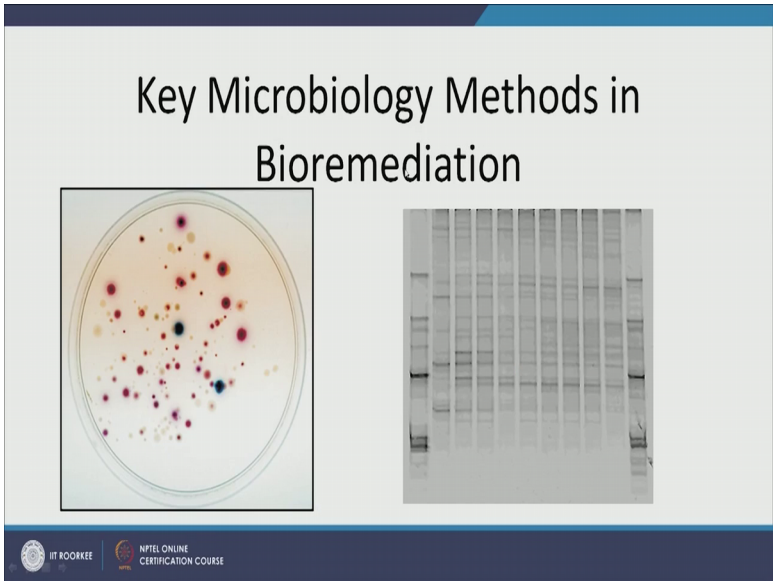
Applied Environmental Microbiology
Dr. Gargi Singh
Department of Civil Engineering
Indian Institute of Technology, Roorkee

Lecture – 32
Techniques in Environmental Microbiology II

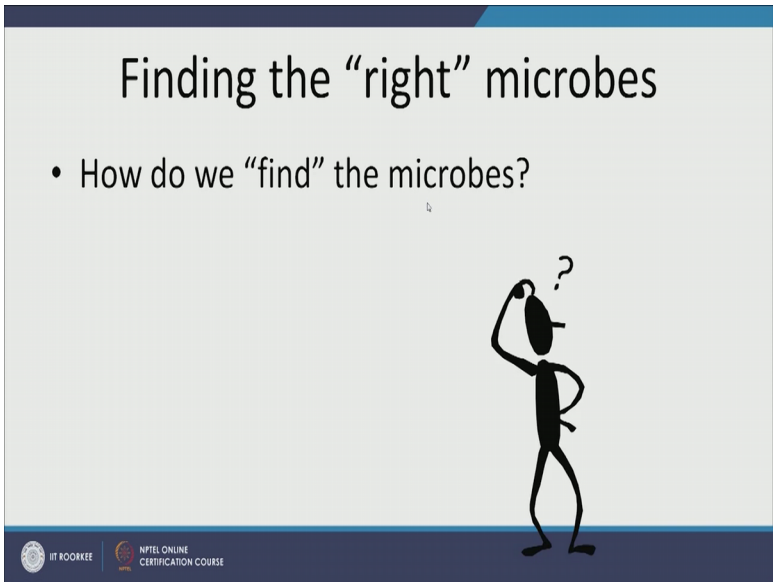
Dear students. In today's class we are going to continue from previous lecture and cover rest of the most relevant and popularly used microbiological techniques for understanding environmental problems and environmental issues. So, let us get started.

(Refer Slide Time: 00:40)



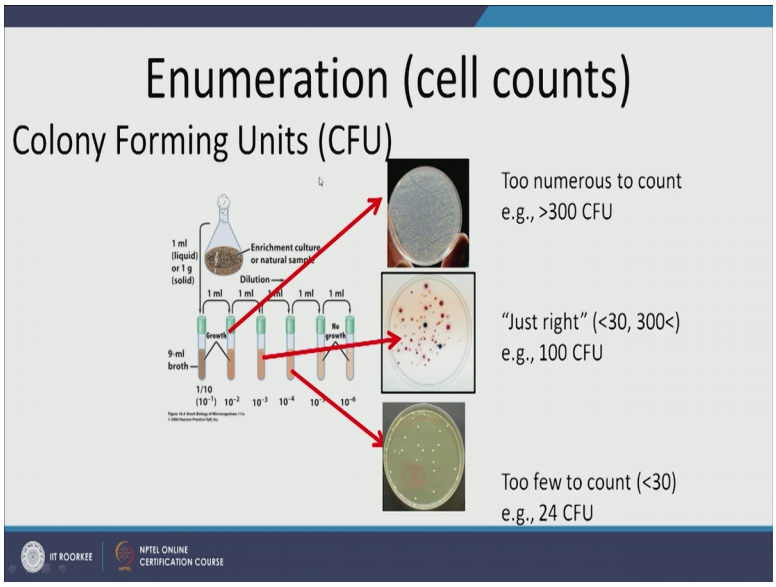
So, what we are talking about, since last lecture is are the key microbial methods used in environmental problem such as bioremediation.

(Refer Slide Time: 00:46)



So, the question is how do we find the right microbes?

(Refer Slide Time: 00:48)



We talked about most probable number in the last lecture then today let us talk about colony forming units. So, remember in NPN when I talked about, NPN in the last lecture I mentioned how we undergo serial dilution and until dilution to extinction.

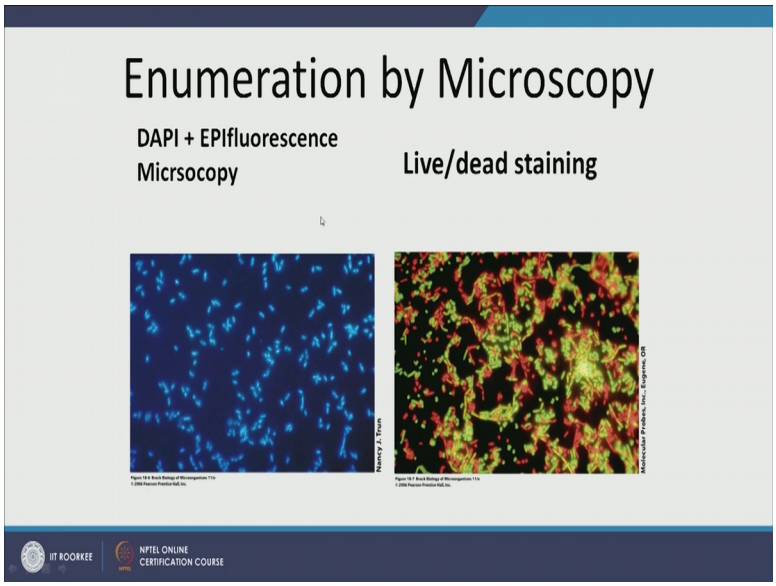
So, here initially if I plate the microbes from the first 2 tubes or the first tube or from the direct natural sample, I will get too many colonies and it should look something like this more than 300 CFU CFU is colony forming units by the way. So, this number of colonies

basically, and remember we assume that each colony comes from one bacteria. So, if you had 300 colony forming units, you have 300 bacteria in the volume that you have put here.

So, what I can do is these are too much to count I cannot count it. So, more than 3 hundreds CFU we do not count it, somewhere here in the middle will have just the right amount between 30 and 300 like 100 CFU 100 colonies. So, we can count them, but towards the dilution will have too few like 24.

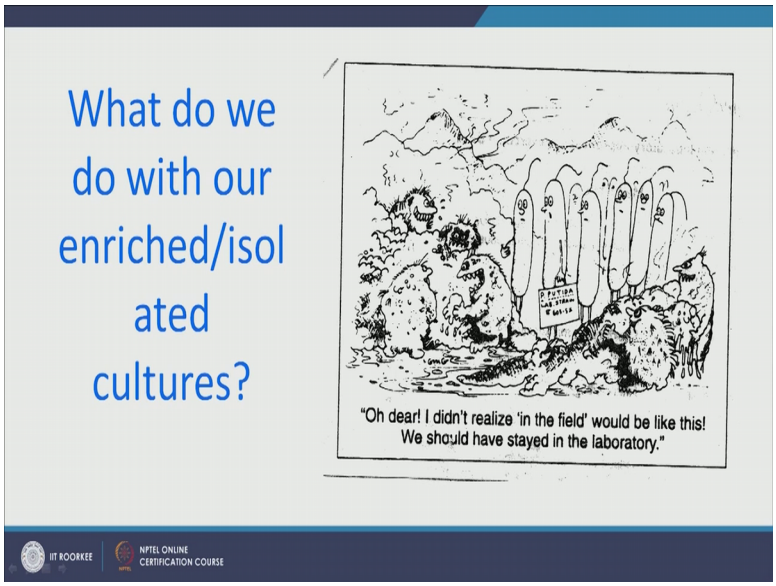
So, if it is less than thirty CFU we assume that it might just be background error it might be contamination or it is too few to be statistically significant. So, when we are doing cell counting by forming colonies on plate by plating, then we want to want to make sure that number of micro number of colonies on our plate are between thirty and 300.

(Refer Slide Time: 02:12)



Alright the other technique is microscopy. So, you can actually look at microsco under microscope and count number of cells this one is very commonly used live dead staining. So, the live cells and the dead cells can be told apart by their color the green was being alive and the red ones being the dead ones. So, you can actually take an image like this and then count how many red cells you have and how many green cells you have? DAPI EPI fluorescence microscope is very very commonly used. So, again here we can give we can stain our microbes with DAPI, DAPI and EPI and then we can count how many microbes you have.

(Refer Slide Time: 02:50)



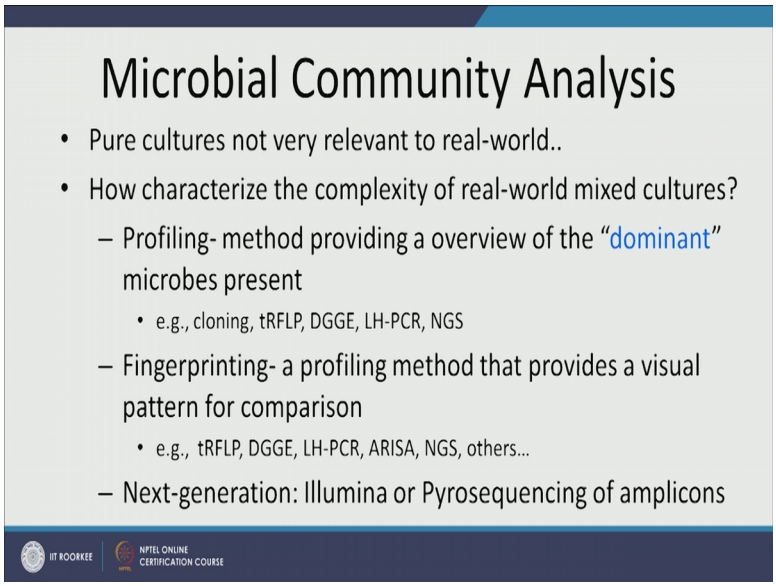
Now, the other question is I know how to count them manually by NPN I know how to count them by plating, I know how to count them by microscopy. And my dear students I must have told you in previous lectures we use also QPCR to count on microbes. Now what do we do with a enriched isolated cultures, I have enriched them, I know how many they are now what do I do with it.

So, what to do with them pure cultures are not very relevant to real world. So, let us say I took cow dung like in previous lecture I was talking about taking cow dung and then enriching the microbes that are resistant to arsenic or arsenic tolerant.

So, finally, let us say I got 5 microbes that are arsenic tolerant and then if I put them in conditions, where there are there is plenty of arsenic. I cannot expect them to thrive them in thrive in a real environmental challenge and I have this, a very nice cartoon that my adviser used to use and I am very fond of it. So, please take a look here these are ceruminous putida strains that for grown in the lab. When they come out in the real world, to clean up the mess that they were cleaning very happily in the lab, they realized that outside it is really difficult.

So, it says oh dear I did not realize in the field would be like this we should have stayed in the lab.

(Refer Slide Time: 04:05)



So, we have to understand that pure cultures are not very relevant to real world; unfortunately many parts of the world people still buy and sell pure cultures believing that they will do the job, but only realize that in lab they are much more efficient than they are in real world. So, the question is how do I characterize the complex microbial real world microbial communities?

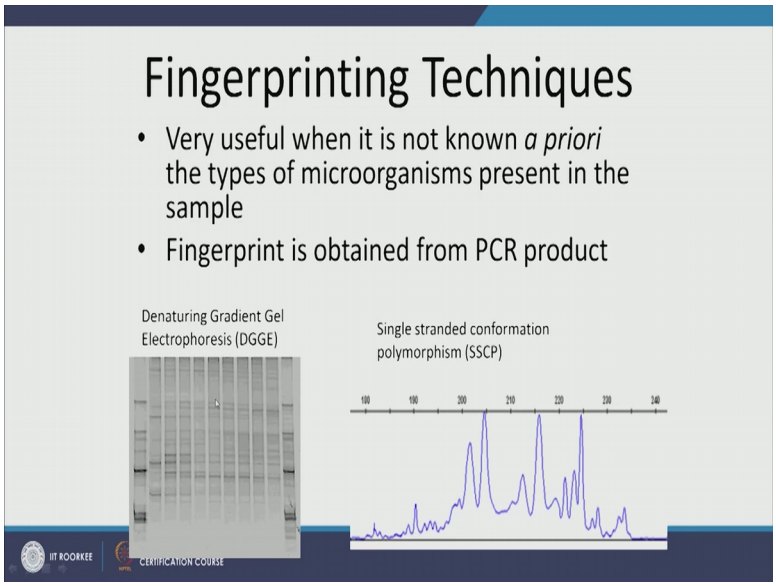
So, if you look here these are very easy to grow on plate they will form similar kind of colonies; I can count them; I can stain them, because I understand, I can count them under microscope, I can even, because I know what kind of food they like to eat? I can do NPN based techniques, but what do I do with these fellows here, who are more environmentally relevant? Who are more likely to be the one that was surviving the environment? How do I characterize them how do I enumerate them?

So, one method is profiling method it gives me an idea of the dominant microbes present. So, most profiling methods they are they do not have very high sensitivity. So, I get an idea of the microbes that are overall present; such as cloning tRFLP, DGGE, PCR, NGS. NGS is very very sensitive nowadays. So, I can get an idea of NGS by the way next generation sequencing. I can also get an idea of the less dominant microbes, but tRFLP for example, DGGE, I have already talked about it they give me an idea of the dominant microbes.

And then I have fingerprinting fingerprinting will give me an visual profile of microbial community that, I can use for comparing 2 different microbial communities again tRFLP DGGE PCR arisa they give me, I have put NGS here NGS does not give visual patterns. So, I should remove it, but we can create visual pattern from NGS data.

And then we had next generation sequencing, such as Illumina Antoine Pyrosequencing Nanopyrosequencing and we can sequence amplicons of interest. So, dear students I have said this before, but I am going to write it again. So, that you drill it in your head because this is the single most important genetic marker for enumerating how many microbacteria we have?

(Refer Slide Time: 06:10)



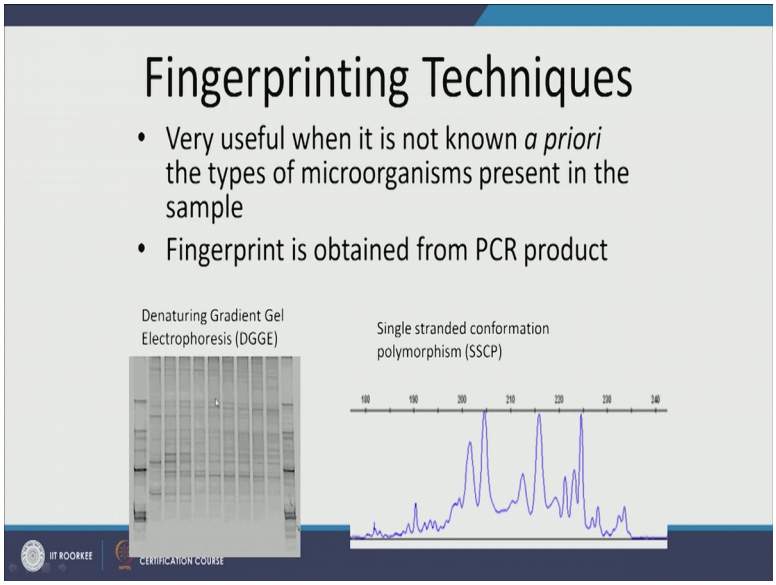
16 S r RNA gene not only tells me the number of bacteria in a sample, but it also helps me understand who is present. So, if I can get like thousand base pair 2000 base pair sequence of 16 S r RNA gene from my sample; I can align it with the well accurate databases and I can find out who is present and if I do a QPCR on it quantitative polymerase chain reaction.

So, if you go back to previous lectures I have talked about QPCR I can also get an idea of the number of bacteria present.

So, in next generation sequencing for microbial community analysis, I can amp amplify 16 S r RNA? So, instead of doing QPCR I can do PCR on it and then I can sequence it

and once I have sequenced it as, I mentioned we can align it with known databases and we will get a microbial community profile.

(Refer Slide Time: 07:19)



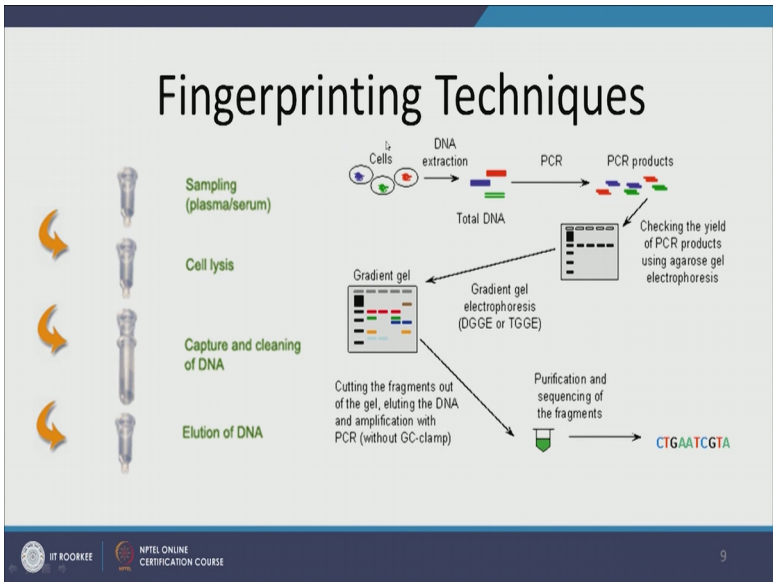
So, let us look at fingerprinting techniques most of them are visual. So, here you have a DGGE, I have already talked about it and here you have single stranded conformation polymorphism. The difference between the 2 is here as I mentioned in previous lectures that we have the gradient of adenylation agent.

So, DNA being negatively charged is pushed away. So, here we have put a samples and now because it is negatively charged it is pushed away, from the well and it flows in these straight lines, but as it flows it encounters higher and higher concentration of denaturing agent. So, they undergo denaturing.

And once the DNA under fragment undergoes denaturing it cannot it will not run away from the negative charge anymore. So, whenever they completely denature they stop there. Now the denaturing process depends upon the GC bonds. So, there are more GC bonds it will travel longer, because it will take longer to break the tRFLP strong bond.

So, GC and SSCP we get a similar information, but instead of getting bands like this we get electrophoretograms and ideally each of these peak talk about one particular kind of bacteria and then I can compare these peaks between sample and get understanding of alpha and beta diversity.

(Refer Slide Time: 08:26)



Now, all of these fingerprinting techniques, whether it is DGGE or SSCP your RF, tRFLP the first step is always extraction of DNA. So, I take my microbes my cell my environmental cell and I extract DNA. So, extraction of DNA the 3 major steps of DNA extraction are Cell lysis.

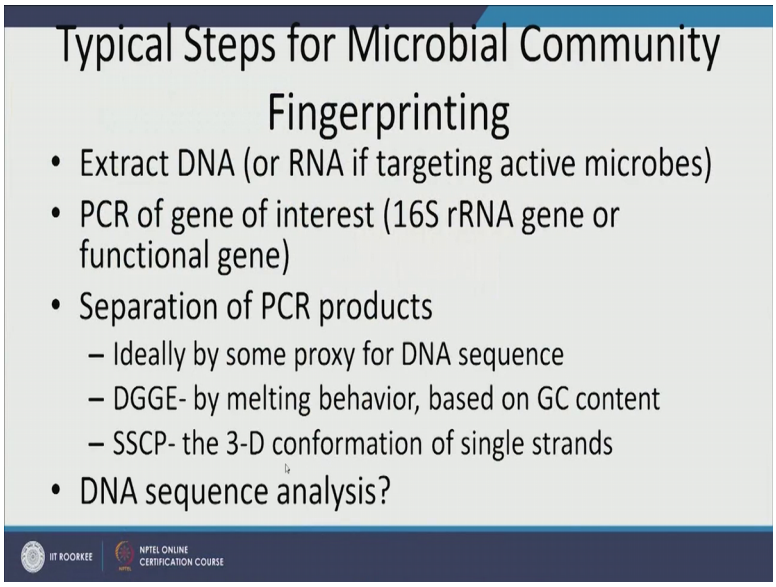
So, you would lysis the cell wall. So, that you can have DNA in suspension and then we capture and clean the DNA and this is usually a multistep process and then we elude the DNA. Typically in either in deionised molecular biology grade water or we do it in a buffer that will keep it DNA safe for long time.

So, once I have my DNA see I can do PCR. So, I get I amplify the zone of interest. Typically for DGGE especially if I of SSCP, if I am interested in characterizing microbes like who is present, then I do amplification of 16 SRRNA gene. And then I run a gel this is gel electrophoresis, I run a gel and I notice if the, my amplicon is the right size, because sometimes, I might get more primer dimers. I just talked about primer dimers. So, I think it is a good idea for me to introduce you to what primer dimer is.

So, when I do my PCR sometimes, but the forward and reverse primer will do instead of attaching to my DNA, they will attach to themselves and they will make primer dimer. So, primer dimer and they will keep amplifying themselves this typically happens when I have a low DNA starting concentration compared to the concentration of primers.

So, I will check the quality of my PCR product; I will check the yield; I will check the amount of primer dimers, I have here if everything looks perfect like; how this just looks perfect then, I would run DGGE or TGGE or whatever else makes sense to me and then I can get a fingerprint. And whatever bands look interesting to me I can cut them out I can elute the DNA and then I can sequence it and or I can amplify it and I will know what is present here. So, this is the sequence.

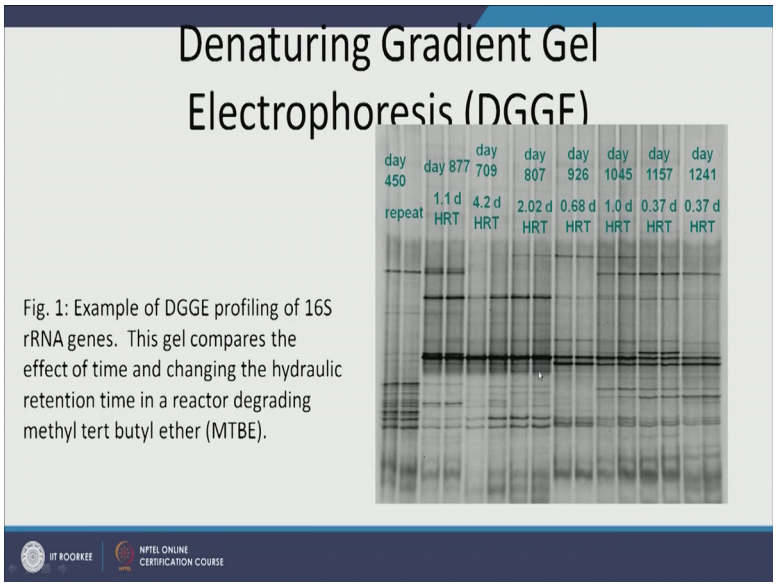
(Refer Slide Time: 10:35)



So, the typical steps for microbial community fingerprinting in summary are extract DNA or RNA, if you are targeting active microbes, because remember DNA. Even in a dead microbe or in a microbe that is in a spore form or that is not very active you can get signal you will get DNA signal, but RNA will be only present in active alive microbe.

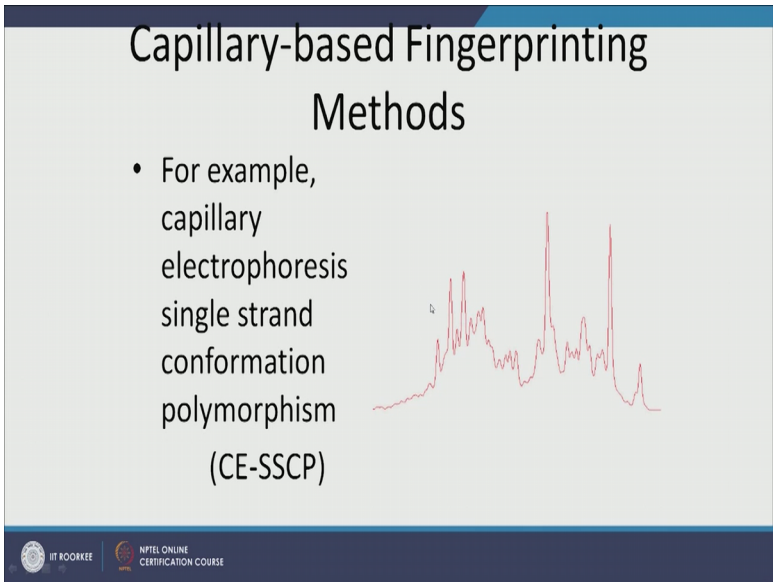
And then I can amplify the gene of my interest using PCR either it can be 16 S r RNA gene if I want to know who is present or how many different kinds of microbes are present, but if I am interested in function gene like sulphate DSR which is specific for sulphate deduction or MCR a which is specific of methanogenic methanogenesis. So, I can amplify these genes and then, I separate my PCR products ideally by some proxy for DNA sequence or by DGGE using a denaturing agent or SSCP and then, I can do my DNA sequence analysis.

(Refer Slide Time: 11:28)



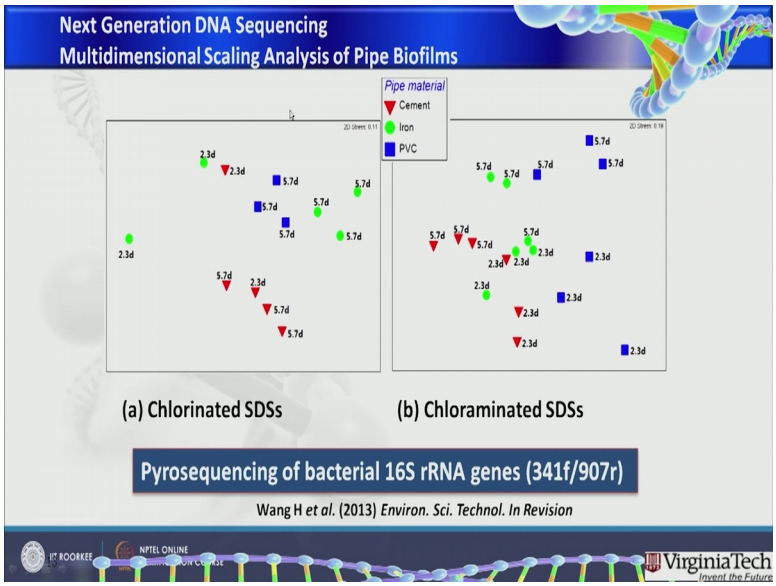
So, this is my DGGE gel and you see how in each lane they have run in duplicates and the duplicates are very similar to each other and usually distinct from their neighbours. And I can get an idea of how the samples have changed with time.

(Refer Slide Time: 11:44)



We can do similarly, we can do capillary based fingerprinting.

(Refer Slide Time: 11:49)

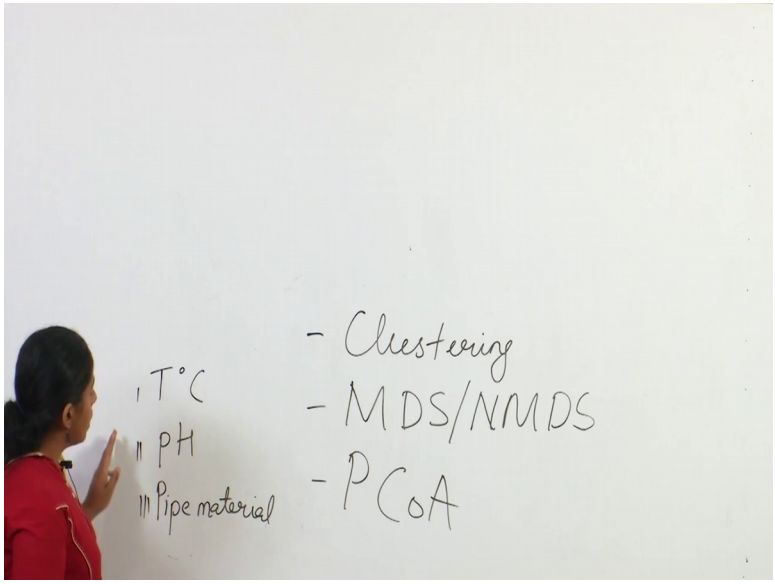


And now we have next generation DNA sequencing. So, this is actually a figure from Wang H et paper in Environmental Science Technology and she did pyrosequencing of 16 S r RNA genes that were formed by biofilms. And after she generated sequences she analyzed them and this is MDS plot by the way. And you can see that the dots that are closer to each other they are they are microbes that are similar to each other.

So, the microbial community all the bacteria that represent in the cement pipe with water edge 5.7 days, this similar to the cement pipe biofilm that had water edge of 2.3 days and all of them are seemingly clustering together.

So, and all these are chlorinated and these are chloraminated. So, we can do analysis such as MDS we can do hierarchical clusters like I shh talked about in the last lecture, you can do all these analyses from your sequencing data. So, this is more of understanding the beta diversity than alpha and then clustering the similar microwave communities together. So, the 3 most commonly used statistical techniques were while doing the microbial community analyses are.

(Refer Slide Time: 13:07)



So, the 3 techniques that are used of statistical techniques often used for microbial community analyses are clustering MDS or NMDS and PCOA. So, clustering clusters the similar microbial communities together and I can also get a p value that will tell me how significant that clustering is or is not.

In the last lecture I told you about some statistical packages that are used for clustering the same in software are also has packages for doing MDS and NMDS, basically these are dimensional scaling of our microbial community sequences. And they again do the same thing, whether you do it in 2 dimension or in 3 dimension they give you an idea of what different kinds of microbial communities are present. The closer ones will be represented by dots that are closer to each other the ones that are more dissimilar to each other will be represented by dots that are far away from each other. And we can do PCOA principal component analysis or it is also called as PCA.

This is very helpful when I am trying to understand how different factors affect my microbial community. So, I might have let us say I have set up a very well designed experiment and I have different factors that are varying such as temperatures PH, pipe material. So, I have 3 factors that I am varying in my experiment and for each of these 3 variations, I have microbial communities that are distinct from each other.

So, once I have derived the sequences for microbial communities that are that for grown at different temperature PH and a bit different pipe material. So, these are biofilms and by the way.

Then I can do PCOA analyses and then notice how the, what effect of temperature PH and pipe material was there on the microbial communities? Which was the more important one, which at what under what conditions, which factor was more important. PCOA is also used actually it is currently being used in sub by some research groups in USA to understand the source of microbes.

So, this is very much like the source apportionment studies done in air pollution, where I know the characteristics or the signatures of different sources in the environment and I know the signature or let us say the composition of air pollutant in the city. And then I can do source apportionment and get an idea of how much pollutant and what kind of pollutant came from different sources.

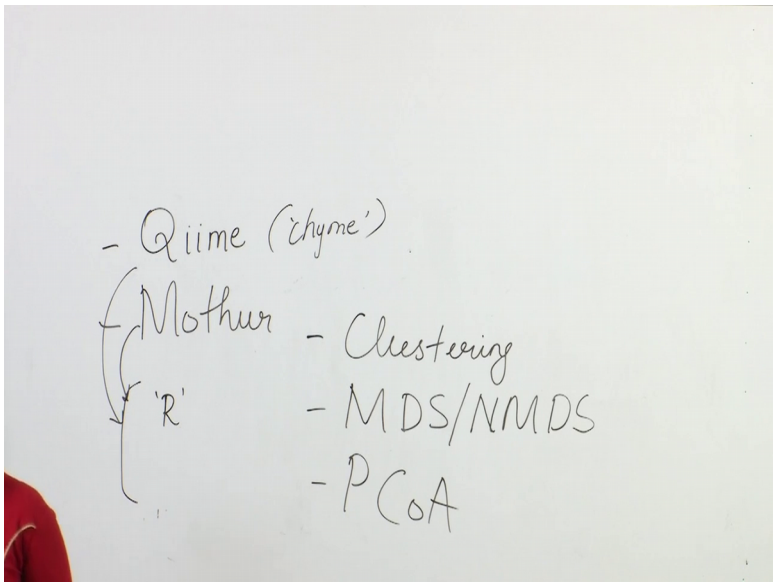
So, similarly PCOA let us say I have microbial community analyses of tap water, I know that what the tap water microbes are and I know what kind of microbes are present in the soil, in the sewage, in the droppings of wild animals and what kind of microbes are growing in biofilm, and what kind of microbes are being discharged from the water treatment plant? Then I can do a source apportionment study to understand, what component of tap water microbiome is from different sources who's more responsible is it the litter dropping of animals that somehow infiltrated into the pipe system or is it the biofilm grow that is the culprit.

So, the source apportionment studies can be done using these statistical techniques note down students keep in mind that these are not the only 3 techniques that are you, but these are the 3 most popular ones and if you plan to do environmental microbiological studies it is a good idea to familiarize yourself with them.

Now, staying with statistics here when you generate next generation sequencing data or like the slide as I showed you had pyrosequencing data or you do Illumina sequencing antorian sequencing, whatever kind of sequencing platform you use nowadays we tend to generate 1000s if not millions and millions of sequences per sample.

And when we do that it is not very easy or intuitive for us to make sense which is a good quality sequence which is not. So, in the previous lecture I just showed you a computer and the computers making sense out of it and it is telling you what it is and then you can align it to a database, but it does not work in this black box manner, where you just put it in your computer in your computer just knows. Usually there are platforms that people use to make sense of their high throughput and the consequence in data.

(Refer Slide Time: 17:38)



So, the 2 most popularly used platforms are chyme and Mother. So, it is written as Qiime, but pronounce is 'chyme'. And this is mother both of them are free chyme in my experience is more computationally intensive. So, it requires you to install the whole environment on your computer and into; mother on the other hand was more computationally light.

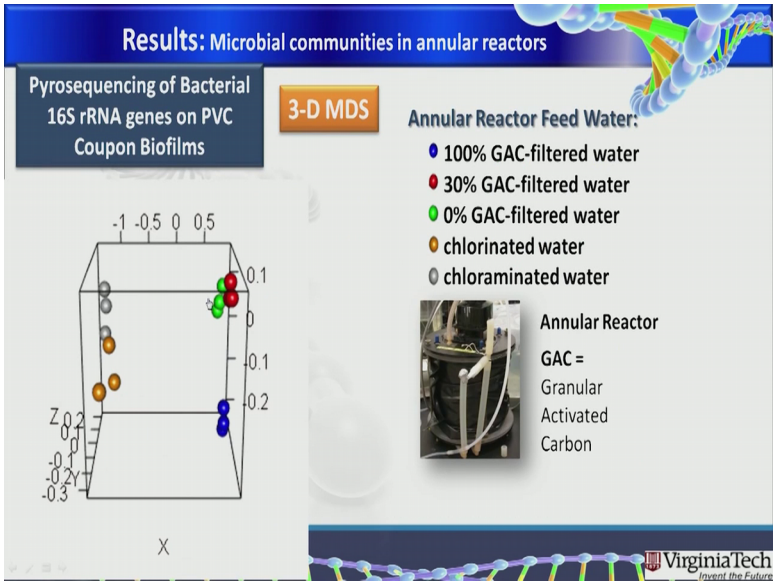
So, I prefer mother, but chyme is more user friendly and makes really pretty plots. So, and then if you have simple pyrosequencing data one you can use mother and once you derive your, if you either from chyme or from mother once you have your good quality sequences; you have separated them you have checked for chimera you have aligned them in and rotated them you have done everything, for your statistical work you can use R, which is again an open source software package that I prefer people have used S SAS SPS in different statistical packages, but I recommend R just, because it is free and it is

very easy to use definitely do not use excel. So, use any other software package except excel alright.

So, here you have once again this is from pyrosequencing data of 16 S r RNA gene of biofilms that were growing in chlorinated drinking distribution system and chloraminated water system. And the numbers on the top of these dots represent that water edge and the color and the shape of the dot represent the pipe material here.

So, you can do some very meaningful data if it is chlorinated then cement pipeline to have a particular kind of biofilms, PVC another and iron pipe another biofilm whereas, if you have chloraminated then they have a very different structure.

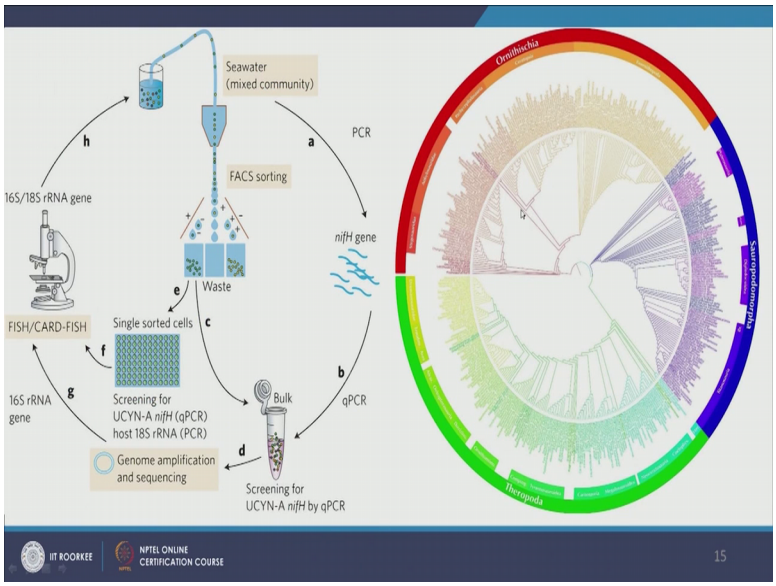
(Refer Slide Time: 19:36)



So, here you have NMDS plot in 3 dimensions. So, you have x y z and again by the same researcher and the dots represent different kinds of conditions. So, we have hundred percent GAC filtered water, 30 percent GAC filtered water, green no filtration, yellow is chlorinated water and grey is chloraminated water and in 3 dimension you can see the actually this is a movie it is not playing, but again from pyrosequencing information you can see how the blue ones clustered together and all of them clustered together.

So, if we notice that whether you are doing GAC filtration, which is granulated granular activated carbon filtration and chlorinating or chloraminating water will make a big big difference in the microbial community of your drinking water.

(Refer Slide Time: 20:24)



So here I have another pipeline that I want to briefly go through and on the right before we go through this pipeline, I am going to talk about this diagram on the right. So, on the right I have put the picture which is microbial community from sea seawater sample and, but I want you to notice the immense amount of data that has been generated and these are not very fine classification, but your this is the dendrogram by the way.

So, what I am trying to suggest by this particular diagram is that your high throughput sequencing data will be a lot of data. It is hard to visualize, it is hard to make sense out of it. So, you really need to invest some time in bioinformatics and thankfully for you we will have 2 lectures at least 2 lectures later in this course, where we will be I will be walking you through the bioinformatics platforms.

So on the left you have a panel where basically you have seawater here. So, from seawater you can do many things the first thing you can do is you can amplify nitrogen fixing, let us say I am interested in nitrogen fixation and I also want to understand the community that is here.

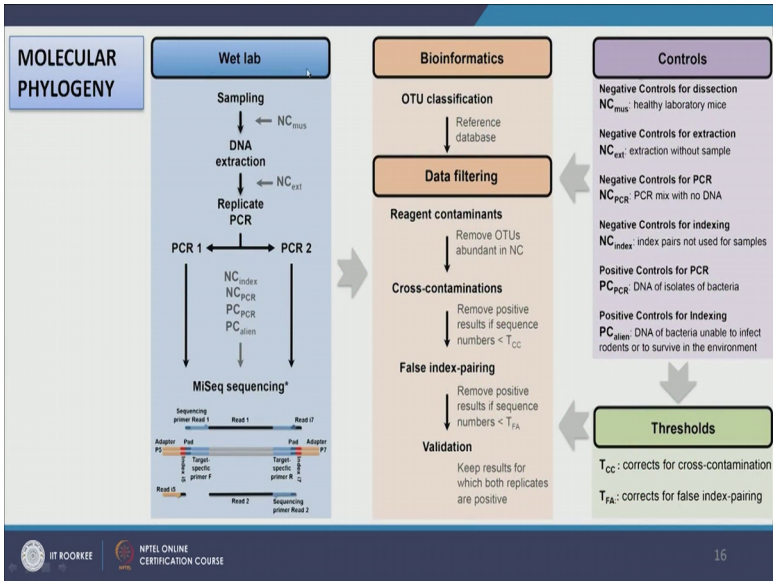
So, I can look at it in a microscope or I can take the microbial community amplify the gene functional gene of interest nifH gene, which encodes for nitrogen fixation and I can quantify it using q PCR. I can screen it using different probes I can amplify the nifH gene or 16 S r RNA gene or 18 S r RNA gene depending on what I am interested in. And I can sequence them and then I will get my profile I can also use fish card or fish card fish.

So, these are fluorescent based techniques for microscopy, I can scream them on microarray, I can also directly sort them seawater microbial community using FACS sorter. Now FACS sorter is a very interesting instrument that allows us to separate the cells based on their qualities. So, I can now using my information from this pathway for example I know what is the flour of flour or the fluorescent dye that will attach to nifH gene.

So, any microbe that has nifH gene I in the seawater even before doing any idea in extraction or anything, I can make this probe attached to that microbe or it is genome. And when it goes to the FACS sorter it will be pulled towards a, because if there is a laser in the FACS sorter, which will receive the, which will elicit the signal from the fluorophore and I can remove the micro I can separate the microbes that have my gene of interest.

So, this way FACS sorter can be used for sorting cells counting cells first of all and then sorting them according to the fluorophore according to their specificity to fluorophore. And then once I have the single sorter cell, I can do micron array analyses, I can look for I can analyze them; because I have sorted them according to their functional characteristics and I can do various kinds of analysis.

(Refer Slide Time: 23:27)



Now, here you have another very nice flowchart for doing molecular phylogenetic analyses of your sample. So, you get your wet from wet lab you get your sample you

extract DNA; you do replicate PCR this is very important. And I want to talk about what NC is here.

So, this is very very important slide. So, please take some time to look over it when I am doing my sampling? When I am doing my analyses? We have we use negative controls and positive controls. So, negative control my dear students is something that should not give you a signal ideally it should be 0, it is reading should be 0 it should give you a blank signal because your target is not present in it.

So, for example, if I am extracting DNA a negative control would be extracting DNA from pure water. So, have my deionized molecular grade pyrogen free DNA free RNA free water and if I extract DNA from it I should not get anything, but if I get a signal in my negative control then I no contamination happened.

So, negative control helps me saves me from false positives. So, during sampling I have my negative control during DNA extraction I have a negative control and then I am doing replicate PCR. Now why this is analytical replicates. So, I have the same sample, but now I am doing PCR in 2 ways or in 2 tubes to make sure if they are very similar to each other then I know, that I can trust my PCR or q PCR readings typically q PCR is done in triplicates by the way. And even here, I will have my negative control for PCR and all of these negative controls will also be amplified and then I can do my sequencing and there here they are showing my 6 sequencing, which is allumina base sequencing.

And now when I have my sequencing I will it will go through the bioinformatic pathway. In bioinformatic pathway I will do OTU class I will remove first of all the noise all the poor quality reads, I will remove the chimeras, then I will classify the OUT, then I will do data filtering, I will remove the contaminants. For example, let us say I am analyzing sludge microbial community, I should not get human cells in it, but if I am getting human cells, then most probably it is the contamination and I can ask my computer to remove them, I can remove for cross contamination.

For example, if my negative control is showing some signal what I can do is, I can subtract that signal, I can subtract those sequences from all the samples assuming that all the other samples were also contaminated to similar degree I can also remove false index pairing and I can do all kind of QC QA steps.

And then after this I can I will go and through another pipe line, which is the statistical analysis pipeline, which is not represented here in statistical analysis pipeline, I can calculate the sequencing data I can understand, I can make my collectors curves to understand if my sequencing was sufficient or not. I can do alpha diversity analysis beta diversity analysis I can do clustering, I can do NMDS, I can do PCA analysis and make sense out of my data and answer the questions I was asking.

So, dear students this is a summary 2 lecture summary of the popular and relevant microbiological tools that are used to address environmental challenges.

Thank you very much.