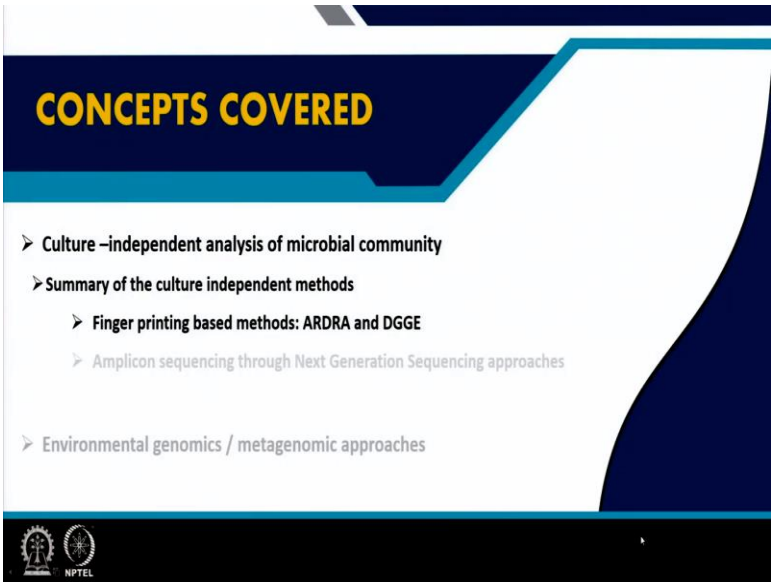
Environmental Biotechnology
Prof. Pinaki Sar
Department of Biotechnology
Indian Institute of Technology, Kharagpur

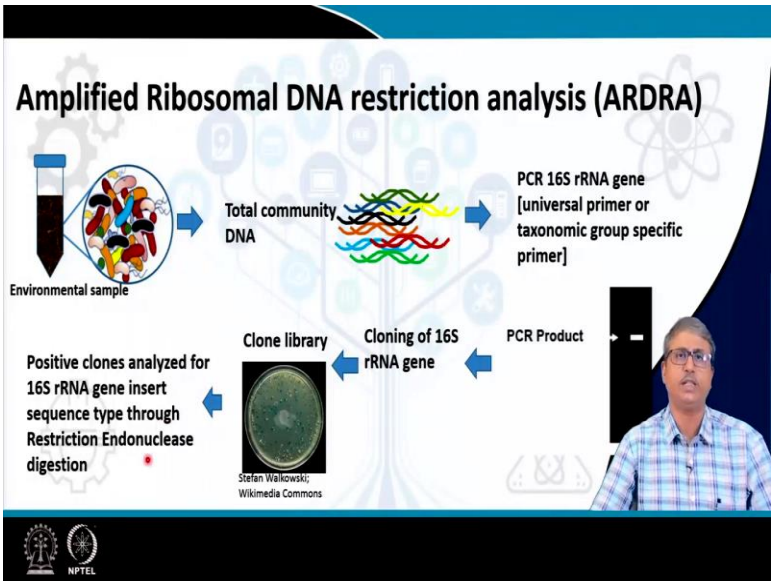
Lecture – 31
Methods in Microbial Ecology with Relevance to Environmental Biotechnology (Contd.,)

(Refer Slide Time: 00:34)



Welcome to the next lecture on the methods in microbial ecology with the reference to environmental biotechnology. And in this particular lecture we are going to talk about the cultured independent analysis methods particularly the fingerprinting based methods which enable us to perform the community profiling in any kind of environmental systems through two important approaches which are the ARDRA and DGG.

(Refer Slide Time: 00:58)



So the ARDRA or amplified ribosomal DNA restriction analysis the first one that we are going to discuss in this class is basically this will allow us to identify the phylotypes present within a community through preparation of the clone library and then subsequent restriction enzyme digestion of the 16s RNA fragments which are actually cloned from the environmental sample.. So, as you can see here the name implies this is the amplified ribosomal DNA.

So, the ribosomal RNA gene is amplified from the sample and then it is planned in order to segregate the different species member of the phylotypes into different clothes and then each of these clones are actually used to obtain the regime in charge within their plasmid and they are subjected to a restriction endonucleus digestion and then followed by distribution endonucleus the profile that we get that is actually compared.

So, the intrinsic assumption in this case is that environmental sample consists of numerous microbes. So, each of the few micrometer of the samples will have several 1000 organisms or several 100 organisms. So, when you take a sample we take multiple samples of these samples extract the DNA total DNA and that represented the total DNA that means the DNA should be representative of all the organisms present here.

And once we have this total DNA with us the DNA is subjected to the PCR application or that is the PCR of 16s ribosomal RNA gene using the universal primer or taxonomy to specific primer.

Now what is universal primer? Universal primer is a primer pair which is um optimized to amplify all bacteria. So, irrespective of its taxonomic identity all bacteria DNA 16s gene will be amplified or potentially be amplified by this set of primers.

While there are some group specific primers group means taxonomy. Group specific planners I suppose I want to only amplify psudeomonouos or I want to amplify only the gram positive bacillus type of bacteria. So, similarly there are many taxonomical specific mark members which could be used similarly. And also there is rather a primer set for archaea as well. So, that would be bacteria specific primer and archaea specific universal primer or there could be a broad primer or universal primer which will amplify all kind of prokaryotic 16s ribosomal RNA gene.

So, once we have that PCR reaction. we are going to have if we run the sample on a gel that is the agarose gel we will get a band and it is a clear band because 60 minutes based on the primers that we have selected a single size would be amplified and of course this might consist of the 16s RNA genes from all of the different members suppose in this case if you have 8 or 10 different members. So, all the members 16s are in the genes are amplified. So, ideally within this single band all the amplified products are included.

Now here we go for the cloning process and this cloning of these RNA will be done using a standard microbiological or molecular biology floating procedure using the plasmid vector. So, each of the plasmid molecule the linear molecule of the plasmid will be ligated with this P R product and the concept is one plasmid will be will be able to take one only one this year product of the 16s RNA gene and it will circularize.

So, ideally one circular plasmid circularise up on the inside the operation inside means the 16s ribosomal RNA gene in this case is inserted properly. So, it is circularised and when it is successfully transferred through the DNA transformation experiment uh. So, as you can see the blue red colony selection is used here. So, now the positive clones are picked up and all these positive clones.

Suppose I am able to pick up 100 positive clones out of the colour selection based method. So,

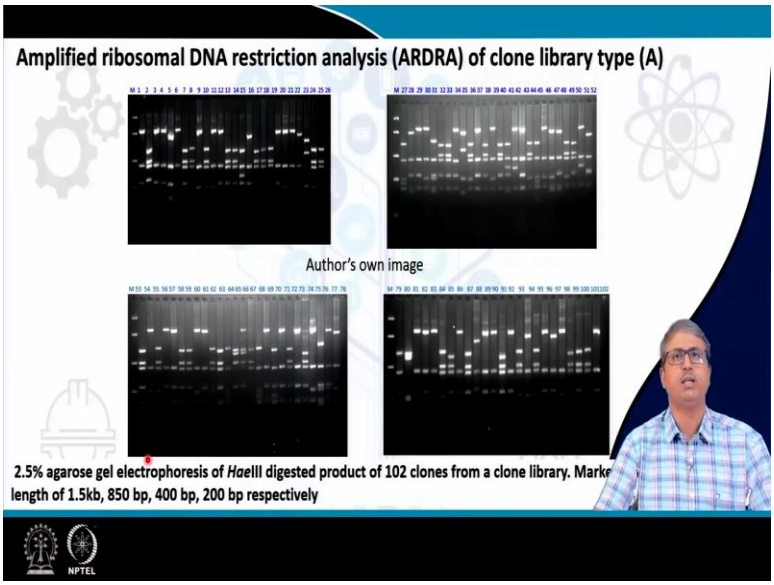
all my 100 clones are supposed to each of these 100 clones will have individually the plasmid inside them and each of these plasmids are having one 16s RNA gene insert. So, if we believe that each of these 100 clones are having a different 16s just we believe that or we take it for example then there should be actually 100 different phylotypes or 100 different species but in reality.

That may not be the case may not be the case we may find out that a particular phylotype or these species was more abundant. So, if we have taken 100 clones in a clone library we may find that actually 20 clones are containing they are plasmids having the same 16s RNA gene insert and rest 80 clones are having some the different types of 16s RNA gene insert it. It depends upon the the relative abundance of the 16s RNA gene.

And also at it is also connected to the to the original relative abundance of course and it is also partly connected to the lysis efficiency because all the cells may not be lysed in with equal efficiency. So, some cells might lies very well in the lysis condition. And some cells might be reluctant and they need some harsh condition maybe to get them our actual and the DNA to come out. So, it requires some kind of optimization.

So, the scientists who work on these kinds of things they know that type of samples they're using. So, they optimize the protocol. So, that the maximum cell lysis and all representative clone libraries can be constructed. So, it is ideal that as many as possible clones positive clones are selected so, that we can have a broad coverage of the different types of arrival of RNA genes.

(Refer Slide Time: 07:39)



Now once we have the clones with us these clones suppose in this case showing you a data with more than 100 clones. So, each of the 100 clones are having one plasmid. So, the plasmid DNA can be instructed and then can be another round of PCR can be made to re amplify the sixteenth from the plasmid and then that PCR product will be subjected to restriction in enzyme digestion.

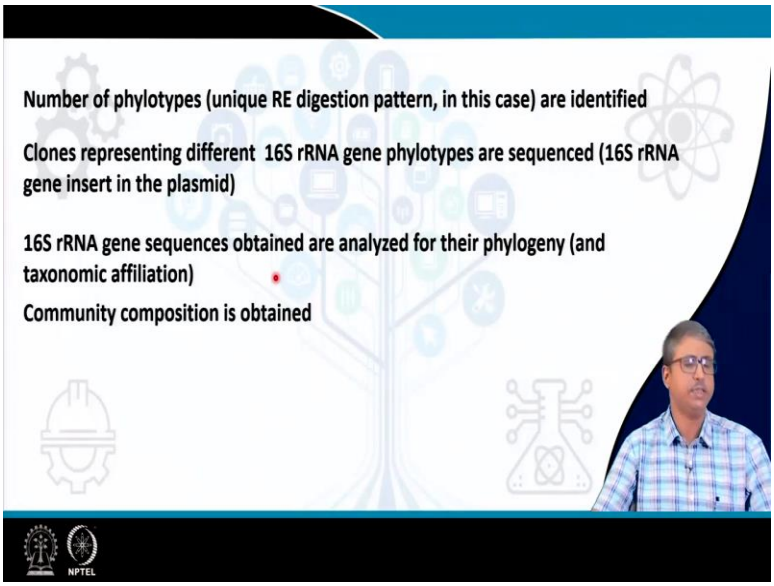
So, as we digest the sample with the restriction enzyme based on the sequence composition of the 16s gene insert derived from the plasmid. If they have similar type of sequence they will show a similar restriction digestion pattern but if they do not then possibly they will cut at different sites and essentially we will lead to a different type of restriction patterns. So, like you can look at this any of these digestion patterns and you will be able to find out that somewhere you have quite a similar type of like these two particular clones.

If you see here they are having almost a similar type of digestion that means these two 16s RNA genes are likely to be similar. Similarly if we see there are some other positions are there like this one this is also similar to the other one and you can see this in this set of clones we have some other clones which are having a digestion pattern of very distinct items like these two clones are similar while these two clones are quite similar with respect to their digestion pattern.

So, we look for this similar digestion pattern and we group them. So, as we group them we possibly reach to the phylotype of information that we group the similar digestion pattern

containing clones and their data together.

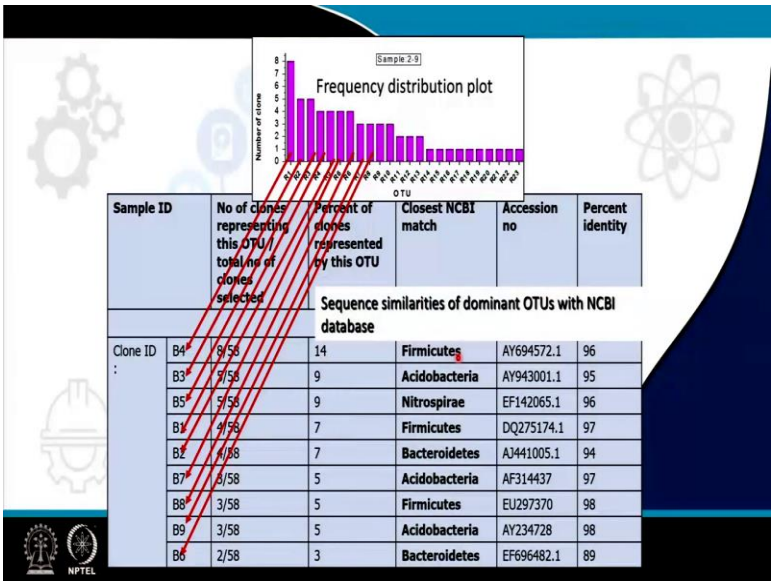
(Refer Slide Time: 09:36)



And if needed we can have actually more number of digestion so, that we it is not a single enzyme digestion separately we ideally we should digest three enzymes and then we take the data that which clones are having a similar pattern restriction digestion pattern. Even after digestion with three separate enzymes in a separate direction experiment because if two inserts are same in with respect to their sequences they will produce similar patterns with different enzymes even.

So, the number of phylotypes that is the unique restriction enzyme digestion pattern in this case are identified and the clones represented the different 16s ribosomal energy and phylotypes are sequenced and the 16s gene sequence obtained then are analyzed for their phylogeny and taxonomic affiliation and then a community conversation is obtain.

(Refer Slide Time: 10:34)

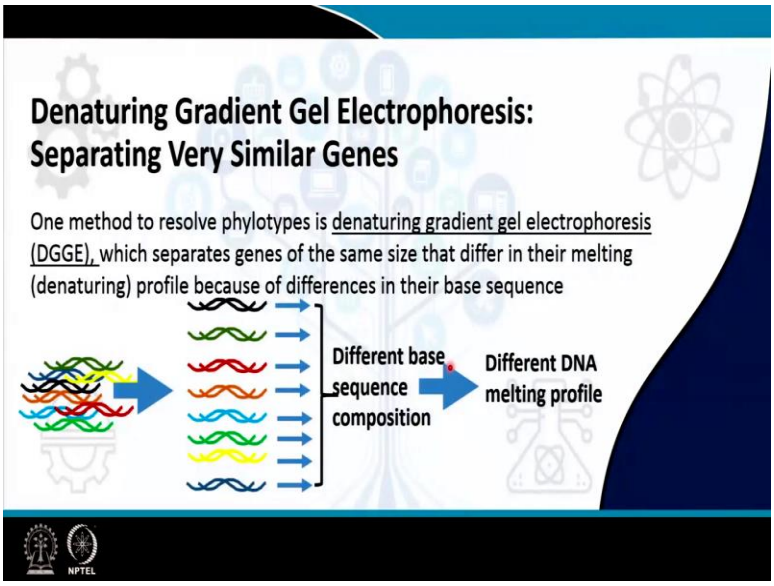


So, I will give you one example of our own study where do we work with just 58 clones we could pick up in that sample. We are just optimizing and standardizing all these things and see seeing how it is doing. So, we could have this data where 58 clones are there and you can see one particular type of restriction pattern is found to be most abundant. So, 8 out of 58 is having that particular pattern 8 clones I plasmid are having the same the same pattern.

While some other patterns are moderate and rest and it followed the pattern like this. Now some representative clones are sequenced for the 16s ribosomal RNA gene. As we sequenced or keep on sequencing these from this we are able to identify through a similarity search in the database NCBI database that the first member or the most abundant one belongs to the Firmicutes that is a group of bacteria.

The second is the Acidobacteria third is a nitrospirae, fourth is again Firmicutes. The next one is bacteria and so on and so forth. So, we can have actually up to genus level information if our sequencing quality is good and we intend to do so.

(Refer Slide Time: 11:51)



So this is one method that is very regularly followed and I will come later in this lecture with some case studies that how this is being actually used. So, the second one is a very interesting method where we do not apply the cloning based procedure in the beginning. So, rather we try to segregate the different phylotypes or different species based on their DNA composition or DNA based sequence. So, this is called denaturing gradients gel electrophoresis separating very similar genes.

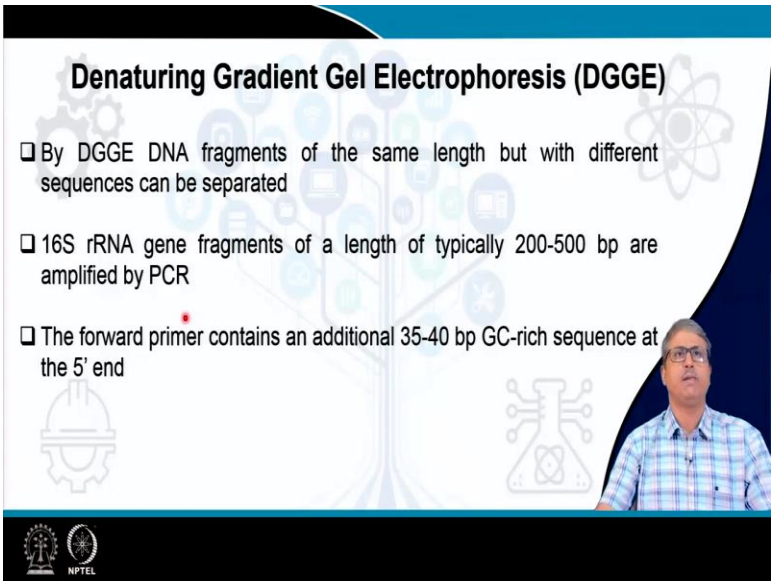
So, 16s ribosomal RNA gene for example is a very similar gene. So, it is like in a pool of 16s RNA genes which are amplified you can assume that there are many 16s RNA genes are there but they are coming from different bacteria. So, it is expected that they might be different in their sequence because they are coming from different bacteria. So, different bacteria would necessarily have a different sequence.

Now how do we segregate them no one fundamental assumption again that each of these different 16s ribosomal RNA genes actually coming from different bacterial species or taxa they have different base sequence composition they are all same gene but they have some variations in the nuclear type based composition. So, now if we; expose them these DNA fragments that is 16s RNA genes to a range of different conditions which actually allowed the DNA to be melted like to be denatured.

So, each of the DNA will show different melting behaviour melting means DNA denaturation behaviour like the DNA which will content higher hydrogen a higher number of hydrogen bonds like because of the present of more GC possibly will require higher temperature or higher denaturing concentration rather than the DNA which is having a low GC percentage or low GC concentration. So, this is one of the methods to resolve the phylotypes.

This is denaturing gradient gel electrophoresis which separates the genes of the same size because a lot of same size because we have used the same primer amplification that differ in their melting profile because of the difference in their base composition.

(Refer Slide Time: 14:04)

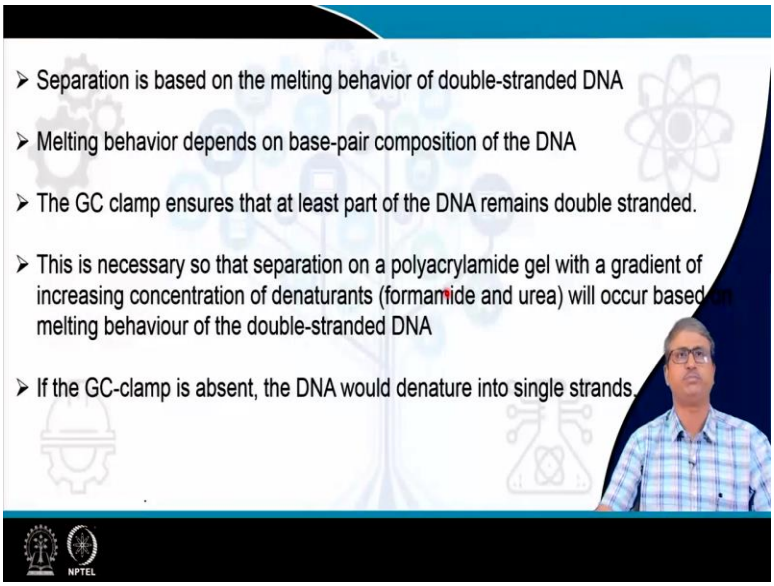


Now by DGG DNA fragments of the same length but of different sequences can be separated and generally the length of the 16s RNA gene that is amplified to P R is a 200 to 500 base pair and the forward primer in the PCR reaction contains an additional 35 to 40 base pair GC reach sequence at the 5 prime end in order to segregate and to have a sharp bands on the on the gel later when we run it on the gel.

Like for in this case and you can see that this is the first round up. So, we have maybe organism wise we have some separate organisms if you can see. And one is the mixture of these organisms as you can see that based on all are 16s ribosomal RNA gene. The same region is amplified but because they are 16s RNA gene base sequences are different some is forming band over here.

Well some other is forming band here and if we mix together we will be able to have a bending pattern like this. So, this is a kind of ideal DGG banding pattern.

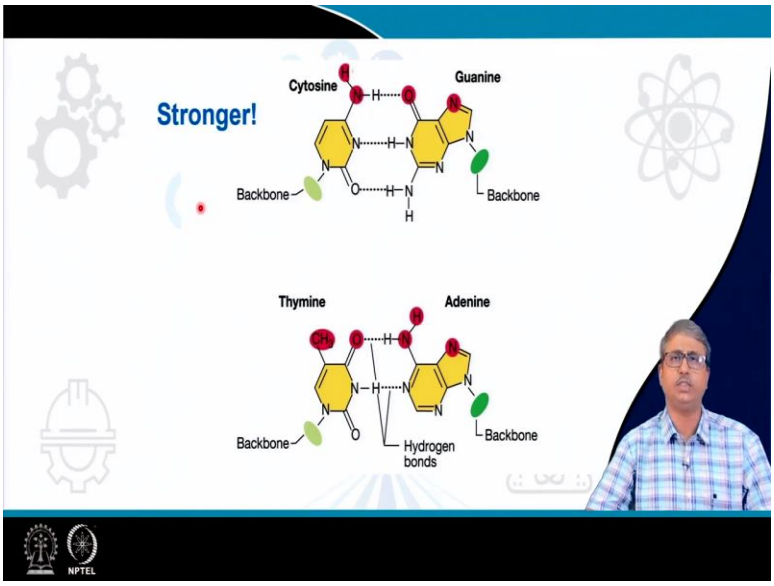
(Refer Slide Time: 15:13)



Now the separation in this case of the philosophy of different 16s based on the melting behaviour and melting behaviour depends on the base pair composition. However the GC claim ensures that the least at least part of the DNA remains double stranded. I will explain it because if we allow them to denature completely then the formation will be like the two single strands will be found and in the gel we will not be able to detect them. So, they will maybe they will migrate very fast then.

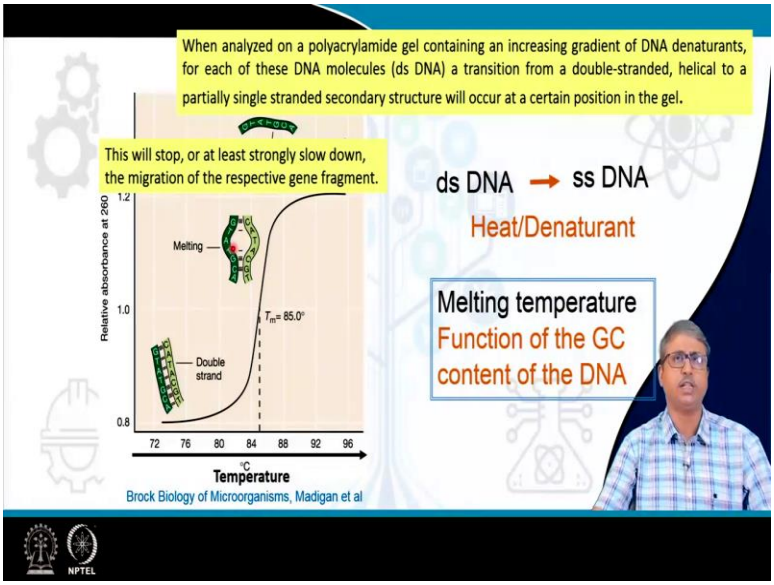
So, we do not want to lose them rather we want to have a distinct band for GC clamp is used you know so, that they denature but part of the small part of them is not the nature because of the very high GC content and then they stay or rather the structure is form like that that they will not move in the gel anymore.

(Refer Slide Time: 16:07)



Like to elaborate this like this is the DNA double helix structure and as you know that some the GC bonds are more stronger. So, if you have more number of GC in your sequence it will it will require higher concentration of the denaturant of temperature.

(Refer Slide Time: 16:26)

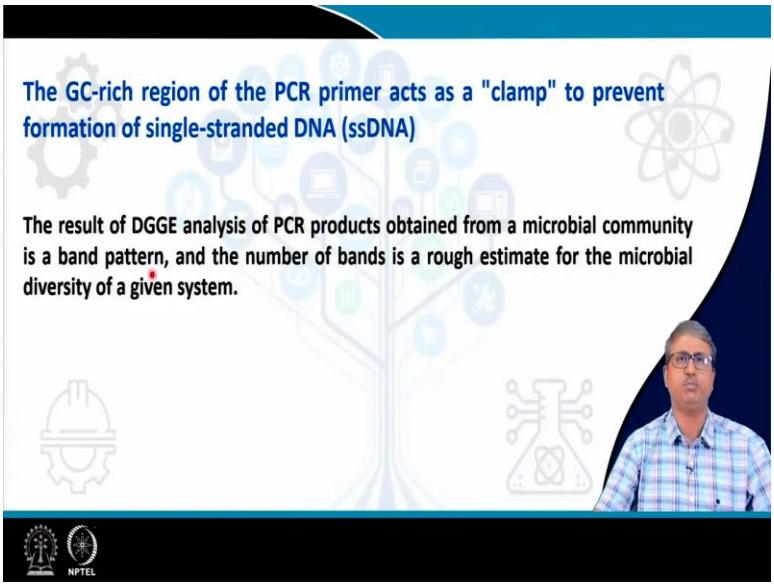


So, let us suppose we have a DNA double centered DNA and we expose it to higher temperature. slowly we increase the temperature. We will see that the double stranded structure will first start melting and this melting will occur in initiate at the AT rich region rather than the GC rich region because the GC rich region these regions will content three triple bonds sorry three hydrogen bonds are rather than the AT rich regions and eventually to break the GC bonds you need to raise the temperature to a significantly higher temperature.

And then they will be too altogether to separate it out to two single strand. So, that is called the remelting behaviour. Now in this case we try to actually use this information that if we melt up to this single stranded then possibly you will not be able to detect them rather if we have some structure like this using the primers not at the actual DNA a part of the DNA. So, we use the primary which has the very high GC content.

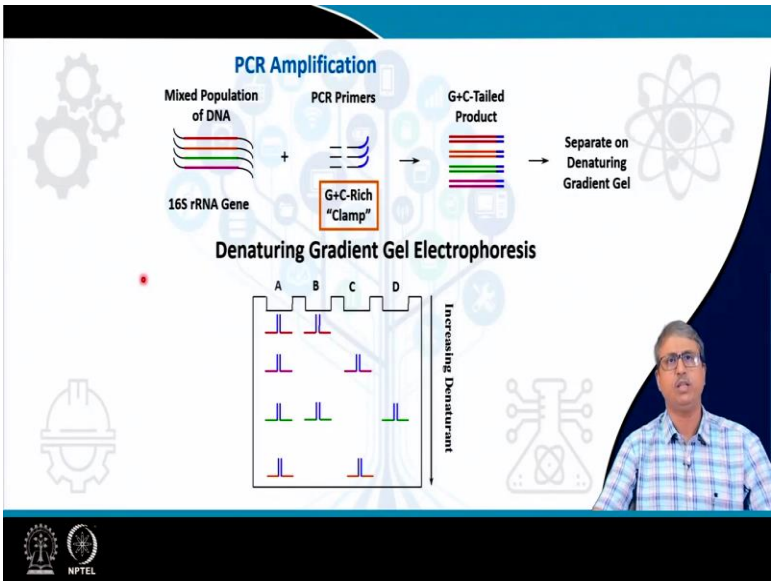
So, that part will not melt but the rest of the part will melt. So, it will form a kind of a y english Y separate structure. So, this Y separate structure and because of this structure it will not be able to move in the gel stay there.

(Refer Slide Time: 17:46)



And later when you stay in the gel will be able to see this so the basically the GC rich region of the primer is used as a clamp. And the result of the DGG analysis of PCR products obtain from a microbial community is a band pattern and the number of bands is a rough estimate for the microbial diversity of a given system.

(Refer Slide Time: 18:04)



So, ideally we have like 4 different types of 16s ribosome RNA genes having 4 different types of GC containing gc content then upon the DGG analysis we can have actually four different bands.

(Refer Slide Time: 18:20)

So, I come to the picture then we will be able to see it. So, here we are actually using some different environmental DNA. So, these are all the total community DNA from the different samples and as you can see that the DGG patterns are varying significantly. And now if I say that these samples are actually from a very extremely low pH to moderate to high pH then possibly that makes some more sense that with respect to changing pH or changing other environmental conditions.

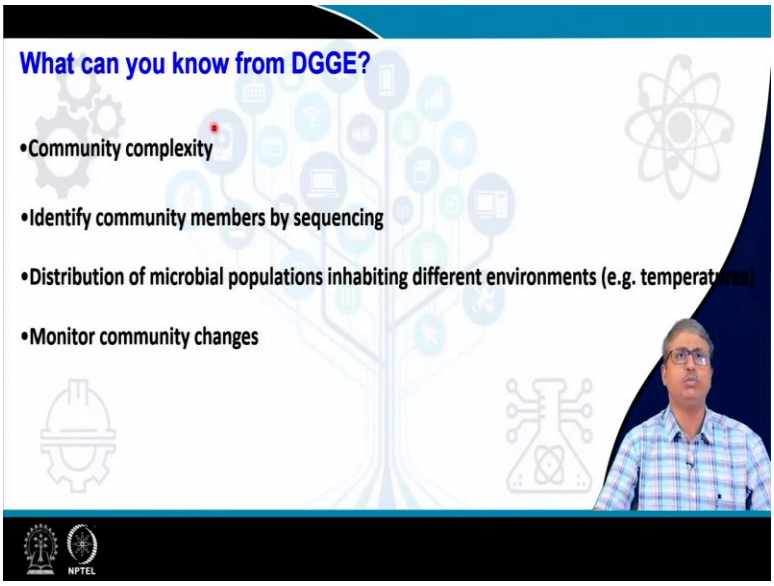
How the bending patterns are shifted. So, you can see that some bands are consistent while in sample was numbers are very the some bands are consistent while some bands are appearing in some of the samples. So, this might indicate the presence of some groups which are present only in some samples or some groups which are consistently present in more number of samples.

(Refer Slide Time: 19:12)

So, ideally each of these brands can be obtained or the bands cut with knife or sharp blade and they are processed exercised and these exercise bands are used to further amplify clone and sequence. So, that we know that what is there exactly in this band. So, ideally these band could we made up of 16s ribosomal RNA gene fragment of multiple organisms. Why multiple organisms because it is because of the same GC content they have stop that does not mean that it will contain only one bacteria.

So, one band is not indicative of one bacteria or one species rather it could be multiple species whose 16s RNA genes in this case are having same GC content. So, based on that we need to actually further in order to if we want to analyze more we can actually exercise this and a clone and sequencing.

(Refer Slide Time: 20:16)



So, what can we do we know from the DGG. So, we can actually know the community complexity because more number of bands will reflect the complexity, identify the community members by sequencing individually each bands can be cloned and sequenced distribution of microbial populations inhabiting different environments like temperature pH etcetera. Particularly in water distribution system or different type of other different processing system.

These kinds of methods are found to be very useful because it is the same water flowing through a different treatment processes. So, how the treatment process is actually working on it or a similar type of environment but there are certain differences with respect to some parameters like monitoring the community changes.

(Refer Slide Time: 21:02)

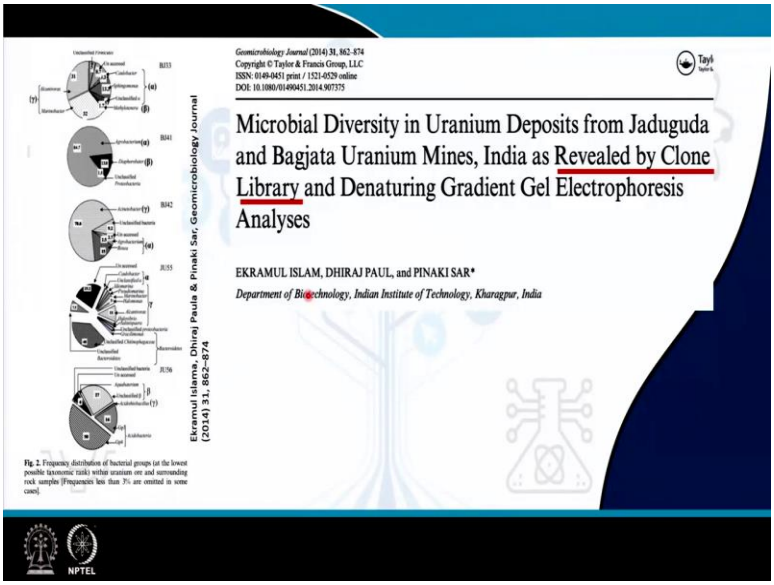
So, it has a very broad application and it is the advantage is that it is reliable, reproducible rapid and somewhat inexpensive in nature. Multiple samples can be analyzed concurrently so, the current practices that we can actually run around 15 to 16 samples at a time or even more than that. It is particularly useful to detect population change addressing the question of stability and dynamics of microbial community.

For environmental biotechnology purpose this stability and monitoring the community becomes an important question to answer. So, DGG was very successfully used before the advent of the

next generation sequence method which was found to be more precise. More deep able to identify our characterize almost all the organisms present there. So, DGG remains a very popular method in order to characterize the community composition and community dynamics.

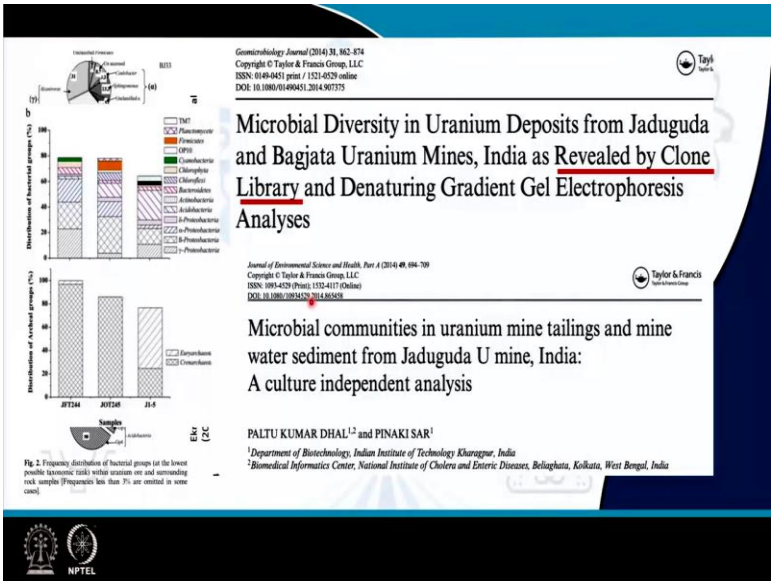
And if this is actually done with rival zone sorry with RNA instead of a DNA then it will actually reflect what types of organisms are active in this case.

(Refer Slide Time: 22:08)



Now here I present some of our studies that done using the DGG and clone library method. So, we can see that how the clone library method can give us the community composition.

(Refer Slide Time: 22:14)

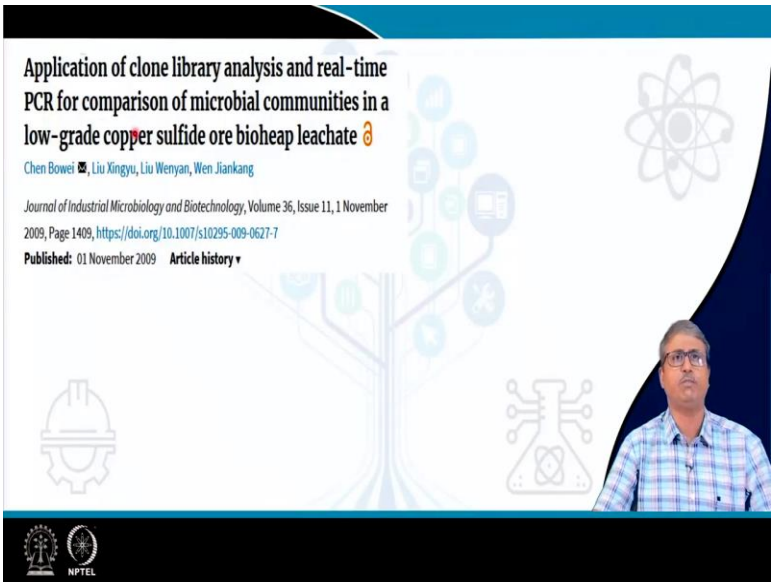


And I can take you to another so, we can have the three different type of mine tailing you can see that how the community is varying from the different type of mine tailing switch are present in the uranium mine.

(Refer Slide Time: 22:35)

And similarly there are other numerous newspapers where we see that the PCR, DGG and the closed library both are very popularly used in different waste water, waste water effluent.

(Refer Slide Time: 22:51)

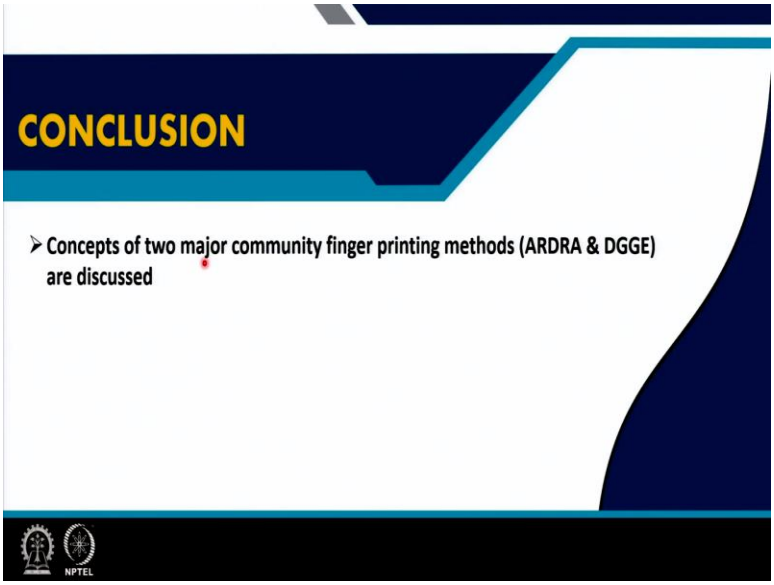


And different other processes which are having like low grade copper sulfide or bioheat leeches or maybe or maybe different other environments like succession of the communities within a very particular environment.

(Refer Slide Time: 23:04)

So, for this part in for the general text or the methods that we have discussed the Brocks Biology of micrograms will be sufficient otherwise the papers that those are being shown may be consulted for some more details.

(Refer Slide Time: 23:20)



So, in conclusion the concepts of two major community fingerprinting methods like the amplified ribosomal DNA restriction analysis and denaturing gradient electrophoresis DGG those are discussed and their applications are highlighted, thank you very much.