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: 09

Lecture-41: Unidirectional and Bidirectional Promoter, Numerical Question on Promoter Analysis

Welcome back, so we will continue again. Now we will discuss about the unidirectional and bidirectional promoter. In unidirectional promoter we have already told that, the within the eukaryotic promoter the TATA box region is available, the CAAT box region is available. So, initially the RNA polymerase binds over there in the TATA box, CAAT box those regions. Then the DNA available in the further upstream region that is folded, that DNA is folded. Suppose, some upstream sequences are available over here, in those sequences different transcription factor binds and those transcription factors basically interact with the RNA polymerase.

And eventually, they will initiate the process of transcription, they will fasten the process. So, over here our gene is available, once this mechanism is completed then the transcription is initiated. In addition to that within a eukaryotic gene you know that introns are there, in some cases the intron mediated gene expression is also found for some genes. For some genes within the terminator region also some sequences were found to be available that play role in gene expression.

So, terminator also helps in gene regulation. So, this is the thing about a unidirectional promoter. Now in case of bidirectional promoter from the name itself it signifies that this promoter will be active in both the directions. Ok! So, in case of eukaryotes it has been found, let us assume it is a chromosome. Over here a particular gene is available, over here another gene is available, then over here a gene is available in this orientation, over

here another gene is expressed in this orientation.

In this way within chromosome different genes are there, some are expressed in this orientation, some genes are expressed in this orientation, in this way different things are available within the eukaryotic chromosome different chromosomes. So, for each and every gene it is supposed to have a particular promoter region, that will control that particular gene. If you think about this promoter this gene, this will be the promoter while for this gene, this will be a promoter region. So, if the span of this promoter is less than 1 kb, then generally it is considered as a bidirectional promoter, means a region, a promoter region which can control the expression of two different genes in opposite orientations. If the length of this intergenic region, if the length of this intergenic region is less than 1000 bp, then in most of the cases it is a bidirectional promoter which can control the expression of two genes, in two different orientations.

But in some cases it has been found that in eukaryotes, some bidirectional promoters are there and the size of those bidirectional promoters are more than 1 kb, more than 1000 bp also. It can be even up to 2 kb, 2000 bp, if we see a 2000 bp span in between two divergent orientation or divergently oriented genes, then we can think that, that region might be working as a bidirectional promoter. So, how gene is transcribed from a bidirectional promoter? So, suppose it is a bidirectional promoter region and, in this side a particular gene is supposed to be expressed, let us assume its sequence is ATG… in this strand, its TAC..... in this side, if it has to express this gene also, then for this gene it will be the 5′ to 3′ direction.

Here we have to, have the ATG sequence or earlier sequence the UTR sequence, some part of the TATA box and, those things might be there, but basically the 5′ to 3′ the coding sequence will be in this strand and, other bases will be there and over here it will be TAC. So, once the transcription will be taken place from this bidirectional promoter from here, the mRNA sequence will be AUG.... basically this strand mRNA will be like this, other than T here will be uracil (U) residues. While, from for this gene let us assume this is gene 1 and this is gene 2, under a particular bidirectional promoter. So, for gene 2

the mRNA sequence will be AUG 5′ to 3′ end in this way different mRNA bases will be there like this strand. Ok! In spite of T there will be uracil (U) residues. So, in this way 2 genes will be transcribed and eventually it will be translated. Ok!

So, this is the case in which unidirectional and bidirectional promoters are available. Now, let us discuss little bit about the preparation of promoter reporter gene construct. Now how we need to make such type of gene constructs? Earlier we have discussed about a promoter region, in promoter region the TATA box binding factors are there, binding regions are there, then CAAT box regions are there, some upstream regions are there, other different transcription factor binding sites are there. So, for characterizing a promoter that how much region it may be? 500 bp from the transcription start site means, where from the gene is transcribed that is known as transcription start site. So, before the transcription start site basically the promoter is initiated before that, it will be -1, -2 in this way different upstream regions are there.

So, up to which part is considered as a promoter of a particular gene, if we have to find out then we need to make such type of construct. So, in this way different types of constructs are made suppose we have such type of promoter. Ok! Then by keeping, this is the 5′ end and this is the 3′ end of the promoter. So, by keeping the 3′ end intact we can make different 5′ deletion construct. Ok! In the 5′ deletion construct in the first construct suppose, we will be having up to this part where upstream element is there where CAAT box is there and, TATA box binding region is there. This is supposing the construct number 2, this one was the original construct.

Then we can make another deletion construct, using only this part where only CAAT box will be there and, TATA box binding region will be there. Then we can make fourth construct, only this part will be there TATA box binding region will be there. Similarly, we can make some 3′ deletion constructs also. In case of 3′ deletion construct basically the 3′ part is gradually deleted. So, here this is our first construct the original one, the second construct will be over here, the transcription factor binding sites over here, the upstream element and over here the CAAT site is available.

Suppose TATA box binding region has been deleted then, we can make construct 3 where, sorry its size will be smaller. We can make this type of construct only the TATA box sorry, only the transcription factor binding sites and the upstream element binding sites will be there. We can make another deletion construct, another 3′ deletion construct only the transcription factor binding sites might be there. Ok! So, in this way after preparing the deletion constructs, we can put some reporter gene over here. In case of reporter gene, different genes could be used we can use *GUS* gene, if you have to see the plant tissue specific expression or you can use *GFP* gene or *RFP* gene for detection of those protein through fluorescence microscopy.

So, in this way the reporter gene, a particular reporter gene will be placed at the end of all this construct and finally, we have to make this type of graph. Suppose this is the construct 1, this is the construct 2, this is the construct 3 and, this is the construct 4 in this case. And suppose, in all this case we have put *GFP* gene, we are detecting the fluorescence activity. Suppose in case of construct 1, 100% activity was observed. In case of construct 2 the activity was 150%, in case of construct 3 the activity was 50%, while in case of construct 4 the activity was again 50%.

So, what is meant by that? From this analysis we can tell that this construct is having 100% efficiency. While in second case, where only the transcription factor binding sites have been deleted there, we are getting more activity means that transcription factor is basically retarding the expression of that gene. It is reducing the transcript expression or it is reducing the transcript production as well as protein production. Ok! So, in this way we can determine that, up to which region could be considered as best suitable promoter of that particular gene. So, clone 2 that is this part could be considered as best promoter for this gene.

And similarly we have to analyze the 3′ deletion also, suppose this is the clone 1 here, we are getting 100% activity. In clone 2, clone 3 and clone 4, in all this case we are getting 20% activity. It means once we have deleted the TATA box binding region then,

our activity has been significantly reduced. So, it is very important for this particular promoter. So, similarly we can make some internal deletion also, like a construct would be made where the transcription factor binding sites will be there, the upstream element will be there and the TATA box binding region will be there.

While the CAAT box region has been omitted in this way, some internal deletion also could be made and thereafter, we have to fuse our reporter gene to check the subsequent expression, that which part is more important for this particular promoter. Now we will start some discussion, some numerical questions on promoter analysis. So, whatever we are discussing earlier. So, these are some examples, some questions, some of it have come in different competitive exams like CSIR or UGC NET exams. So, let us see what is there, in order to identify the regulatory regions of a novel promoter sequence.

Suppose, we have a novel promoter sequence shown below 450 bp deletion constructs have been made. Here we are talking about internal deletion, internal deletion of this promoter has been made and, transiently analyzed using some reporter gene. Here the luciferase reporter system has been used for transient analysis as indicated below in box, A to D different deletions have been made over here, in this construct A region has been deleted, in this construct the B region has been deleted, over here C part has been deleted and, D part has been deleted. If you see the A part is available in between the transcription start site and -150 region, while B is available in between -150 to -300, C is available -300 to -450 and D is available in between -450 to -600. So, 150 bp span is there, in this way different internal deletion has been done and after transfection the observed level of promoter activity of each construct were indicated in figure.

Identify the best correct combination of regions in the options given below, later on I will show that indicate the presence of a positive and negative regulatory elements, respectively. Now over here if we see once the A region has been deleted. So, our activity drastically reduced. Ok! It means the A region is very important, right? If we delete that region our activity will be reduced. So, definitely some activator might be available in A region, if we see the construct 2. Once the region B is only deleted, our

activity become 96% means it will be minutely reduced not significantly reduced.

So, we cannot tell surely that whether an activator is available over there or not. Let us see, the region C if we delete the region C, our activity is increased from the full-length promoter our activity is increasing, if we delete this part, it means some harmful elements are there, some repressor elements are there, once we are reducing it once we are removing it, our activity is increased. So, this region might be having some repressor the region C, what is there in D? In D again our activity is showing 105% compared to the parental one, in each of the cases we have to compare it with this one, right? Because, these are separate deletion constructs. So, here it is close to 100% not significantly high. So, let us see what are the points available.

So, it has been asked that positive and negative regulatory elements, respectively. So, options are there, B and D, A and C, A and D, A and B. So, if you see the last slide, found the activator in region A and repressor in region C means, the positive and negative regulatory elements are there. So, the option will be A and C. Let us see the next question here, the deletion analysis of a promoter region of a gene, was carried out to identify the regulatory elements in it.

Here again the different deletion analysis has been done in the figure below, the filled box denotes, the areas of deletion from the upstream region means all internal deletion have been made. In this construct extremely part means, extremely 5′ end -250 to -300 has been deleted, the activity is same. Over here the -200 to -300 region deleted, here we are getting activity 185% means, if it is considered as 100% those are just arbitrary numbers. Ok! So, the filled boxes denote the areas of deletion from the upstream region and, the observed activities in arbitrary units of the promoter are shown below. If we see over here in the third construct here also it is showing activity of 185%.

If we see the fourth construct fourth deletion construct here the activity is 60%, over here activity is 60% and, in this deletion the activity is 0. Now from this analysis, first we need to find out if we compare the whole promoter region and, the first construct what

will be happening, is there any variation, no both are 100%. So, over here no important thing is available. Now let us think about the next construct, in this case. Ok! Once this deletion has been done this part has been deleted, we are getting our activity is increasing tremendously.

Activity is increasing upon deletion means definitely some repressor is there. So, what could be the repressor region? It might be in between -200 to -250, -200 to -250. Let us see the next construct here -200 to -250 deletion is there, in addition to that some other deletion is there. Here also we are getting activity 185%, no further change means this region was playing an important role. Once it is deleted our activity was getting more.

Now coming to next part for this deletion, we are getting activity 60%, this part is common for this part we may not see something, but for this region our activity is highly reduced, it means some activator is there. Once we are deleting it, we are getting more activity, right? So, this region will be in between -100 to -150, that is this region some activator might be there. Then if we see the next construct here, also deletion is their activity is same. Ok! So, this activator is playing some role and let us think about the last construct here, the 3′ end of the promoter has been deleted means within the -50 region to transcription start site definitely this will be the transcription start site definitely some more activator will be available.

The TATA box binding elements, those things might be available. Let us see the points, based on the observations the following statements were made, the region between -100 and -50 a positive regulatory element, is there. The region -200 and -250 a negative regulatory element, is there, region between -150 and 200 a positive regulatory element, is there. So, let us see what we found -200 to 250 the replace that means, negative regulatory region -100 to 150 positive regulatory region means, 100 to 150 is not available over here. While -200 to 250 is available over there, -200 to 250 is available over there, it was a repressor means negative regulatory region. Ok!

In between 150 to -200, 150 to 200 was it important it was not at all important. So, this

could be rejected -100 to -50, -100 to -50 this region this is also of not much importance. So, the answer will be B only. So, it is another question another numerical related to the promoter deletion construct study, or promoter analysis. So, here a question is given a promoter deletion study was done, in order to determine the binding site for a transcription factor on the promoter which is either activated or retarded on treatment with drug X.

The following constructs were made. Ok! So, these are the different constructs this is construct -1850, this is construct number -1250. So, this one will be construct number -850 and this one is the construct number -450. So, these 4 deletion constructs were made what type of deletion here, 5′ deletion has been done. So, different 5′ deletion constructs have been made and, with all of this promoter construct a luciferase gene, a reporter gene has been added over there just after the promoter. So, this is the TSS the transcription start site of luciferase gene, it is available in all the constructs.

So, once these constructs were made, this different 5′ deletion constructs were made, thereafter, it was treated with a particular drug X and we measured the luciferase activity. So, if drug X binds with specific transcription factor, then transcription factor binding site or specific transcription factor, then ultimately it can modulate the gene expression. So, these are the results, what are the luciferase activity observed in different constructs? Ok! So, in control condition in most of the cases we are getting 100% or an arbitrary value of 100 in all the constructs, but once we have treated it with drug X then differential expression of luciferase activity was detected. In first case the activity was almost 180% if you consider is at 100% then 180% expression was detected upon exposure to drug X.

In second construct it was close to 140%, in third construct it was close to only 40 % of the control condition and over here also we got almost 40% of the control condition. Now from this data what can we infer, that in which region the activator or repressor is present? So, let us try to find out. So, if you think about the first construct, in first construct under control condition we are getting 100% activity. Once we are treating

with drug X its activity is increasing.

It means over here some sequences are available that is helping in the activation of this, right? Means some activator might be available. If you think about the construct 2 here, some regions might be there which is showing relatively higher activity compared to the control 1, but if you compare with the previous construct in between these two regions then its overall activity has been reduced right means that activator was available over here has been omitted in this construct. So, its overall activation has been reduced. So, now let us see the third construct over here in -850 construct again, deletion has been done and we can see a drastic retardation of the activity. It means some activator might be available over here also once we are removing this part then activity is being reduced isn't it.

Once we are going to fourth construct that is -450 what we are observing upon exposure to drug X the activity is same. It means neither activator nor repressor is available in this region because in -850 and -450 upon exposure to drug X, we are getting 40% activity compared to the control condition, right? So, now let us see what are the probable outcomes are there? So, the following statements can be made, first of all region between -1850 and -1250 contains binding site for an activator. Ok! Next one in between -850 and -1250 contains binding site of repressor, then region between -850 and 450 contains binding site for a repressor and, third one region between -1250 and -850 contains a binding site for an activator.

So, 4 options are given out of this maybe 1 or 2 could be, right? So, let us think it once again I am going to the previous slide. If you think -1850 and -1250 indeed some activator was there. So, contains binding site for activator these options will be right think about -850 and -1250, -850 and -1250 they are telling about a repressor, but it is not like that again some activator was there.

So, the option 2 will be wrong. Let us go to the third one the region -850 and -450 contains a binding site for repressor. If you see the -850 activity and -450 activity, they

are same upon exposure to drug X. So, neither activator nor repressor is available over there. So, option 3 will be also not right, then think about the option D between -1250 and -850 contains binding site for activator. We have mentioned it over here active some activator regions are available.

So, the option A and D is supposed to be right. So, now you may get this type of questions in competitive exams that which of the following options represents the correct interpretation of the data both A and B, both A and C, both A and D, and both B and C. So, both A and D will be right. Ok! So, thank you I hope you have enjoyed the promoter characterization related part under this particular NPTEL course. Ok! Thank you.