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: 07

Lecture-28: RFLP

Welcome back. So, we will continue again. So, now let us start our discussion on RFLP that is Restriction Fragment Length Polymorphism. So, it is one of the initially discovered molecular markers and it was found to be highly useful, highly beneficial and it has been used for more than 3 to 4 decades by different molecular biologist, different plant breeders. Now before discussing RFLP first, RFLP in detail first, let us discuss about the process what should be done if we have to do RFLP. So, in the name itself it is written Restriction Fragment Length Polymorphism means, some restriction digestion is occurred, is done over there.

So, you guys can… you guys can understand what are the different steps for conducting RFLP? First, we need to isolate genomic DNA, first, we need to isolate genomic DNA from our target organism. Let us assume, we are discussing about plants, first we need to isolate the genomic DNA from plants, then we have to check the quality and quantity of the DNA. We have to check the quality of the DNA by running on 0.8% agarose gel. Generally, genomic DNA is checked by running or by loading on 0.8% agarose gel, its size is larger and if we have to do RFLP analysis, generally, large amount of DNA is needed, at least 10 microgram of DNA is needed for each sample. Thereafter, we need to digest the DNA, gDNA stands for genomic DNA, then we need to digest the genomic DNA with a specific restriction enzyme, either one restriction enzyme could be used or two restriction enzymes could be used it's as per requirement. So, commonly *Eco*RI, *Hin*dIII, *Sac*I, *Bam*HI, these restriction enzymes are most commonly used, but other than

this, some other restriction enzymes could be used also or a couple of restriction enzymes also could be used that is as per requirement.

But generally, we will be discussing this course by using a particular restriction enzyme. So, after digesting the DNA suppose over here, we have isolated the genomic DNA we have loaded it on 0.8% agarose gel and it will show a particular band that is supposed to be more than 10 kb. If the DNA quality is good, we will see almost intact band, if the DNA quality is bad, then we can see some smear formation, that is bad genomic DNA. So, the genomic DNA quality is not good.

So, then what we had to do, our next step was digestion of genomic DNA with restriction enzyme. So, let us assume this is our undigested genomic DNA, and this is the digested genomic DNA. If we digest the genomic DNA, hundreds of bands will be available basically, we can see a smear on the gel using any specific enzyme *Eco*RI, *Hind*III whatever. So, our next step will be after genomic DNA digestion, we have to do electrophoresis on 0.8 % agarose gel, then we have to transfer the DNA from gel to membrane, we have to transfer the DNA from gel to membrane.

Now, what membrane I am talking about, different nitrocellulose membrane or nylon membranes are used for transferring DNA from gel. Ok! So, then I will be discussing about the steps how the DNA could be transferred from gel to membrane. So, after that, after transferring the DNA from gel to membrane, we have to fix the DNA on membrane. We have to fix the DNA so that the DNA will not be washed out easily from the membrane, it should stick to the membrane and then we have to do hybridization. So, here we have to do Southern hybridization, it was initially discovered by E. L. Southern, ok! from his name the 'Southern hybridization' term came. So, once the DNA is transferred on to the membrane, then we have to do southern hybridization using specific DNA probe. So, another new term came probe. So, probe is a known DNA sequence and through this RFLP process, we will try to find out that whether that specific sequence, specific DNA sequence is available on the genome or not that is our target. Ok! So, or if it is available on the genome, how many copies of that particular region is

available on the genome, it may be in 2 copies, it may be in 3 copies, those things could be identified through RFLP analysis.

So, the detailed procedure, means, how RFLP could be utilized for identification of different diversity study, later on, we will be discussing, but these are the steps we need to follow; first, we have to isolate the genomic DNA, its quality should be high, means, lack of garbage should be there and less smearing should be there and thereafter, we need to digest the genomic DNA with a particular restriction enzyme or a couple of restriction enzymes. Then after restriction digestion, we have to do electrophoresis on 0.8% agarose gel, then we need to transfer it, we need to transfer the DNA from gel to membrane thereafter, we have to do the hybridization process. And once the hybridization is done thereafter, we have to develop the autoradiogram. In most of the cases, in RFLP analysis, or in southern blotting, the radio-labelled probe are used, in most of the cases the radio-labelled probes are used, but nowadays, since the last one decade almost, the non-radiolabelled probes are also being used, but we will be discussing about the radio-labelled probe related problems, that is the classical RFLP procedure.

So, upto genomic DNA digestion we have mentioned, thereafter, we have to transfer this DNA from gel into membrane. Ok! So, two types of transfer mechanisms are there, one is by capillary method. By capillary method, basically a tank is used, within that tank we have to put our particular buffers, specific buffers are used then over here, we have to put our gel where the undigested DNA and different digested DNA is available. Then on the top of that gel, we have to put the membrane, we have to put our membrane over there and over the membrane we have to put a stack of papers with some weight. So that, whatever solution is available in the tank it will come through the gel into the membrane.

Basically, those solution will be absorbed by the blotting papers or stack of papers available on the top of the membrane. So, in this way, whatever DNA is available on the gel through overnight it will be transferred onto the membrane. While, another method is commonly used that is the vacuum transfer method. In vacuum transfer method, basically we have to have a vacuum blotter, there some space is available in between and

this blotter basically is attached with some pump. So here, first, we need to put our membrane, our membrane is placed initially, then over the membrane we have to put our gel, over the membrane we have to put our gel and then we will use different buffers on the top of the gel and due to the suction of this pump, suction created by this pump, those buffers will go inside through this gel into the membrane and finally, the buffers will be released from here, a pipe will be available over there.

So, in this way whatever is available on the gel that will be transferred onto the membrane. So, these are the two transfer mechanism through which we can transfer the DNA from gel to membrane. Ok! So, once the transfer process is done, then first, we need to mark the different lanes available on the gel before removing the gel from that particular location, first, we need to mark the lane otherwise we cannot tell in which lane the undigested sample was there, in which lane the digested sample was there that will be difficult to understand later on. So, once the transfer is done, we need to mark the lane on the membrane and then the first lane, last lane, in this way, we can mark it and thereafter we have to fix the DNA on membrane by UV cross-linking that is another very important part; until and unless we fix our transferred DNA on the membrane, once we will go to hybridization process, the DNA may be washed out from the membrane eventually because during hybridization process, a lot of things has to be done. I will not discuss about those things because we will mostly focus on the RFLP principle, RFLP mechanism. So, during hybridization process, some pre-hybridization buffers, different hybridization buffers, wash buffers are used again and again.

So, if the DNA is not stuck properly to the membrane, it will be difficult to carry out the full process. Ok! So, UV cross linking is must to fix the DNA on the membrane. So, once UV cross linking is done, then gradually we will go into the hybridization part, some specific DNA probe will be used to identify its position on the available DNA. So, in next, later on we will be discussing about this part. So, suppose a particular rice variety or any other variety, suppose, we are talking a particular crop where 3 chromosomes are available, ok, only 3 chromosomes are available; chromosome 1, chromosome 2 and chromosome 3, 3 chromosomes are available.

And within chromosome 1, this is just our assumption, this is for your understanding, let us assume within chromosome number 1, 1 *Eco*RI site is available over here, this is a particular variety, 1 *Eco*RI site is available over here. While in chromosome number 2, another *Eco*RI site is available over here, while in chromosome number 3, two *Eco*RI sites are available over here. Ok! And let us try to discuss about the size of the chromosomes in different positions across the *Eco*RI sites. Suppose its size is 2 kb, its size is 4 kb, its size is 7 kb, its size is 1 kb, its size is 1.5 kb, its size is 2.5 kb and its size is 0.5 kb. So, if you assume this thing, suppose, in a particular crop 3 chromosomes are there in chromosome 1, only 1 *Eco*RI site is available, in chromosome 2, 1 *Eco*RI site is available and in chromosome 3, 2 *Eco*RI sites are available. So, we have isolated the genomic DNA of this particular plant and it has been electrophoresed on the gel, this is the undigested DNA, suppose we are getting a band over here. While, once the DNA has been digested with *Eco*RI you can think that how many fragments will be generated 1, 2, 3, 4, 5, 6, 7, 7 different fragments will be generated, right!

So, the lowest one will be 0.5, next will be having 1.0, next will be having 1.5, next will be having 2.0, next will be having 2.5, next will be having 4 and next will be having 7. 1, 2, 3, 4, 5, 6, 7; 1, 2, 3, 4, 5, 6, 7 these 7 bands will be obtained from this particular genomic DNA upon *Eco*RI digestion. Now, generally a single *Eco*RI site will not be available in any chromosome, hundreds of *Eco*RI sites will be available, a number of *Eco*RI sites will be available. And whatever the bands will be getting after *Eco*RI digestion, it will look like a smear any distinct 7, 10 bands will not be available if we digest the genomic DNA. So, this is just for your understanding. So, this could be the scenario if we digest the genomic DNA of this plant with *Eco*RI. Ok!

Now, suppose I am talking about the same plant, I am just drawing it once again, this is the three chromosomes, suppose these are the three chromosomes. Now, we are doing RFLP analysis, we are trying to know about a specific DNA sequence which will be used as a probe. If you recall the RFLP procedure you can understand first DNA quality should be checked, it will be digested, then it will be transferred on to a membrane,

thereafter, we have to hybridize the membrane with a specific DNA probe. Now, suppose, a specific DNA probe we have taken from here, we are trying to find out that whether this region is available over here or not. Suppose, this probe is related to actin gene in rice and it is a touch-me-not plant.

So, we do not have any genomic information of touch-me-not plant, but we have the genomic information of rice actin gene at NCBI database we can get that information. From that, we have cloned rice actin gene and we have, somehow, amplified that particular rice actin gene and we have prepared a rice actin gene specific probe. So, if you think about actin gene, actin gene have some specific function that is a housekeeping gene, it should be active in all the tissues in all the developmental stages. Ok! So, in rice, if actin is there, the actin available in touch-me-not might be related to the similar function, they might be having the similar sequences. So, suppose, the rice actin gene has been taken and it is being used as a probe to hybridize over here.

So, this digested DNA has been transferred into membrane, in membrane we cannot see anything right! Initially, we can, we may see some EtBr stain, but later on, once the pre hybridization or hybridization process will be conducted, we will not see anything. But once the hybridization process will be done, our actin specific probe will be binding to this 7 kb region. The other bands, other bands will not be visible because our probe is supposed to bind only this DNA fragment, right, where the target probe specific region is available. While in the undigested plant, in the undigested genomic DNA, it will bind again because here, all the chromosomes are available. Within this chromosome also actin sequence will be available.

So, in autoradiogram basically, we can see this is the undigested band and this is the digested band in this way, the RFLP is done basically. Now, we will try to find out how we can identify the variability through RFLP process. So, let us think that we have the same, we have the same touch-me-not plant what we were discussing, earlier here one *Eco*RI site is available over here, one is here and two *Eco*RI sites are available over here. In chromosome 1, chromosome 2 and chromosome 3. Ok! Suppose, this is an Indian

origin touch-me-not plant and some European origin touch-me-not plant we have isolated.

Just assume in this plant over here, the *Eco*RI site is available on chromosome 1, as it is available in Indian origin. In chromosome 2, this *Eco*RI site has been mutated, this *Eco*RI site has been mutated, while in chromosome 3, the two sites are available at its specific position. So, suppose over here the sequence was GAATTC in 5' to 3' orientation, then only the *Eco*RI can cleave it suppose over here, due to some mutation, the sequence of this region has been GGAACC some modification has been occurred. Ok! So, due to this modification or suppose the sequence has been GAATTG, a single nucleotide change has been occurred due to that the restriction site has been abolished. Now, let us see what will be the digestion pattern of these two plants.

This is the Indian, this is the European. Ok! In Indian, we will be having similar type of banding pattern after digestion in both of the DNA, here has been digested with *Eco*RI restriction enzyme. So, over here we can see 7 kb bands then 4 kb bands then 2.5, 2, 1.5, 1 and 0.5. Let us assume this is 0.5, this is 1, this is 1.5, this is 2, this is 2.5, this is 4 and this is 7. These bands are available in European touch-me-not, what could be the banding pattern over here, we will get this band.

This band means two bands will be getting from chromosome 1, from chromosome 2 no *Eco*RI site is available, in chromosome 2, no *Eco*RI site is available. So, only one band will be coming from here, right, while from chromosome 3 will be having 1, 2 and 3 band means 3 and 2, 5, 6 bands will be there. So, here we will be getting 0.5 bands from here, then we will be getting 1.5 kb bands from here, we will be getting 2.5 kb bands from here. From chromosome 3, we are getting 0.5, 1.5 and 2.5 kb bands, right! From chromosome number 1, we are getting 2 kb bands, that is this one and another band of 4 kb, 4 kb bands will be getting from there, this is 4 kb, 2.5, this is 2 kb, 1.5, 0.5. So, from chromosome number 2, earlier we are getting one 7 kb band, one 1 kb band, but as here the *Eco*RI site has been mutated.

So, this digestion will not be there. So, finally, we will get a 8 kb band over here, right? 1 kb band will not be available, 7 kb band also will not be available. In spite of that, we will be getting a 8 kb band. Now, suppose once these two plant have been originated, the Indian touch-me-not or European touch-me-not, the actin gene was initially available on chromosome 2 in both the case. Now means thereafter, some mutation occurred in European origin means in the touch-me-not plant available in the genome of the, sorry, the touch-me-not plant genome available in the European continent.

So, now let us assume we have digested the genomic DNA from Indian sample and European sample and we have used this actin gene specific probe what we were using earlier. Ok! Once we will use this particular probe in the autoradiogram, what will be our scenario? So, before autoradiogram, what we have to do, we have to isolate the genomic DNA from both the plants separately, we have to digest the genomic DNA from them separately, although I am showing 7 bands over here and 6 bands over here. But if we think the real scenario, you will see smear in both the cases ok, because not a single *Eco*RI site will be available in chromosome number 1, 100s *Eco*RI sites will be available.

So, ultimately, we will be having smear over here. Once we will do the autoradiogram, once we will do the hybridization process, thereafter, we can tell that what should be the banding pattern. So, in the autoradiogram once we are using the actin specific probe, then in Indian one, in Indian touch-me-not plant, that 7 kb band will be highlighted, isn't it? Because on the chromosome number 2, the actin gene is available over here. While in case of the European sample, the 8 kb band will be highlighted, because here one *Eco*RI site has been mutated. So, we will be getting 8 kb fragment and our probe, actin gene specific probe will bind over there. So, in this way, we can distinguish these two varieties. Ok!

We can distinguish the genotypes of touch-me-not from Indian origin or European origin. So, this is the process through which we can use the RFLP to detect the polymorphism. Hence, its name is restriction fragment length polymorphism, means, it is

based on restriction enzyme digestion and based on that, we have to detect the polymorphism.