Comprehensive Molecular Diagnostics and Advanced Gene Expression Analysis

Prof. Aritri Bir

Dr. B.C. Roy Multi-Speciality Medical Research Centre

Indian Institute of Technology Kharagpur

Lecture 16 : DNA Microarray

Hello everyone. Welcome back to the lecture series in the NPTEL platform for Comprehensive Molecular Diagnostics and Advanced Gene Expression Analysis. So, we are going to start another week of tools for molecular diagnostics and gene expression analysis. Today, we are going to discuss DNA microarray. So, the concept today which will be covered are the basic concepts and principle of DNA microarray mechanisms, how the DNA microarray actually happens and finally, what are the different applications in terms of medical field and healthcare facility. So, let us see now DNA microarray you can see there are multiple names for this method like DNA chips, gene chips, DNA array, gene array, biochip.

So, this term chip is basically coming from the semiconductor industry and that is being combined with the life science department. So, here we are talking about seriously we are talking about chip which is being used in the healthcare facilities diagnostic part. Now DNA microarray is a kind of technology in which specific DNA sequences are arranged in an orderly fashion over a surface mostly of glass or silicon to detect not only detect to also quantify genes. So, you can see we were talking about chip.

So, you can see this chip can be hold in between these two fingers this small amount of chip you can see here. So, this is the chip over which this DNA microarray is done. Now the basic principle for DNA microarray is the charge off rule that is nucleic acid hybridization where the two strands of DNA can be joined by hydrogen bond based on the complementary, complementary sequence complementary. So, basically this concept once again you can see it is just similar to southern blotting technique and the first the microarray the mini miniaturized version of microarray for gene expression profiling was first reported in the year of 1995 and in 1997 one complete eukaryotic genome that is saccharomyces cerevisiae was published through microarray. So, here you can see the basic principle is nucleic acid hybridization.

So, these DNA chip these part these part of the chip basically contains multiple single

stranded DNA sequences. So, if we can modify this picture just we can expand the picture. So, you can see these square shaped chip among them among them there are multiple such squares you can see these are the multiple squares and these squares are known as spot or features and these features contain millions of copies of identical DNA strands single strand and those single strands are called probes which represent a single gene. So, what you can see in these these are the features and in such a single feature there are multiple single stranded DNA and these are called probes and these square is called the feature. Now, to that probe definitely some target will be there and those targets are our sample DNA.

So, what we do we collect the sample from which we isolate and purify the sample DNA those DNA are converted to cDNA via micro RNA that is via you know reverse transcription and those cDNA are coupled or joined with some dyes some fluorescent dyes and finally, those coupled sample DNA strands are made to hybridize with the probes which are located over the chip. So, here you can see there is hybridization between the tagged sample and the probe and finally, on hybridization there is some signal which is generated and detected over the scanner and interpreted in some analyzer which is definitely some computer based analyzer. So, these are the basic steps of DNA microarray. Now so, what happens basically? So, there is a target, target is our sample DNA which is labeled with the fluorescent dye and it hybridizes with the probes. Probes are basically captured over the array surface and on hybridization there is some signal and by that signal we can detect whether some gene sequence present or not even we can relatively quantify relative concentration can be measured as well.

So, there are different types of DNA microarray based on the fabrication process. How the chips are generated? How the chips are fabricated based on that there are two types of DNA microarray one is spotted microarray another is in situ microarray. Now in spotted microarray the probes or the specific sequence which we want to detect those single stranded DNA those are spotted means what we do in case of spotted DNA microarray at first we synthesize the probe via PCR mostly and those probes can be some cDNA can be oligonucleotides and those pre synthesized probes are spotted or deposited over the array surface. How they are deposited? By different there are different methods, but the commonest are pen tip method you can see these are the pen tips based by which we are depositing the probes over the array surface this is the array surface this is the array surface or there can be inkjet method as well. Now initially those were used to be done manually, but presently there are robotic arms or the machine version based on which we can spot the probes over the microarray surface.

Now what is then in situ microarray method? Definitely in in situ microarray we do not need the pre synthesized probe rather over these microarray surface the target sequence or the I mean the over the sequence we which we want to detect those are synthesized

directly over the plate or array surface. So, that is why it is called in situ and what are the method that is photolithography method or even inkjet method can be done and of course, in case of in situ microarray in most of the cases the probes are short oligonucleotide sequences and those are present in high density. So, sometime this in situ microarray are also known as high density microarray method. So, you can see this is the inkjet method. So, there are different just like the pen nibs there are different ink bottle types container and those are sprayed or can by can be spotted by the tips by the inkjet method and this is the photolithography method.

So, we are going to discuss those just a short time later. Then you can see these are the advantages. So, basically when there is in case of spotted microarray what happens we can design our target sequence based on which we want to detect the sample. So, those are relatively inexpensive because we are pre synthesizing the materials and spotting directly over the array surface and they are flexible we can design our own probe, but what is problem with in situ we cannot design our own probe rather we need to purchase because those in situ plates are designed or synthesized by some specialized machineries, but there are some advantages like those are less time consuming because in case of spotted microarray we need to convert the DNA to mRNA to form the cDNA and then we need to level this all these stages are already done over the commercialized in situ array plates and they come with high specificity, but definitely because the sequences are fixed the flexibility to detect different sequences are not present. We need to detect the we can detect those sequence which are the sequence which are available over the microarray plate and definitely because there are specialized machinery required we the method is costly the product is costly.

Now we are going to discuss the photolithography technique. So, in situ DNA microarray as I told you they can be generated by photolithography or also can be generated by inkjet method. In photolithography what happens the commonest probe is oligonucleotide probe and the company which is commonly generating this photolithography based in situ DNA microarray is Affymetrix Affymetrix gene chips. Now what happens this in situ probes are directly generated over the array surface and that is one nucleotide at a time per spot and many such spots are generated simultaneously and the method is definitely photolithography which means there is a pattern in the light when we say photolithography means using the light we are designing some pattern. Similar such method we are following here what happens here you can see the UV light is directed through a photolithographic mask.

So, this is our photolithographic mask. Now what is the specialty of this mask? This mask is designed in such a way that the exposure sites can be chosen as per requirement. So, you can see over this mask suppose this is our spots and our masks are generated in such a way that we can choose that we just want to synthesize the oligonucleotide over

this spot. So, these features or spots can be masked. So, this is the specialty of this photolithographic mask.

So, here you can see over this mask these areas are masked whereas, these part are unmasked and these unmasked part UV lights are going UV lights are exposed over a solid support. Now this is our solid support or chip which are mostly generated by glass or silicon over this solid support there are covalent linkers present covalent linker molecules which are protected by some protecting groups and what is the specialty of those group those are those group can be removed by light. So, those are vulnerable to lights. Now what happens when there is exposure of UV lights those these are the protecting groups. So, you can see in these regions only UV lights are exposed through the photolithographic mask and they are the protective groups are removed and they are activated by addition of hydroxyl groups.

So, here are hydroxyl groups which are . So, the covalent linker here is activated. Now there will be some protected nucleotides which are exposed over those unprotected covalent linker. So, you you can see there are this protected nucleotide which are joined here only to the unprotected areas. What is the next step that is coupling? So, again we are moving the protecting protective mask or the photolithographic mask to the next spot in such similar multiple sequences or cycles different protected nucleotides you can see here are added.

So, basically this addition of nucleotide occurs only on the chosen unprotected or unmasked side and we can go for multiple cycles for this and can use another new mask to unprotect or expose different sites. Now what happens there are simultaneously there are 4 mask present per round. Why 4 mask? You can see one mask is required to deprotect where at the other 3 mask basically prevent the light to the same spot other 3 mask are preventing the light from the other 3 nucleotides. So, there are 4 nucleotides you all know. So, when there is exposure of one area to one nucleotide other nucleotide addition should be protected.

So, there are basically 4 mask one is for unmasking and another another 3 for masking. So, there are multiple similar such cycle by which using the photolithographic plates we can this in situ photolithography based DNA microarray plate is generated. But in case of ink jet spotting method what happens there are ink tanks there are 5 ink tanks 4 contains nucleotide 4 nucleotide precursor along with catalyst and also there are coupling and deprotection steps as well, but what is not there that is photolithographic mask. So, ink jet spotting method is used to treat the photolithographic mask here rather this ink jet sprays the nucleotide over the array surface in picoliter volume and those are accumulated base by base printing. And because this ink jet spotting method are they do not require specific photolithographic mask synthesis of longer molecule longer

oligonucleotide strands can be possible whereas, in case of in situ photolithographic technique mostly 25 mirror lengths oligonucleotide can be synthesized here it is 60 mirror length.

So, you can see probes can be of 2 different types one is cDNA another is oligonucleotide probe. So, what happens what is the basic difference between this c nucleotide and cDNA and oligonucleotide probe. So, in case of cDNA that cDNA is basically synthesized from PCR. So, via PCR we are getting the target sequence. So, the synthesis of cDNA probe is flexible can be done from the gene bank sequences can be selected from the gene bank or gene library and one additional step is required that is for tagging the fluorescent dye here you can see tagging the fluorescent dye after getting the total RNA from the PCR product and we are getting the cDNA probe.

Whereas, in case of oligonucleotide probe these probes are basically biotin labeled basically this oligonucleotide probe they are mostly synthesized in in situ pattern. So, the in situ pattern synthesis cannot be done directly via directly tagging the fluorescent dye instead of that what we do we do biotin labeling and those biotin label are next attached to some other fluorescent detector by streptavidin based dye. So, an additional step is required in case of oligonucleotide probe. So, what are the different application of DNA microarray definitely DNA microarray is the method these days are utilized for intense gene expression and profiling where we can detect presence or absence of different specific genes. We can compare the expression of different genes from different sources.

We know how these external genes affect can be affected via the external stimuli. We can detect the gene expression pattern of different disease such as cancer on exposure to certain treatment or certain environmental toxin we can detect whether there is any changes in the gene expression. Also this dynamic changes of gene expression under different physiological condition or cellular processing can be detected different signaling pathway which can influence the gene expression can be detected different regulatory mechanisms can be detected via this gene expression profiling. Now, there is a term called comparative genomic hybridization. So, this is the whole genome which can be compared via gene expression profiling how the whole genome is stacked with the fluorescent dye.

This comparative genomic hybridization we will discuss in detail in our next class that is fluorescent in situ hybridization technique in that class we will discuss this comparative genomic hybridization in details. But for this moment you just need to know that in comparative genomic hybridization the whole genome the expression of the whole genome can be checked whether there is any loss of gene whether there is in there is any amplification aneuploidies deletion mutation all the things can be detected in comparative genomic hybridization. So, if we can see specifically what are the things

related to what are the application related to medical field for DNA microarray definitely it can be used for biomarker discovery. So, the specific gene which changes or the expression of which changes in some disease or there is some gene which is overtly expressed in some disease based on that we can say this is these can be considered as the biomarker for this disease. So, this is the basic of biomarker discovery via DNA microarray.

It is extensively used in cancer research and profiling what happens we can detect the molecular signature or the pattern of different cancer in different cancer how different specific gene expressions are changed. Then we can identify different genetic variations different expression patterns along with the cancer subtype different infectious disease monitoring can be possible via monitoring the genetic pattern and variation of different infectious agents we can classify different pathogens based on this genetic profile. Now what happens if there is any mutation in the pathogen pathogen genetic profile that can cause outbreak. So, the epidemics outbreaks those can be tracked via DNA microarray and also the very basic of antibiotic resistant or drug resistance can be studied. So, the genetic mutations based on which the antibiotic resistance appear that can be detected.

Then it is hugely used in prenatal screening. So, analysis of fetal DNA possible for different aneuploidies chromosomal anomalies and based on that based on the early detection the decisions can be taken and the genetic counseling can be done. Functional genomics are immensely helped with DNA microarray. So, systematic analysis of different gene function identification of different gene related to different biological process gene interactions between gene to gene functional roles of different genes in cellular and signaling pathways can be thoroughly studied via DNA microarray. Finally, pharmacogenomics is one evolving phase now where the comparative analysis of different genes in a person can be studied.

How a drug is reacting or changing the expression of those genes, how a drug is metabolized based on a person specific gene profile, how the bioavailability of a drug is changed based on a person's genes profile that can be studied and based on that the specific dose for a specific person can be designed. So, that is our personalized medicine. So, this is a very new concept here that is personalized medicine where a person's genetic profile is thoroughly studied and based on that the individualized medicine is designed for him. So, that is the pharmacogenomics and the role of DNA microarray, but there are few limitations of DNA microarray or as well. So, the measurement of measurement or the relative concentration detection is basically dependent on the kinetics and linearity range of the assay system.

What happens when there is high concentration of sample DNA the probes are saturated. So, they cannot give the accurate result and when there are there is very small

amount the detection range is very low in that case the linearity does not help. There is chances of cross hybridization when there is similar sequence pattern or sequence homology between two different gene one probe can detect apart from the target gene other genes can be detected that is cross hybridization. And that cross hybridization can create problem in case of reference strain bias as well. What happens when there is a reference strain based on that the probes are designed the other classes of the same species might not behave in that specific way.

So, if there is any changes in the gene sequence in other strain of that species the reference strain might not be able to hybridize. So, one is cross hybridization where the other genes of similar sequence homology can be detected and also in the reference strain bias when there is changes in the sequence the detection can be hampered. Then there is problem with detection of splice variant basically this accurate probes synthesis of accurate probes are complex in terms of when we want to detect the splice variant we need to design a probe specific to the exon or the or in between the two or in between the design the sequence should be covering the splicing region. So, that accurate designing of probes sometimes might not be possible rather it is very complex. Also this is dependent the DNA microarray is only dependent on the availability of the sequence the sequence can be designed based on the identified gene sequence.

So, the non coding region detection sometimes faces a lot of difficulty and also there is limited genome coverage because the genes are designed and those pre designed sequence we can use in designing the probe. So, let us see what we have learned in this class that is the very basic principle of DNA microarray that is nucleotide hybridization based on the complementarity of the sequence there are two types we have discussed in situ and spotted technique in in situ make DNA microarray the pre synthesize sorry in in situ microarray the probe is directly synthesized over the array plate whereas, a pre synthesized probe is spotted or deposited in case of spotted technique and the DNA microarray has immense role in diagnosis of cancer as well as personalized medicine. These are my references you can go through this these are very well discussed books and journals. Thank you. .