Course Name: Basics of Crop Breeding and Plant Biotechnology

Professor Name: Dr. Joydeep Banerjee

Department Name: Agricultural and Food Engineering

Institute Name: Indian Institute of Technology Kharagpur

Week: 11

Lecture-50: Microarray and Other Screening Methods

Welcome back. So, we will continue again. Now gradually we will move into DNA microarray. It is another important technique nowadays, microarray has been little bit obsolete, but once it came into the market it was highly popular for different analysis. So, so far whatever we have analyzed that we are trying to check whether, our transgene has been delivered through Southern, whether our transgene is expressing properly. Ok! At transcript level or at protein level through Southern, Northern and Western blot. Now through microarray basically we can check, the expression of a large number of genes together, not a single gene through Western we can check a single gene expression.

Through Northern, using a particular probe we can check single gene expression, but for microarray at a time we can check the expression of a large number of genes. Ok! Large number of proteins, basically, or their transcript, basically, the transcript is detected to microarray. Ok! So, first part of doing microarray. So, let me describe the process once again. Suppose we have generated a transgenic line, in that particular transgenic line we have over expressed a gene that is known as *OsCaM1*.

So, rice calmodulin gene has been over expressed in the transgenic lines. Ok! Suppose the transgenic line $P_1$, $P_2$ we got we have the untransformed plant also with us, through Southern, through Western we have confirmed that, in these two plants our transgene has been integrated or it is expressed properly. Ok! Suppose we have generated the tobacco over expressed lines. Now, we would like to know how other different genes are expressing in these transgenic lines, how they will, means, how the total network the transcriptional network is being changed. Because, let me give you a hint, once we have

grown this particular tobacco plants under salt stress condition this plant is surviving, this plant is also surviving while the untransformed plants cannot survive.

It means, some set of genes might be there which are showing such type of activity. Our transgene is there, but if you recall about *OsCaM1*, *CaM* stands for calcium modulated gene calcium modulated gene. Ok! Under any types of stress in a cell the calcium signaling is initiated. Ok! Once the calcium signaling is initiated, this type of CaM proteins, it interacts with a large number of proteins, large number of downstream proteins. And upon binding, they activate some gene expression, they repress the gene expression, they basically trigger some hormonal pathway, they will trigger some transcription factors in this way a lot of gene expression is finally changed, because the plant has to adapt in that particular stress condition.

So, these proteins, they will modulate itself, they will interact with other proteins and they will try to produce some defensive proteins, some SOS related proteins which will exclude the excess sodium ions from the cells. Ok! So, for analyzing which genes are basically expressed in this particular plant, or which genes have been reduced in this particular plant we have to do the microarray analysis. So, in microarray analysis basically two samples are taken together. Suppose here from we have taken this plant and we have taken this plant because we need an untransformed one, otherwise from which we will compare that this gene A C or E whichever is being overexpressed or under expressed how can I know. So, for comparison we need a control one, we need a treated one treated sample.

Now, we need to collect the sample, what sample is collected, their tissue is collected. Suppose, we have given the stress to the plant and we have taken the leaf tissue from both of this plant. From this leaf tissue, we have isolated the total RNA and thereafter we have isolated the mRNA, the mRNA isolation is done. Then we have to prepare cDNA over here as we are taking mRNA from two different types of plants one is sample, one is treatment. Ok! So, their mRNA will be, sorry, their cDNA will be labeled with different probes generally Cy-3, Cy-5 these fluorescent dyes are used for labelling the cDNA. Ok! So, suppose it has been used with a particular dye suppose Cy-3 has been used over here and Cy-5 has been used over here.

Now, this fluorescent dye once it will be excited properly it will show a particular fluorescent band, fluorescent signals during the time of emission. So, suppose this one is giving red color, while this one is giving green color, in this way different labeling could be done for treatment. One type of labeling of the cDNA for sample one type of treatment one type of labeling of the cDNA. Now, in microarray analysis we need to make a slide like things that is the most important thing. First, we have to prepare a set of genes for which we are interested to find the expressional variation. Ok! You know that in rice approximately 30,000 genes are there. Ok!

Suppose in our microarray analysis we are not interested to consider all the genes suppose we are trying to play with only 5000 genes. So, we have to prepare the probe for that you have heard about the probe preparation in Southern, Northern those things. Ok! Over here the probes are used in stationary condition. So, first of all in these particular slide-like things in this way small wells are available. Ok! In this well basically we have to put the short oligonucleotide sequence of our target gene suppose we are trying to screen these 5000 genes. Ok! So, those target gene sequence should be known to us.

So that, based on that we can design probe specific probe of smaller single stranded DNA sequence. Ok! In this way so, over here suppose this is gene A, this is probe for gene B, this is probe for gene C, probe for gene D and each thing will be available in replication also. Ok! Suppose we have used gene A, B, C, D over here B, D, C, A in this replication, we have made this particular block. Ok! Initial slide for some plants, for rice, for *Arabidopsis*, this microarray this chip, this is known as basically chip, this basic chip is available. But we can design this chip based on our requirement, we can tell that we are interested to know only the *CDPK* gene, calmodulin dependent protein kinase gene. We are interested to know all the *CBL* gene, calcineurin B like protein gene basically those genes are interactor of *OsCaM*.

So, based on that we have to design the target gene. Ok! So, once this thing is done and our labelling is completed, then we have to do the hybridization part. In this hybridization process, basically we have to mix equal amount of labelled cDNA from the sample as well as treatment. We have to mix equal amount of labelled cDNA from sample and the treatment. Then this labeled cDNA will be used for hybridization of this particular chip.

So, in microarray chip we know that 4 genes are there. Ok! And over here whatever the genes being expressed in the treated plants or in the sample its cDNA has been made. Now, I am coming to this part that is data collection and analysis. So, suppose over here gene A is available in the treatment as well as in the sample. Suppose we are talking about gene A that is available in treatment as well as sample. So, over here this will bind, because the treated is colored with red fluorescence and this will also bind right?

If both these genes are expressed in, if the gene A is expressed in both the conditions. So, it will bind over here, as well as over here we can see its binding. Suppose gene B, gene B is only expressed or up regulated in the treatment what will be the scenario? In the treatment this specific cDNA, this labeled cDNA will bind over here as well as over here. Suppose for gene C, suppose for gene C, only it is expressed in the sample, in the treatment its expression is very less. Ok! So, in sample we can see its binding in the position of C and suppose the gene D, is not at all expressed in both of these plants the treatment as well as sample.

So, it will be blank. Now, we have to detect it once this hybridization this thing we have talked about the hybridization process. Ok! Now, how can I detect it? Once we will detect it if you think about the position of A gene ok, there both of the dye will be available Cy-3 as well as Cy-5. Ok! So, in those case we can see some yellow color. Ok! In those cases where only Cy-3 is available. Ok! That is the gene B.

So, there we can see this type of color in case of gene B, where only sample tissue containing transcript is available in C. So, in C we can see only green dye. Ok! In C only we can see the green dye while in D none of them has been bound. So, in D we can see the black color. So, in this way based on the over expression if we see over expression under treatment then we will see the red color.

If we see retardation of transcript in treatment then we will see green color, because from the sample cDNA its population will be more, right? If their expression is almost similar, we will see yellow color and if it not at all bind, we can see black color. So, in this way the microarray analysis is done. So, the principle already we have discussed. So, in this way basically you can see on the microarray plate once we will check it with the scanner or

something then this type of data is available not present in the cells, it is present in both cells, present in normal cells only and its pathological cells.

In this analysis a normal cell and a pathological cell has been used for microarray analysis using different sets of genes. Ok! So, now coming to the last part of the screening of different transgenic lines. Ok! Already we have discussed about Northern blot. In Northern blot I have told that, the gel become highly fragile. Ok! And we have to use the formamide formaldehyde on the gel that is also hazardous. In addition to that we have to use a radio-labelled probe mostly in the Northern blot analysis that is also toxic or hazardous.

So, nowadays most of the laboratory they will analyze their plant through quantitative PCR. So, quantitative PCR basically it is also known as real time PCR. Ok! Or basically over we have to check the gene expression through this quantitative PCR. It is commonly known as qPCR sorry qPCR or it is a part of RT-PCR reverse transcriptase polymerase chain reaction because over here we need to start our experiment from RNA or especially from mRNA. So, if you start the experiment from mRNA then gradually, we will make cDNA, and through PCR we will check the gene expression. Ok! Then another thing we will be discussing that is the transcriptome profiling, mostly these two things will be discussed and other different methods are there like whole genome sequencing, the flow cytometry analysis ELISA. Ok!

It could be done also for analyzing the putative transformants. So, let us discuss about the quantitative PCR little bit in detail. So, suppose this is our gene of around 600 bp, we have used in the over expression construct, before that we had the promoter and after that we had the terminator. Ok! So, as we have done the over expression through Northern we are supposed to get the enhanced transcript band, but through quantitative PCR or real time PCR we can design some gene specific primers. And using those gene specific primer basically for real time PCR the primer length should be close to 20 mer, generally 20 mer primers are used the GC content should be more than 50 percent. Ok!

And generally the Tm is set for the primer that is close to 60 ºC. Ok! These things are mostly followed. Ok! So, this one was our gene, from this gene if the transcription is taken place we are supposed to get this mRNA. So, in real time PCR or quantitative real time

PCR first we have to make the cDNA, ok using oligo-dT primers or random hexamer primers. So, once the cDNA is formed then we have to use the real time primers or gene specific primers. And over here one primer can bind it will start amplification in next cycle it will be separated it will be separated and again from here we can see amplification.

In this way our target region will be amplified again and again which is done in case of PCR. So, suppose we have over expressed the *OsCaM1* gene through real time experiment, we are trying to check the expression of that particular gene in control plant and over expressed plant. In control plant suppose this is our endogenous gene expression in over expressed plants we are getting higher fold of gene expression. Ok! So, in this way, we have to tell that how much fold has been changed in the expression of a target gene? That could be detected through real time PCR or quantitative PCR. Ok! And basically Δ Δ Ct (ΔΔCt) method is used for detection of the over expression or the relative quantification of the transcript Δ Δ CT (ΔΔCt) method is used.

So, you guys can go through it little bit later on, because this is not within the scope just once the Northern become tough, then under that particular circumstances we can analyze the transcript using gene specific primer designing and real time PCR. Now, transcriptome profiling that is another thing which is highly popular nowadays suppose we have a plant that is growing in Sundarbans region. Ok! It is the coastal region while another plant of same variety is being grown in close to Himalayan region. So, definitely their gene expression will be highly changed. So, first of all if we plan an experiment for knowing which genes are over expressed in coastal region, and which genes are over expressed in hilly region, for that analysis we cannot design any primer. So, we cannot go to real time PCR analysis we cannot even plan the microarray analysis also.

So, for that thing we can directly go through transcriptome profiling. Here also we need to take the tissues then we have to make the mRNA, then using mRNA we have to make the cDNA, then cDNA library is prepared and thereafter it is cloned, it is sequenced and then different software are there, using that we can tell that in our sample which set of genes are over expressed and which set of genes are under expressed. Similarly in our treatment which set of genes are under expressed, and which set of genes are over expressed

even which genes are giving similar results, that thing also could be done from here that is known as DEG analysis, differential expression of genes differentially expressed genes. Ok! So, this analysis could be done to know which genes are differentially expressed in the coastal area and the hilly area. Ok! So, one sample we should take as a sample and another one could be considered as treatment for such type of analysis.

So, these are the briefs. Thank you. I think you have understood the topic that how we should screen different transgenic plants. Ok!