patterns indicate identical strains while similar patterns suggest related strains. This technique is valuable for identifying specific bacterial strains particularly in infection clusters to determine if they originate from the same strain.



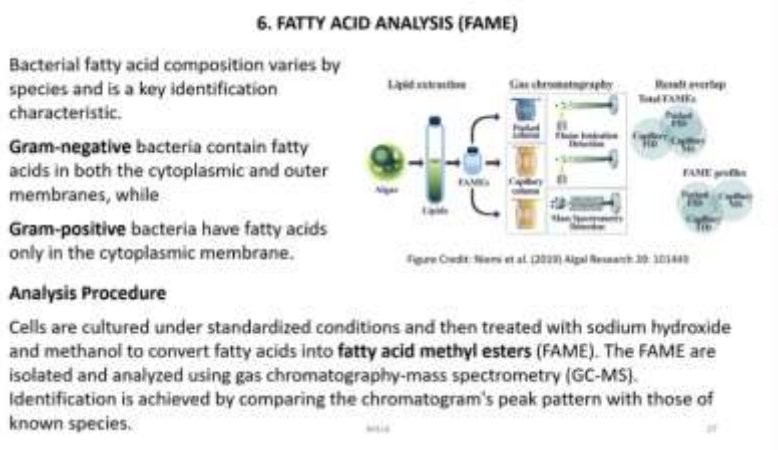
Next, let us discuss about the double agar overlay plague test. So, the purpose of these particular tests is to quantify and characterize phages. The procedure includes overlaying a bacterial lawn. with a fast suspension in soft agar on a hard agar base. Plex indicate fast induced bacterial lysis.

So, in A you can see a double agar overlay plaque assay. In B you can see a spot test. So, in spot test We applied drops of fuzz suspensions to a bacterial lawn on an agar plate. Clear zones indicate fuzz activity.

So here you can see this double test, this spot test in B. And here you see in figure C a double agar overlay plaque assay. The dilution illustrated here for example is 10 to the power minus 9. D1 shows the spot test using 10-fold serial dilution of the fuzz. So, we keep on diluting it and then we carry out the test.



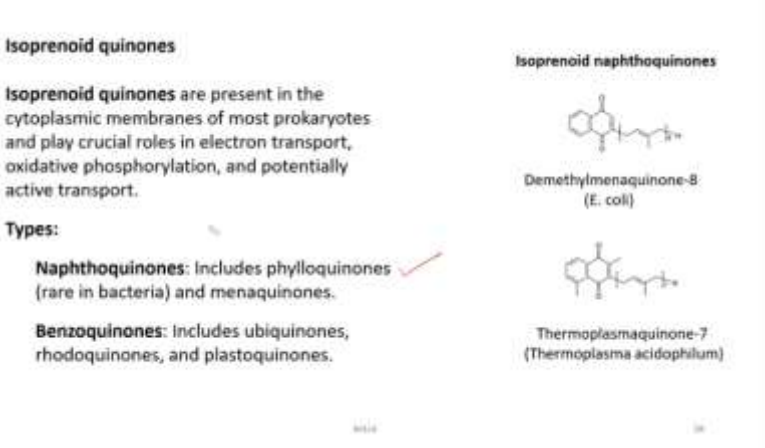
The next most important technique is the fatty acid analysis or FAME. We call this the FAME test because of the fatty acid methyl esters. So let us see what this technique is all about. Bacterial fatty acid composition varies by species and is a key identification characteristic. Gram-negative bacteria contain fatty acids in both the cytoplasmic and outer membranes, while gram-positive bacteria have fatty acids only in the cytoplasmic membrane.



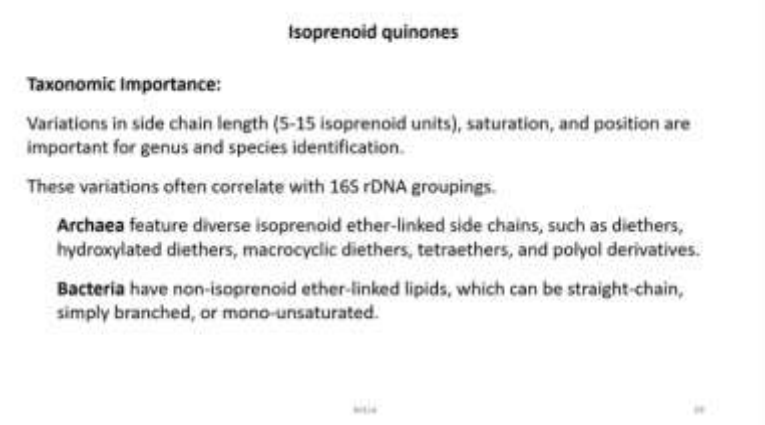
Analysis procedure. Cells are cultured under standardized conditions and then treated with sodium hydroxide and methanol to convert fatty acids into fatty acid methyl esters or FAME. These FAME are isolated and analyzed using gas chromatography, as you can see in this picture, or gas chromatography mass spectrometry. Identification is achieved by comparing the chromatogram's peak pattern with those of known species.

Let us now discuss the importance of Isoprenoid quinones for identification and classification. These Isoprenoid quinones are present in the cytoplasmic membrane of most prokaryotes and therefore play crucial roles in electron transport, oxidative phosphorylation, and potentially active transport. We may have different kinds of Isoprenoid quinones. For example, naphthoquinones include phylloquinones. It is rare in bacteria and menaquinones.

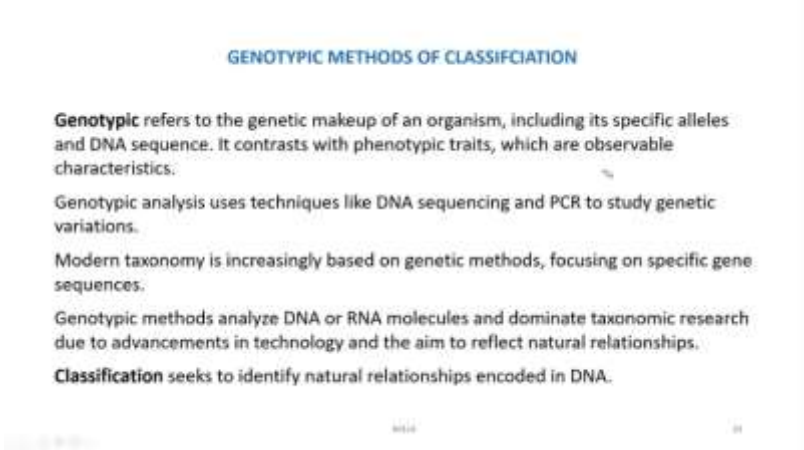
Then we have benzoquinones, which include EB quinones, rhodoquinones, and plastoquinones. So here you can see some of the isoprenoid naphthoquinones in this figure on the top and bottom. So, these serve as tools for taxonomy or classification. So, variations in the side chain length, as you can see here, have side chain lengths which may range from around 5 to 15 isoprenoid units. Then saturation and position are important for genus and species identification.



These variations often correlate with 16S-RDNA groupings as per research reports. Archaea feature diverse isoprenoid ether-linked side chains, such as diethers, hydroxylated diethers, macrocyclic diethers, tetraethers, and polyol derivatives. Bacteria have non-isoprenoid ether-linked lipids, which can be straight-chain, simply branched, or mono-unsaturated. Let us now discuss section 3, which is about the genotyping methods of classification, where we will be speaking about DNA base sequencing, hybridization, restriction enzyme typing, then pulse-field gel electrophoresis, ribotyping, and then many other similar unique methods. So, what is the genotypic method of classification?



Genotypic refers to the genetic makeup of an organism, which includes its specific alleles and DNA sequences. It contrasts with phenotypic traits, which are observable characteristics. Genotypic analysis uses techniques like DNA sequencing and PCR to study genetic variations. Modern taxonomy is increasingly based on genetic methods, focusing on specific gene sequences. Genotyping methods analyze DNA or RNA molecules and dominate taxonomic research due to advancements in technology and the aim to reflect natural relationships.



Classification seeks to identify natural relationships encoded in DNA. One of the important aspects of this genotypic classification is DNA base content or the amount of base composition. Particularly, the GC content, which is the mole percent of guanine and cytosine taken together in DNA, varies among organisms, and these variations are used as a parameter in bacterial characterization. It is calculated with a very simple formula where we add up the total content of G and C (guanine and cytosine) and divide by the total base composition, and this fraction is converted into a percentage to yield the GC content. So, what are the various analytical methods employed?

Number one is thermal denaturation. The melting temperature of DNA correlates directly with G plus C content. Because G plus C pairs have stronger hydrogen bonding—they have three hydrogen bonds compared to AT bonding, which has only two. DNA with higher GC content will require more energy and therefore melts at a higher temperature. Another method is buoyant density in cesium chloride.

DNA's buoyant density, determined through cesium chloride gradient centrifugation, is influenced by its G plus C content. Chemical hydrolysis and HPLC: here, DNA can be hydrolyzed to release free bases, which can then be quantified using high-performance liquid chromatography to determine the G plus C content. Now, what are the various characteristics? Melting point: GC base pairs have three hydrogen bonds, as already mentioned, and AT pairs have two. Higher G plus C content leads to a higher melting point due to more hydrogen bonds.

1. DNA BASE CONTENT

The G+C content (mole percent of guanine and cytosine) in DNA varies among organisms and is used as a parameter in bacterial characterization. It is calculated as:

$$G+C \text{ content} = [(G+C)/(G+C+A+T)] \times 100\%$$

Analysis Methods:

- **Thermal Denaturation (Tm)**: The melting temperature of DNA correlates with G+C content, as G+C pairs have stronger hydrogen bonding. DNA with higher G+C content melts at higher temperatures.
- **Buoyant Density in CsCl**: DNA's buoyant density, determined through cesium chloride gradient centrifugation, is influenced by its G+C content.
- **Chemical Hydrolysis and HPLC**: DNA can be hydrolyzed to release free bases, which can then be quantified using high-performance liquid chromatography (HPLC) to determine G+C content.

So, this is the DNA melting curve over here. So, you have the double-stranded DNA. Here, you have the totally opened-up DNA strands—the single-stranded DNA, I mean—reading over there, and this is the transition from the double strand to the single strand. So, this is basically a spectrophotometric graph. DNA melting is tracked by UV absorbance at 260 nanometers.

Characteristics:

Melting Point: GC Base Pairs have 3 hydrogen bonds, AT pairs have 2. Higher G+C content leads to a higher melting point due to more hydrogen bonds.

Spectrophotometry: DNA melting is tracked by UV absorbance at 260 nm. Single-stranded DNA absorbs more UV light at 260 nm than double-stranded DNA due to the increased exposure of nitrogenous bases to the surrounding environment, a phenomenon known as **hyperchromicity**.

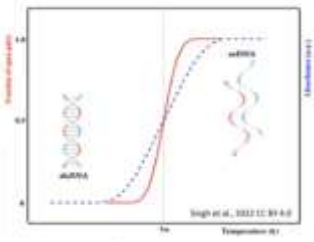
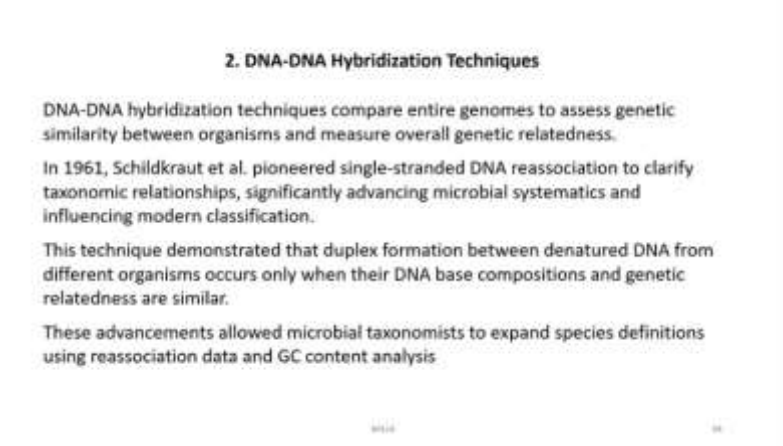


Fig. **DNA melting curve:** The plateau in absorbance indicates complete denaturation. Typically, G+C content varies by around 3% within a species and 10% within a genus. In bacteria, G+C content ranges from 24% to 76%.

Single-stranded DNA absorbs more UV light at 260 nanometers than double-stranded DNA. Due to the increased exposure of nitrogenous bases to the surrounding environment, this phenomenon is known as hyperchromicity. So, in brief, if you look at this graph, which is a DNA melting curve, this plateau in absorbance indicates complete DNA saturation, and this is the double-stranded DNA, as already mentioned. Typically, GC content varies by around 3% within a species and 10% within a genus. In bacteria, G plus C content ranges from 24% to 76%.

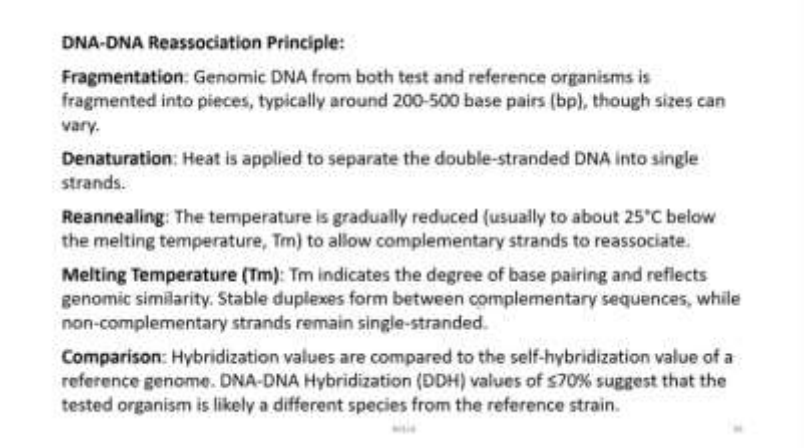
This graph will always be influenced by the G plus C content, as we have mentioned, due to the greater number of hydrogen bonds; if the GC content is higher, the curve will shift

toward the right. Another technique is DNA-DNA hybridization. These compare entire genomes to assess genetic similarities between organisms and measure overall genetic relatedness. In 1961, Childercourt et al.



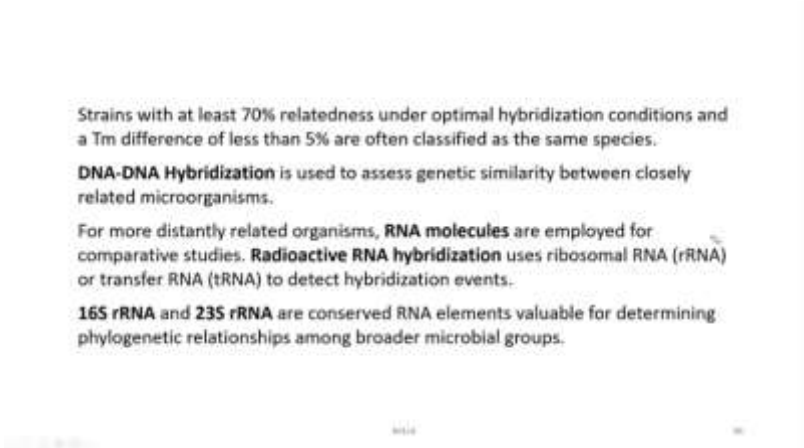
Pioneered single-stranded DNA reassociation to clarify taxonomic relationships, significantly advancing microbial systematics and influencing modern classification. This technique demonstrated that duplex formation between denatured DNA from different organisms occurs only when their DNA-based compositions and genetic relatedness are similar. These advancements allowed microbial taxonomists to expand species definitions using reassociation data and GC content analysis. So, in DNA-DNA reassociation, we first fragment DNA from both test and reference organisms into pieces, typically around 200 to 500 base pairs, although the sizes can vary. Then we proceed with denaturation: heat is applied to separate the double-stranded DNA into single-stranded DNA.

Then we proceed with re-annealing: the temperature is gradually reduced to about 25 degrees Celsius below the melting temperature to allow complementary strands to reassociate. The melting temperature indicates the degree of base pairing and reflects genomic similarity. Stable duplexes form between complementary sequences, while non-complementary strands remain single-stranded. Then we proceed with the comparison step, where hybridization values are compared to the self-hybridization values of a reference genome.

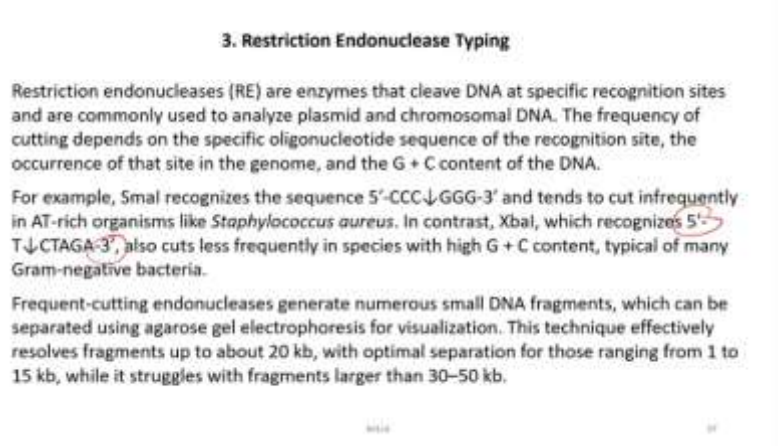


DNA-DNA hybridization values of less than 70% suggest that the tested organism is likely a different species from the reference strain. Strains with at least 70% relatedness under optimal hybridization conditions and a melting temperature difference of less than 5 degrees are often classified as the same species. And this is one important point that is exploited for this segregation and delineation. DNA-DNA hybridization is used to assess genetic similarity between closely related microorganisms. For more distantly related organisms, RNA molecules are employed for comparative studies.

Radioactive RNA hybridization uses ribosomal RNA or transfer RNA to detect hybridization events. 16S rRNA and 23S rRNA are conserved RNA elements which are useful for determining phylogenetic relationships among broader microbial groups. The next technique is restriction endonuclease typing. Restriction endonucleases are enzymes that cleave DNA at specific recognition sites and are commonly used to analyze plasmid and chromosomal DNA. The frequency of cutting depends on the specific oligonucleotide sequences of the recognized site, the occurrence of that site in the genome, and the G+C content of the DNA.



For example, let us consider the case of SMAR-1, which recognizes this particular sequence 5' to 3' triple G, and the restriction site is just in the center. This tends to cut infrequently in AT-rich organisms like Staphylococcus aureus. In contrast, XbaI, which recognizes 5' to 3' T. CTAGA and cuts between T and C, also cuts less frequently in species with high GC content, typical of many Gram-negative bacteria. Frequent-cutting endonucleases generate numerous small DNA fragments, which can be separated using agarose gel electrophoresis for visualization.



3. Restriction Endonuclease Typing

Restriction endonucleases (RE) are enzymes that cleave DNA at specific recognition sites and are commonly used to analyze plasmid and chromosomal DNA. The frequency of cutting depends on the specific oligonucleotide sequence of the recognition site, the occurrence of that site in the genome, and the G + C content of the DNA.

For example, SmaI recognizes the sequence 5′-CCC↓GGG-3′ and tends to cut infrequently in AT-rich organisms like *Staphylococcus aureus*. In contrast, XbaI, which recognizes 5′-T↓CTAGA-3′, also cuts less frequently in species with high G + C content, typical of many Gram-negative bacteria.

Frequent-cutting endonucleases generate numerous small DNA fragments, which can be separated using agarose gel electrophoresis for visualization. This technique effectively resolves fragments up to about 20 kb, with optimal separation for those ranging from 1 to 15 kb, while it struggles with fragments larger than 30–50 kb.

This technique effectively resolves fragments up to 20 KB, with optimal separation for those ranging from 1 to 15 KB, while it struggles with fragments larger than 30 to 50 KB. Another important technique is pulsed-field gel electrophoresis, or PFGE. Pulsed-field gel electrophoresis effectively resolves large DNA fragments, typically greater than 30 to 50 KB, by applying an alternating electric field from various directions, usually perpendicular to one another.

Unlike standard electrophoresis, a PFZ can separate DNA molecules as large as 10 mega base pairs, which is quite huge. The periodic changes in the electric field causes DNA molecule to reorient within the agarose gel. Larger molecules adapt more slowly to these changes than smaller ones, resulting in improved size-based separation. The efficiency of separation in PFZ is influenced by the duration and frequency of the electric field changes, making optimization of these parameters crucial for achieving optimal resolution. So, there are various ways for bacterial subtyping using PFGE, so here is one of the methods.

4. Pulsed Field Gel Electrophoresis (PFGE)

Pulsed-Field Gel Electrophoresis (PFGE) effectively resolves large DNA fragments, typically greater than 30–50 kb, by applying an alternating electric field from various directions.

Unlike standard electrophoresis, PFGE can separate DNA molecules as large as 10 megabase pairs (Mbp).

The periodic changes in the electric field cause DNA molecules to reorient within the agarose gel. Larger molecules adapt more slowly to these changes than smaller ones, resulting in improved size-based separation.

The efficiency of separation in PFGE is influenced by the duration and frequency of the electric field changes, making optimization of these parameters crucial for achieving optimal resolution.

So, bacteria are cultured in appropriate media until they reach the desired density. And these are then encapsulated with low melting point agarose allowed to solidify to form plugs that contain the live bacterial cells. So in the second step we are preparing the agarose plugs which contain the live bacterial cells. And then these plugs are then used for lysis and digestion by treating them with lysis buffer or specific enzymes to break open the bacterial cells and release the genomic DNA. And these are finally subjected to restriction digestion of DNA in the agarose plugs.

And then we prepare the gel and the PFGE is carried out and the PFGE gel pictures are captured. So, after electrophoresis the gel is stained with DNA binding dye for visualization and in the analysis step the DNA bending patterns are captured and analyzed allowing comparison of genetic relatedness among the species. So we use the help of softwares to analyze these results. We calculate two coefficients, namely the Jaccard and the Dice coefficients. And using these we generate dendrograms to find a similarity.

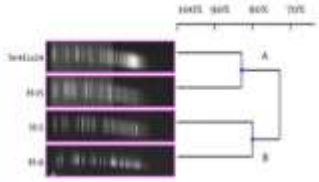
Software analyzes results, calculates Jaccard and Dice coefficients, and generates dendrograms to find similarity.

Strains with identical banding patterns are **clones**, while differences of three or fewer bands indicate **close relation**; four or more differences suggest **unrelated strains** (Tenover et al. (1995)).

Jaccard Coefficient: Measures similarity as the ratio of shared elements to the total unique elements in two sets.

Jaccard= $|A \cap B|/|A \cup B|$ Dice=$2\times|A \cap B|/ |A|+|B|$

Dice Coefficient: Measures similarity as twice the shared elements divided by the total number of elements in both sets.

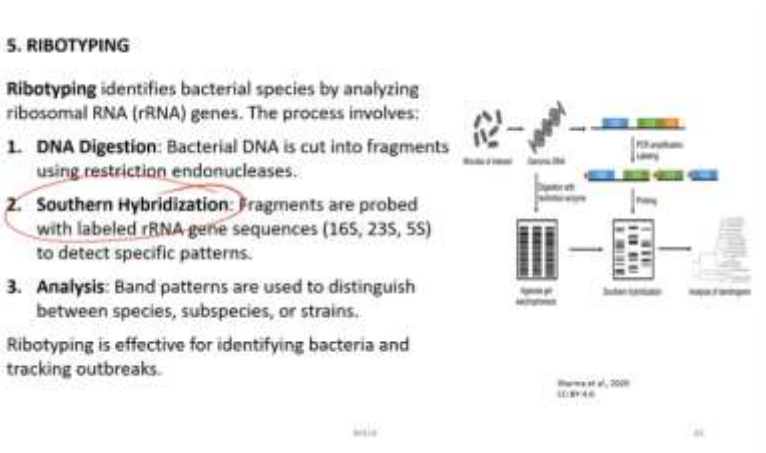


A typical dendrogram, showing the genetic relationships between the Bioluminescent *V. harveyi* Strains based on the PFGE Analysis

Strains with identical banding patterns are clones. While differences of three or fewer bands indicate a close relation. Four or more differences suggest unrelated strains. So, let us see what the Jaccard coefficient is. The Jaccard coefficient measures similarity as a ratio of shared elements to the total unique elements in two sets, which means the two different bacteria.

So, A and B, and the Dice coefficient measures similarity as twice the shared elements divided by the total number of elements in both sets, as you can see in this particular formula. The fifth important technique is ribotyping. Ribotyping identifies bacterial species by analyzing ribosomal RNA genes. The various steps of the process involve, number one, digesting the DNA. Bacterial DNA is cut into fragments using restriction endonucleases.

Then we proceed to Southern hybridization. In between, there is, of course, a step of agarose gel electrophoresis. Southern hybridization is done with labeled rRNA gene sequences to detect the specific patterns, namely 16S, 23S, and 5S. And finally, we do the analysis where band patterns are used to distinguish between species, subspecies, or strains. Ribotyping is effective for identifying bacteria and tracking outbreaks, particularly in epidemics or pandemics.



The sixth most important technique is PCR-mediated DNA fingerprinting. PCR, or polymerase chain reaction, amplifies specific DNA sequences through repeated cycles using defined primers. Some of the important points in PCR-mediated DNA fingerprinting are, number one, control amplification. PCR uses specific oligonucleotide primers to repeatedly amplify the target DNA. In DNA fingerprinting, variable regions, including tandem repeats or restriction enzyme sites, are targeted.

Specific primers are based on known sequences. Random primers are used in techniques like RAPD and APPCR. In REP-PCR, we focus on repetitive DNA sequences, offering strain typing with discrimination similar to PFGE, and the use of automated systems enhances the results. Then, we also consider amplified fragment length polymorphism, another modification. This combines restriction digestion with PCR to detect DNA sequence differences, and we will discuss AFLP in another slide later.

6. PCR-mediated DNA fingerprinting

PCR (Polymerase Chain Reaction) amplifies specific DNA sequences through repeated cycles using defined primers. Key points:

Controlled Amplification: PCR uses specific oligonucleotide primers to repeatedly amplify target DNA.

DNA Fingerprinting: Targets variable regions, including tandem repeats or restriction enzyme sites.

Primers: Specific primers are based on known sequences; random primers are used in techniques like RAPD and AP-PCR.

rep-PCR: Focuses on repetitive DNA sequences, offering strain typing with discrimination similar to PFGE. Automated systems enhance its use.

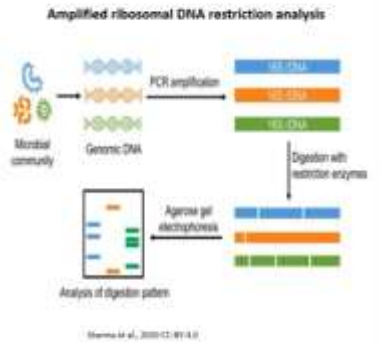
Amplified Fragment Length Polymorphism (AFLP) combines restriction digestion with PCR to detect DNA sequence differences.

Fluorescent labels and capillary sequencers enhance reproducibility and resolution.

Fluorescent capillary sequencers enhance the reproducibility and resolution. There are many other molecular methods; a detailed description of each of these methods is beyond the scope of this particular course. Still, we will have some introductory discussion on some of the important methods, like ARDRA and AFLP, which we just mentioned. DGGE or TGGE, then also TRFLP. So here, in amplified ribosomal DNA restriction analysis, we analyze the genetic diversity of microbial communities, which involves the amplification of ribosomal DNA using PCR, followed by

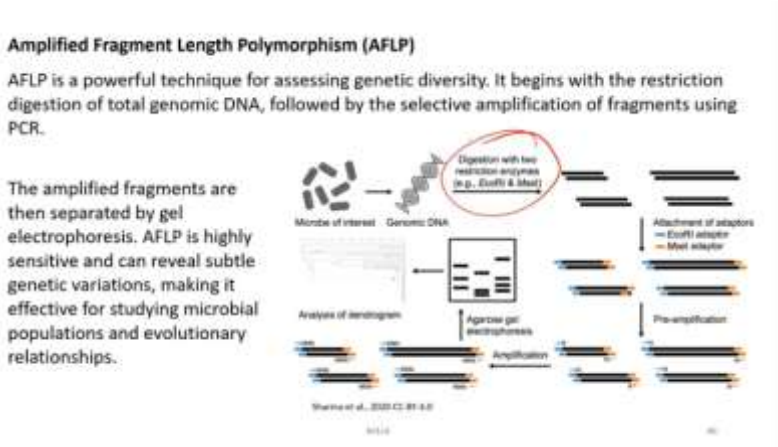
Amplified Ribosomal DNA Restriction Analysis (ARDRA)

ARDRA is a molecular technique used to analyze the genetic diversity of microbial communities. It involves the amplification of ribosomal DNA (rDNA) using PCR, followed by digestion with specific restriction enzymes. The resulting fragment patterns are then separated by gel electrophoresis. The unique banding patterns allow researchers to differentiate between species or strains, making it useful for microbial identification.

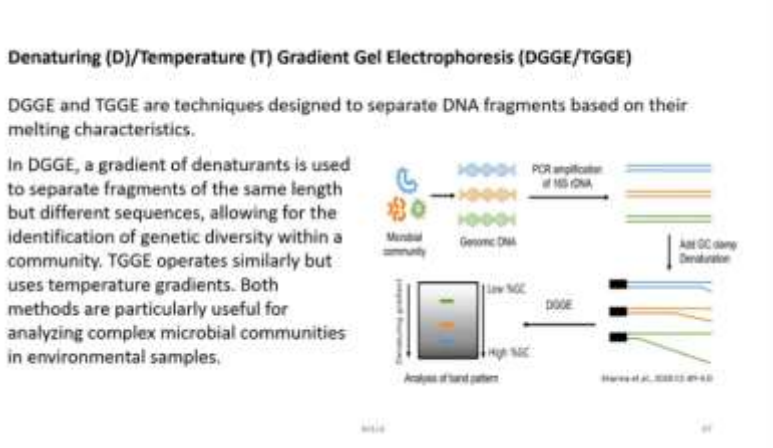


Digestion with specific restriction enzymes, as you can see here, involves the amplification step and the digestion step. And these are finally separated by agarose gel electrophoresis. The unique banding pattern that results from these procedures allows researchers to differentiate between species or strains, making it useful for identification. Another method is the Amplified Fragment Length Polymorphism. It begins with the restriction digestion of total genomic DNA.

So first, we digest this genomic DNA using two restriction enzymes, EcoRI and MseI. Then we attach adapters to the two ends cut by the two separate enzymes. We use an EcoR1 adapter on one side and an MSE1 adapter on the other side, and then these are amplified using PCR techniques. And finally, let us discuss the denaturing temperature or gradient gel electrophoresis (DGGE).

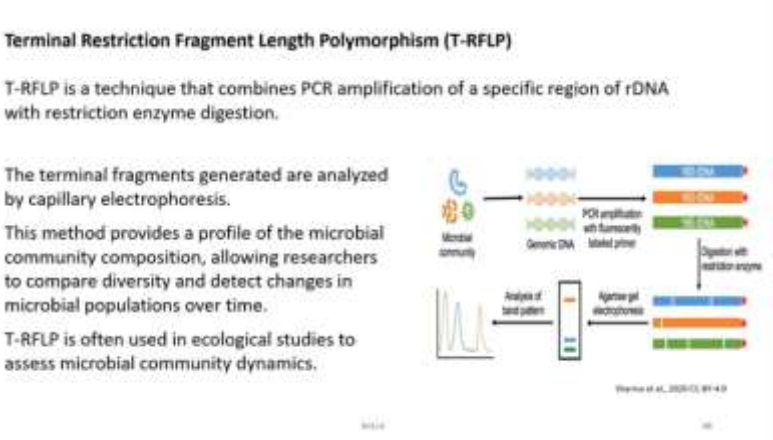


These are techniques: DGGE and TGGE. These are designed to separate DNA fragments based on their melting characteristics. In DGGE, a gradient of denaturant is added to separate fragments of the same length but different sequences, allowing for the identification of genetic diversity within a community. TGGE operates similarly but uses temperature gradients. These methods are particularly useful for analyzing complex microbial communities in environmental samples.



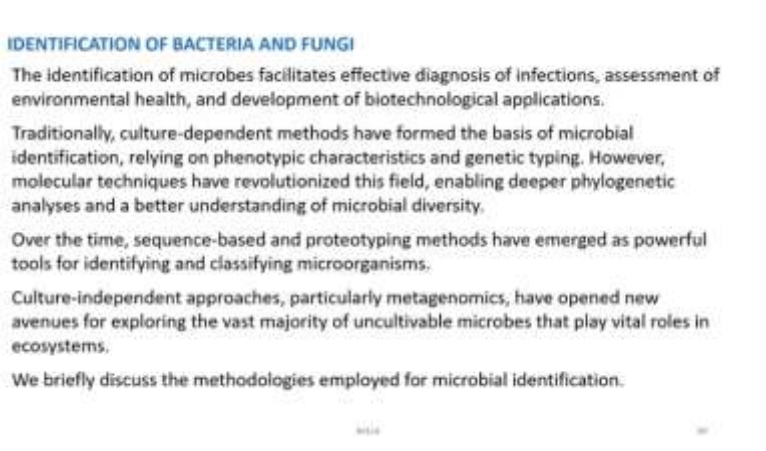
Another method is terminal restriction fragment length polymorphism. This combines PCR amplification of a specific region of our DNA with restriction enzyme digestion. The terminal fragments generated are analyzed by capillary electrophoresis. This method provides a profile of the microbial community composition, allowing researchers to compare diversity and detect changes in microbial populations over time.

TRFLP is used in ecological studies to assess microbial community dynamics. So, with these, we come to the end of the various genotypic methods of classification. Let us now discuss some of the identification methods for bacteria and fungi. We have two types of methods employed for two different purposes. The first one is a culture-dependent method.

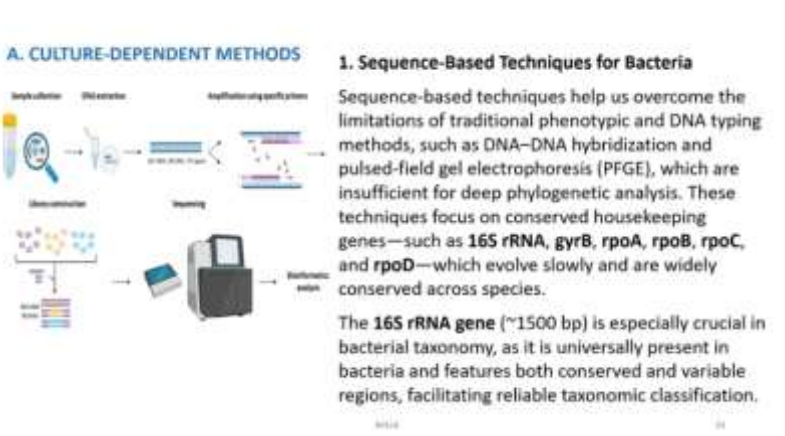


The second one is a culture-independent method. The identification of microbes, as we have already understood, facilitates effective diagnosis of infections, assessment of environmental health, and development of biotechnological applications. Traditionally, culture-dependent methods have formed the basis of microbial identification, which relies on phenotypic characteristics and genotyping, as we have discussed in the last two slides.

However, molecular techniques have revolutionized this field, enabling deeper phylogenetic analysis and a better understanding of microbial diversity. Over time, sequence-based and proteotyping methods have emerged as powerful tools for the identification and classification of microorganisms.



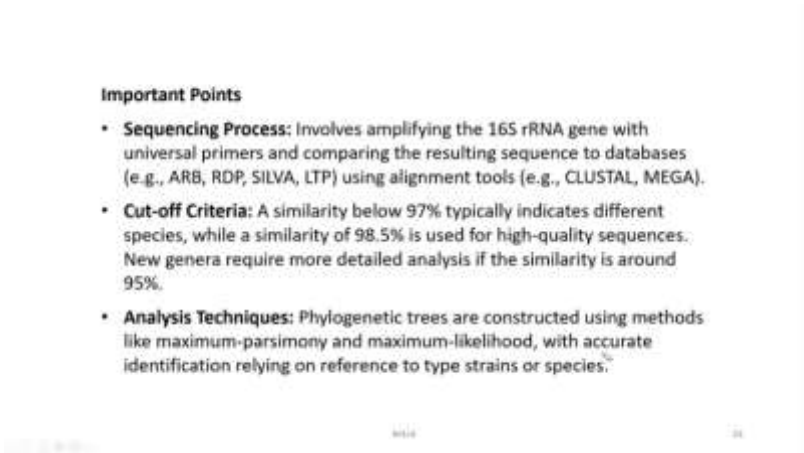
Culture-independent approaches, particularly metagenomics, have opened new avenues for exploring the vast majority of uncultivated microbes that play vital roles in the ecosystem, including our immune system. We briefly discussed some of the methodologies employed for microbial identification. The first one is the culture-dependent method, sequence-based techniques for bacteria. For example, they help us overcome the limitations of traditional phenotyping and DNA typing methods, such as DNA-DNA hybridization and pulsed-field gel electrophoresis. which are insufficient for deep phylogenetic analysis.



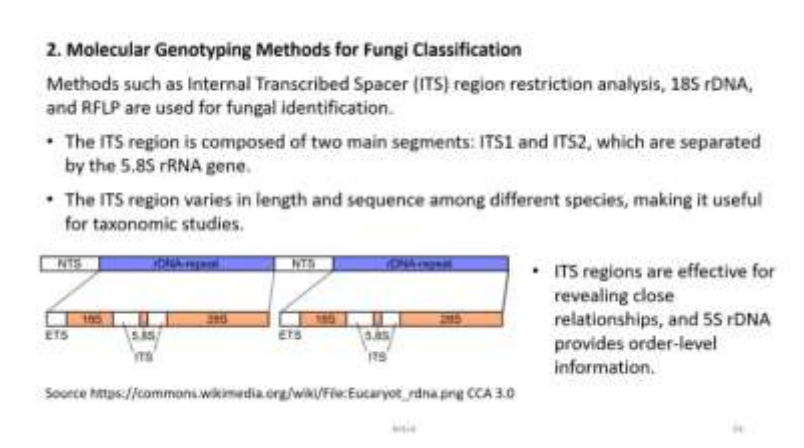
These techniques focus on conserved housekeeping genes such as 16S rRNA, gyrB, rpoA, rpoB, rpoC, and rpoD, which evolve slowly and are widely conserved across species. The 16S rRNA gene, roughly around 1500 base pairs, is especially crucial in bacterial

taxonomy as it is universally present in bacteria and features both conserved and variable regions, facilitating reliable taxonomic classification. Some of the important points in this regard are the sequencing process. This involves amplifying the 16S rRNA genes with universal primers and comparing the resulting sequences to databases like ARB, RDP, SILVA, LTP, and using alignment tools like Clustal and MEGA. Then we need to set a cut-off criterion: a similarity below 97% typically indicates different species, while a similarity of 98.5% is used for high-quality sequences.

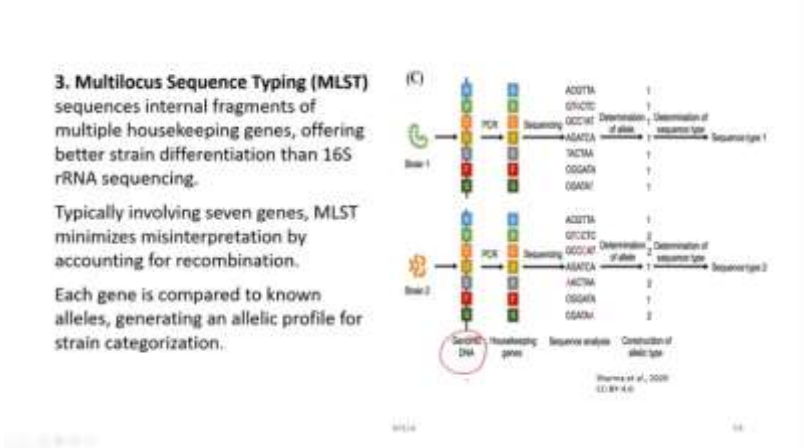
New genera require more detailed analysis if the similarity is around 95%. The other important point is the analysis technique. Phylogenetic trees are constructed using methods like maximum parsimony and maximum likelihood, with accurate identification relying on reference to type strain or species. The second method is the molecular genotyping method for classification of fungi. Here, internal transcribed spacer region restriction analysis, 18S rDNA, and RFLP are used for fungal identification.



The ITS region is composed of two main segments, ITS1 and ITS2, which are separated by the 5.8S rRNA as you can see in the diagram. The ITS region varies in length and sequence among different species, making it useful for taxonomic studies. ITS regions are effective for revealing close relationships, and 5S rRNA provides order-level information. Multilocus sequence typing, or MLST. So, MLST sequences internal fragments of multiple housekeeping genes, offering better strain differentiation than 16S rRNA sequencing.



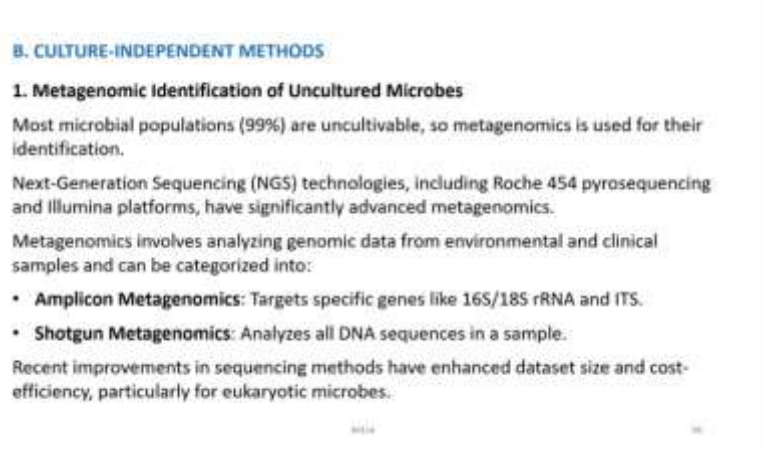
This typically involves seven genes. MLST minimizes misinterpretation by accounting for recombination. Each gene is compared to known alleles, generating an allelic profile for strain categorization. So, we have here strain 1 and strain 2, and you can see PCR is being—this is the genomic DNA. PCR is being generated for these 1, 2, 3, 4, 5, 6, 7 housekeeping genes, and then it is subjected to sequencing. In the sequence analysis, we can detect some of the differences in the bases at different positions.



So, this helps us in the determination of the allele and thereby the sequence type, which may be categorized in this case as sequence type 1 and sequence type 2. So, these steps are the construction of the allelic type. So, a major strength of MLST is its use of well-defined genetic markers—sequences of internal fragments of seven housekeeping genes, as already mentioned, like ADK, ROE, and then PharmC, GDH, GKI, PGI, and TPI are used. However, MLST can be costly and may fail to differentiate unrelated isolates due to the conserved nature of housekeeping genes.

Let us now discuss the culture-independent methods. Number one, the metagenomic identification of uncultured microbes. Most microbial populations, as I have said, 99% are reported to be uncultivable. We cannot culture them. So, metagenomics is used for their identification.

Next-generation sequencing technologies like Roche 454 Pyrosequencing and Illumina platforms have significantly advanced this field. Metagenomics involves analyzing genomic data from environmental or clinical samples and can be categorized into amplicon metagenomics or shotgun metagenomics. The first one targets specific genes like 16S, 18S rRNA, and ITS. The second one analyzes all DNA sequences in a sample.


B. CULTURE-INDEPENDENT METHODS

1. Metagenomic Identification of Uncultured Microbes

Most microbial populations (99%) are uncultivable, so metagenomics is used for their identification.

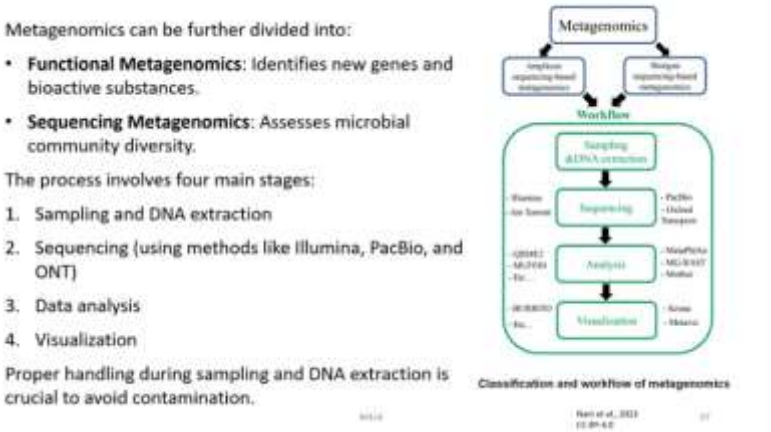
Next-Generation Sequencing (NGS) technologies, including Roche 454 pyrosequencing and Illumina platforms, have significantly advanced metagenomics.

Metagenomics involves analyzing genomic data from environmental and clinical samples and can be categorized into:

- **Amplicon Metagenomics**: Targets specific genes like 16S/18S rRNA and ITS.
- **Shotgun Metagenomics**: Analyzes all DNA sequences in a sample.

Recent improvements in sequencing methods have enhanced dataset size and cost-efficiency, particularly for eukaryotic microbes.

Recent improvements in sequencing methods have enhanced the dataset size and cost efficiency, particularly for eukaryotic microbes. This process involves four main stages. Number one, sampling and DNA extraction in both cases, whether it is amplicon or shotgun metagenomics. Then comes the sequencing, where we have different platforms available. Commercial providers and vendors help us in this step.


Metagenomics can be further divided into:

- **Functional Metagenomics**: Identifies new genes and bioactive substances.
- **Sequencing Metagenomics**: Assesses microbial community diversity.

The process involves four main stages:

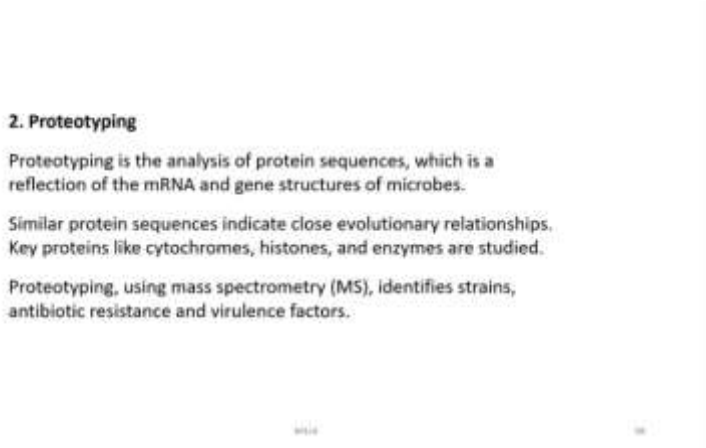
1. Sampling and DNA extraction
2. Sequencing (using methods like Illumina, PacBio, and ONT)
3. Data analysis
4. Visualization

Proper handling during sampling and DNA extraction is crucial to avoid contamination.

Classification and workflow of metagenomics

Then we proceed to the data analysis step and finally the visualization. Proper handling during sampling and DNA extraction is crucial to avoid contamination, which could lead to confusing or false results. The next important technique is proteotyping, which is the analysis of protein sequences. Until now, we have focused primarily on DNA sequences. This reflects the mRNA and gene structures of microbes.

Similar protein sequences indicate close evolutionary relationships. Key proteins like cytochromes, histones, and enzymes are studied. Proteotyping using mass spectrometry identifies strains, antibiotic resistance, and virulence factors. Several mass spectrometry platforms are used in bacterial typing. For example, we use matrix-assisted laser desorption ionization time-of-flight mass spectrometry, or MALDI-TOF-MS.



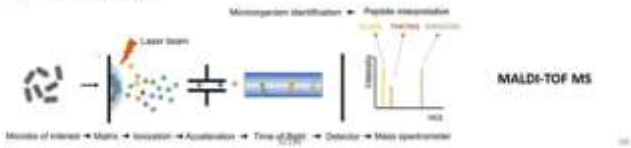
This uses laser radiation to ionize bacteria, generating fingerprints based on ion speed. Then we have LC-MS-MS. Here, peptides are analyzed after being separated and ionized, with fragmentation used for identification. Then we have targeted LC-MS-MS, which employs isotope-labeled standards for specific peptide identification and quantification. MS applications include peptide mass, new amino acid sequences, and information on protein modifications.

So here we have, for example, the microbe of interest, and this is immobilized into a matrix. And then we perform the ionization, and then you can see the representation of the time of flight. There is a detector over here which detects these ionized molecules, and then these are the mass spectrometer results that finally come out. We identify the peptides by analyzing these spectra. So, with this, we come to the end of the various classification and identification tools that are available. So, thank you for your patient listening.

Mass Spectrometry platforms used in bacterial typing include:

- **Matrix Assisted Laser Desorption/Ionization-Time of Flight (MALDI-TOF) MS**: Uses laser radiation to ionize bacteria, generating fingerprints based on ion speeds.
- **LC-MS/MS**: Peptides are analyzed after being separated and ionized, with fragmentation used for identification.
- **Targeted LC-MS/MS**: Employs isotope-labeled standards for specific peptide identification and quantification.

MS applications include peptide mass, new amino acid sequences, and information on protein modifications.



MALDI-TOF MS