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-27: Introduction to Markers

Hello everybody. Welcome to the SWAYAM NPTEL course on Basics of Crop Breading and Plant Biotechnology. Today we will start on Types of Molecular Markers and Application of Molecular Markers. And in lecture 1 under this module, we will be discussing on Introduction to Markers RFLP and RAPD. So, these things will be covered in this particular lecture. First of all, what is marker? We know that in plant breeding different markers are used.

So, what is marker first that concept has to be clear. Then, different types of biomarkers will be discussed here. Then, gradually, we will go into the deeper part the biochemical marker, phenotypic marker and molecular marker will be discussed in brief. Then in molecular marker, mostly today, we will be discussing about the RFLP and RAPD markers.

So, what is marker? If you see this slide, a lot of things are written in different paper cutouts, but if we would like to specify something then we have to mark it. In this way, using a specific marker or using a marker, we can identify a particular individual from a group of organisms, from a group of people, a particular human being could be identified from a group of human beings, a particular plant could be identified from a group of plants. Ok! In this way, we can use this marker in plant breeding also. So, what are biomarkers? Biomarkers, 'bio' name is there means it is related to biological organisms. Ok! So, biomarkers refer to quantifiable indicators that are used to assess various biological phases or growth conditions or divergence.

Means, different biological phases could be identified, different growth conditions could be identified, divergence study could be done. Ok! Suppose, among the human beings, someone might be having some severe disease like cystic fibrosis. Some markers might be responsible for that means if a patient is having cystic fibrosis disease, that particular marker it may be some enzyme may be available in their body, it may be some specific DNA sequence might be available in their body. Ok! In this way, the biomarkers are used. Then biomarkers are measurable indications encompassing the existence or amounts of certain microbes, nutrients or genetic diversity are employed to signify the state of health functionality or quality.

Suppose, a milk product is being analyzed and using some specific marker, we can tell whether the toxic microbes have been grown over there or not. We can use some microbe specific markers. Within the milk product or within any other food products, we can calculate the amount of some nutrients, it may be some heavy metals by using some specific markers. Ok! So, that whether any contamination is available there or not. Similarly, during the study of genetic diversity, suppose, 100 rice germplasms are being maintained in a particular location.

So, among them some variations might be there. So, based on those variations, we can tell that this variety is specifically "Badshahbhog" variety. This variety is basically "Tulaipanji" variety, in this way some markers could be identified, so that, we can distinguish them. So, these are different types of markers. Ok!

Similarly, markers are classified in three types biochemical markers, morphological markers and molecular markers. Mostly, we will be discussing about molecular markers, but I will just try to mention a few things about the biochemical and morphological markers. In biochemical marker, different types of biochemical markers are available either they are isozyme based or they are based on the protein banding pattern. Later on, we will be discussing these things once again. Then in morphological marker, morphological marker from the name itself, it signifies that it depends on the phenotypic

performance, it depends on the phenotype.

Suppose in a field, different *Lathyrus* varieties are being grown. Ok! So, you know in *Lathyrus*, different types of flower colors are available blue, white, pink different flower colors are available. So, based on those particular flower color as a morphological marker we can identify that this is this particular variety. The blue flower is produced in another variety, in this way we can identify easily. Next, third one is the molecular marker.

This is the robust marker it is not mostly affected by the environment. So, the molecular marker could be broadly classified into two parts, either they are hybridization based molecular marker or they are PCR based molecular marker. Ok! Among hybridization-based marker, RFLP what we will be discussing today, then mini-satellites, micro-satellites, those markers are available. While among the PCR-based marker, different markers are there which are commonly used RAPD, Random Amplified Polymorphic DNA Markers, SSR, that is Simple Sequence Repeat Marker, then CAPS, Cleaved Amplified Polymorphic Sequences Marker, then ISSR, Inter Simple Sequence Repeat Marker and SCAR, Sequence Characterized Amplified Region Marker those markers are PCR based marker. Another hybrid marker is available that was found to be highly beneficial to the different plant biologist, different plant breeders that is AFLP Amplified Fragment Length Polymorphism.

In next class, we will be discussing about AFLP. So, let's discuss about the biochemical marker little bit in detail. First of all, the biochemical markers are related to variations in banding patterns of proteins or amino acids. Means, suppose from a particular plant we have isolated total protein samples. Ok! Among those proteins, different protein will be having different molecular weight.

So, if we run SDS-PAGE by boiling the protein with SDS, so that, all charge will be negative before loading on the gel after boiling with SDS, then we can make the charge of all the proteins negative and if we boil in SDS, the protein will be denatured. So, based on its size, based on the amino acid available in a particular protein, it will be migrated on the

gel. If its size is larger, then it will be migrated slower, if its size is smaller, it will be migrated faster. Ok! In this way, in protein gel, we can see different banding patterns of various proteins available in a particular plant. In this way, we can analyze different plant varieties, suppose in rice, different rice varieties could be used in such type of analysis and based on their protein banding pattern, we can identify some differences.

Some successful examples of biochemical markers are isozymes, allozymes or secondary metabolites. Now to identify variations in biochemical markers, gel electrophoresis is done. Ok! We know that a particular enzyme.. its name is SOD, it is an enzyme, its full form is superoxide dismutase. If a plant encounters any type of stress, it may be abiotic stress like salt stress, drought stress, heat stress or any biotic stress. So, what happens within the plant cell? The superoxide molecules are accumulated, then, this SOD, super oxide dismutase enzyme come into the picture.

If try to modify the super oxide molecules into hydrogen peroxides and subsequently by other enzymes, it is converted into water. So, this particular enzyme, it has also different isomorphs means, they are having similar function, but their size might be different. Some of the superoxide dismutase might be working in presence of iron, some of them might be working in presence of magnesium or manganese, in this way, different types of SODs are available in plant system. So, suppose, we have identified the protein samples from a tolerant variety and a susceptible variety. Suppose, it is salt tolerant and this one is salt susceptible variety, rice variety we have isolated the total protein from them.

Now, we are trying to identify some biochemical markers. Let us assume we are talking about the superoxide dismutase band, active bands available in which of them in how many numbers or what are their sizes, we can find out using biochemical markers. So, for superoxide dismutase assay, we have to do in gel SOD assay. In gel SOD assay is done as it is an enzymatic reaction. So, we cannot denature the protein by using SDS.

So, here, we have to use semi-native gel, we have to use semi-native gel. In native gel, no SDS is used, in semi native-gel within the gel only SDS could be used, but within the

sample, it is not boiled with SDS, in this way, different variations are there. So, over here we have to use semi-native gel. So, in semi-native gel, total protein could be loaded and thereafter, different staining procedures are there and based on that, suppose, we are getting this type of banding pattern in the tolerant one and this type of banding pattern in the susceptible one. Then, we can conclude that in the tolerant plant, this specific band is more SOD active band is more, the fourth one that is not available in the susceptible one.

In this way, using biochemical marker we can discriminate different varieties. So, what are the advantages and disadvantages of biochemical marker? First of all, easy process and economic, not too much costly chemicals are used over there, definitely we have to have the protein gel electrophoresis options over there, mostly, it is based on protein and it can be correlated with morphological variations means, as I have told earlier, one variety was tolerant, one was susceptible. In the tolerant one, if we can identify some specific SOD active band then we can tell that based on this band, we can identify the tolerant individuals. So, it can be correlated and it is versatile and efficient. Ok! Now what are the disadvantages of this particular biochemical markers? First of all, less in number we have shown that if we run the total protein samples of rice, we can get approximately 4 bands in 4 SOD active bands. Ok!

So, 4 bands are not a too much in number. So, it is less in number means if the number is less, then the polymorphism will be reduced also. Then, have a lower polymorphism detection rate, as I was mentioning before and they act different, differently to different plant extraction techniques, tissue types and growth phases. So, this is very important part if we isolate the total protein samples from rice leaf or if we isolate the total protein samples from rice root, we may see difference in SOD active banding pattern because in different tissue, different types of proteins might be functional because tissue specific expressions are available in most of the organisms. Now, let us discuss about the morphological markers.

So, from the name itself, it is known that it is based on morphs that is the observation from outside the phenotype. Ok! So, morphological markers are visual indicators that

distinguish phenotypic characteristics including the color, shape, size of flowers, seeds or leaves. Over here, you have seen that two types of flowers are available. Ok! One is dark red in color, another one is relatively pink in color. Ok! This is the morphological difference based on flower color as I was mentioning earlier, in *Lathyrus*, different colored flowers are available it may be blue, white, pink.

So, based on that we can distinguish the genotype. Similarly, if you think about some pulse crop like lentil, over there different sized lentil seeds are available. Ok! Some seeds are macrospermic, macrosperma means their seed size is larger, while, some seeds are microsperma, their seed size are smaller. In this way, we can distinguish the genotypes based on morphological markers. Now, what are the advantages and disadvantages of morphological markers? First of all, easily detected without the need of specialized instruments, no specialized instruments are needed.

If we go to the field, then based on the morphology we can distinguish it. Suppose 100 rice germplasm are being grown, out of them may be 3 to 4 might be having the awns. So, it is a morphological marker based on that, we can discriminate different varieties then no specialized biochemical or molecular techniques are needed. So, it is, I mean very cost effective, means very less amount of cost is needed, just you have to monitor it by yourself or by some labors. Ok! So, no costly chemicals are used over here.

Then, coming to the disadvantages; first, limited quantity, here also the number of variants is less. If you think about the rice awn, either awn will be available or it will not be available or maybe in some cases, the awn size could be larger or smaller. So, 2 to 3 differences are there. Ok! If you think about the flower color, there also 3 to 4 types of difference might be available in nature. Then, it is dependent to the plant growth stages.

Suppose, anyone has identified a morphological marker that is awn in rice. So, based on this particular morphological marker we can screen a particular variety once it is in flowering stage, means, once it is in grain formation stage, grain filling stage, right! Because at initial stage, we cannot tell that which variety is which one, which one will be

having awn or which one will not have awn, we cannot identify it. So, it is depended on plant growth stages and it is sensitive to diverse environmental factors. Ok! Suppose, due to diverse environmental factors, what happens maybe, 100 rice germplasms were being grown and due to some salt stress or due to some microbial infection in most of the plants, the seed production has been hampered.

So, we cannot identify that which plant was having awn or which plant was without having awn, in this way we cannot discriminate them. Ok! It is depended on the environmental factors. So, now coming to the molecular markers. Mostly, our discussion will be on molecular markers. First of all, molecular markers refer to specific nucleotide sequence that can be examined for polymorphic variations among different individuals.

So, few things are there, specific nucleotide sequence. So, in biochemical marker it was based on protein or isozymes those things. In morphological markers, it was based on phenotype, while in molecular marker, it is based on the DNA sequence or RNA sequence means it is based on the nucleotide sequence and that can be examined for polymorphic variations, means we need to check whether any variation is available at DNA or RNA level among different individuals. So, polymorphisms are primarily driven by process like insertion, deletion, point mutations, duplication and translocation. We know that these different things occur: deletion, duplication, translocation, point mutation during DNA replication… these things are taken place.

So, due to this type of genetic modification, polymorphisms are aroused within a population. So, what are the features of ideal molecular marker ok? An ideal molecular marker should be having these features. First of all, exhibit co-dominance. What is co-dominance? Later on, I will describe, but co-dominance is the thing those markers which can distinguish the homozygous dominant and heterozygous, means capital A capital A and capital A small a, this could be distinguished by a co-dominant marker. Ok! While, by dominant marker these things could not be distinguished by dominant marker, it will be almost similar by co-dominant marker we can distinguish them.

Then, it poses an even distribution across the genome. Ok! Let us assume, this is a particular chromosome, suppose this is chromosome number 7. Once we are discussing about the biochemical marker, suppose, some SOD active gene was available over here, another SOD active gene was available over here. So, from those two genes we are getting different bands. If you think about the morphological markers, suppose a flower color gene was available over here ok! based on that, we can see the flower color.

But if you think about molecular marker, it should be widely distributed throughout the chromosome. So that, total chromosomal part could be analyzed that which part is coming from which parent, it could be analyzed easily. Ok! Now, it demonstrates good reproducibility that is the main important factor about molecular markers that it is, in most of the cases, it is having good reproducibility. As it is dependent on the DNA sequence or RNA sequence, due to environmental change we may see variation in the phenotype. A tall plant may not show taller growth due to lack of availability of proper nutrients due to unavailability of irrigation, but its DNA sequence will be same.

So, this means using this molecular marker, we can get reproducibility and possess the capability to identify a greater degree of variation. As this type of markers are available throughout the genome, so large number of variations could be detected. Now, what are the advantages of molecular markers? First of all, abundance, I have mentioned earlier it might be available throughout the chromosome within a genome, means in different chromosomes within a genome, different molecular markers might be available. So, we can explore it based on our requirement. Then, it may show co-dominance so that, the homozygous dominant and heterozygous could be identified easily.

The F_1 hybrid, whether F_1 hybrid after crossing capital A capital A with small a small a suppose, we are getting this one. Ok! Suppose, this is tall plant this one is dwarf plant and F_1 we got tall plant. So, whatever the plants are being grown as F_1, we cannot tell whether they are true F_1 or they are just the selfed seed of this parent, we cannot tell until and unless we can distinguish capital A capital A and capital A small a. So, using this type of molecular markers we can distinguish it, some of the molecular markers we will be

discussing later on. Then phenotypic neutrality, then absence of epistasis, suppose, some morphological features are controlled by two or more genes. Ok!

Suppose, this one was the substrate by the action of a particular enzyme, we can get an intermediate and by the action of another enzyme, we can get the final color. Suppose, it is a flower color. Ok! If E_2, if this enzyme is non-functional, we can see white color in flower, if this one is non functional, there also we can see white color flower, but using molecular marker we can discriminate these two, whether there is the variation within E_1 or there is variation in E_2. So, the epistasis things will not work over here. Ok! Then unlimited in number, means, we can use different numbers of molecular markers as per our requirement, then it is easier to analyze and not influenced by environment, plant developmental stages or tissues.

Suppose, as I was discussing about the awn development in rice, in seedling stage if we can identify some molecular marker which is associated with the awn development, then using those particular marker even at seedling stage, we can tell that variety A may be producing awn while variety B may not produce awn. If we can identify molecular markers which is specific for that particular trait.