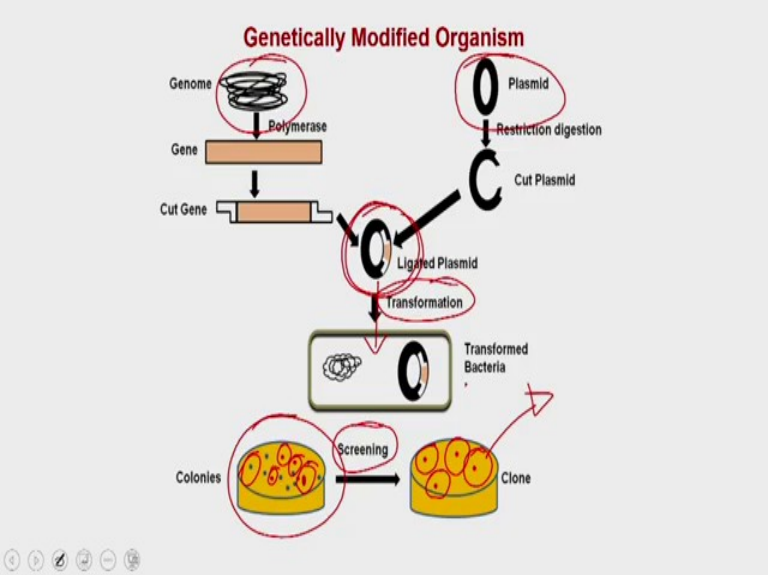
Genetic Engineering: Theory and Applications
Professor Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Module 5
Recombinant DNA Technology (Part I)
Lecture 15
Screening of Recombinant Clones

Hello everybody, this is Dr. Vishal Trivedi from department of biosciences and bioengineering IIT Guwahati.

(Refer Slide Time: 00:38)

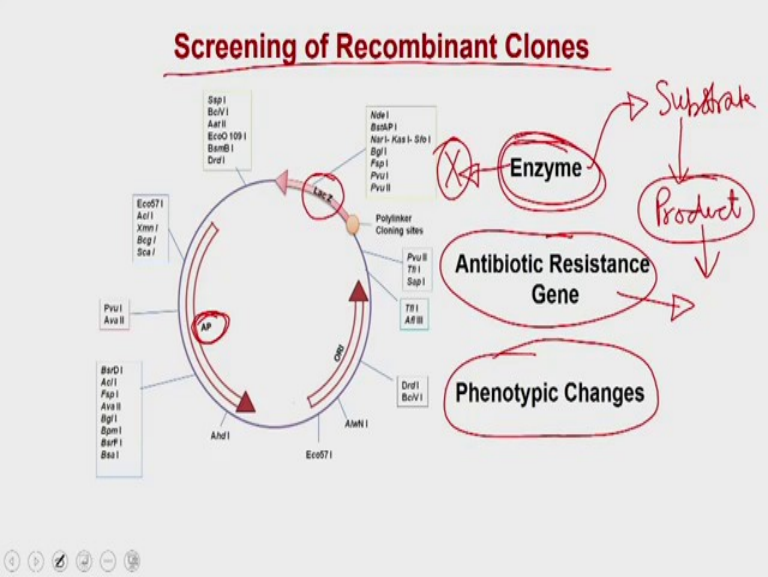


So far what we have discussed, we have discussed about the host cells, we have discussed about different types of transforming agents, we have discussed about different types of machinery which is, we are using and then recently what we have discussed, we have discussed about how to deliver this recombinant DNA into the cell. So once you deliver the recombinant DNA into the cell the next step is you have to select the cells which are going to receive the DNA and the cell which have not received the DNA.

Which means if you are taking the colonies the full colony and then suppose you have done the transfection you have to do a screening to know the cells which are going to get your plasmid or your recombinant DNA because these are the colonies which we are looking for and these are the

colonies which we are interested to grow and further use for downstream applications, such as we can use them for protein productions or you can use them for any type of study which is transcription studies or translation studies. So for the screening you have to have the exclusive criteria which you can use to exclude these molecules from these molecules.

(Refer Slide Time: 02:05)



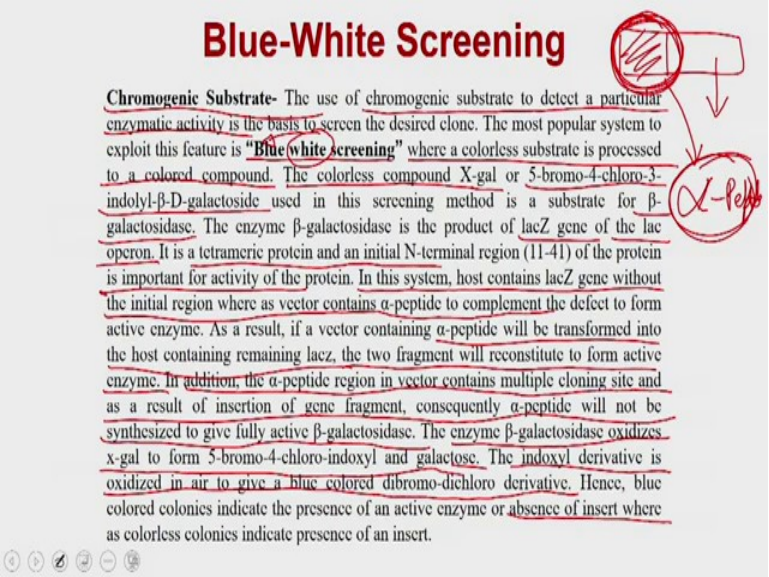
Now if you see a typical plasmid, this is the pUC19. What you will see is that, you have a ampicillin resistance gene, you have an enzyme which is called as the lac Z and then, so these so that is why if you take any plasmids where you have reduced the recombinant DNA you are either going to have the enzyme which you can exploit for two purposes, either you can use a substrate and once you use the substrate, the substrate may get converted into the product by this particular enzyme and this product may be of a unique color or some kind of, it may give you some kind of indications that enzyme is active.

In some cases you may do exactly reverse that you make this enzyme inactive. In those cases the exactly the reverse is going to be happened. So the (molecule) the recombinant DNA is not going to show you this activity. Then you have the antibiotic resistance genes, for example in the case of pUC19 you have the ampicillin resistance so you can use this ampicillin resistance so the if you have generated the recombinant DNA, the recombinant DNA is going to have the ampicillin resistance and that is how if you put these recombinant DNA along with the host on to ampicillin

containing media the cells which have taken up the plasmid will grow whereas the other cells will not grow into go.

The third criteria is that phenotypic changes, for example when you transfect the recombinant DNA into the host, it may induce some kind of typical phenotypic changes into the cell which will be clearly visible. Okay, for example, if the cells become fluorescent or cells become blue color or some cells become red color. So you know that the red cells or the blue cells or the cells which are showing the green fluorescence are actually the cells which have received the DNA, all other cells are not have received the DNA so because, so these are called as the phenotypic changes into the particular type of cell. So all these strategies can be used to screen the recombinant clone after the (translation) transformation or the transfection. So once you deliver the DNA, you can use all these approaches to screen the molecules.

(Refer Slide Time: 04:49)



So first approach is that where you can use the chromogenic substrate of a typical enzyme which you are going to use. So one of the classical screening method which we have discussed in the past also is called Blue-White screening, where a chromogenic substrate is used to detect a particular enzyme activity where a colorless substrate is (progressed) processed into a colored compound which means a white color compound is getting converted into a blue color compound.

This colorless compound which we use for the blue-white screening is called as the X-gal or 5-bromo-4-chloro-3-indolyl-beta-galactosidase and this is the substrate for beta-galactosidase. Beta-galactosidase is an enzyme which is actually be a product of the lacZ gene and from the lac operon, so lac operon is the lactose operon and the beta galactosidase is a tetrameric protein and the initial N-terminal region.

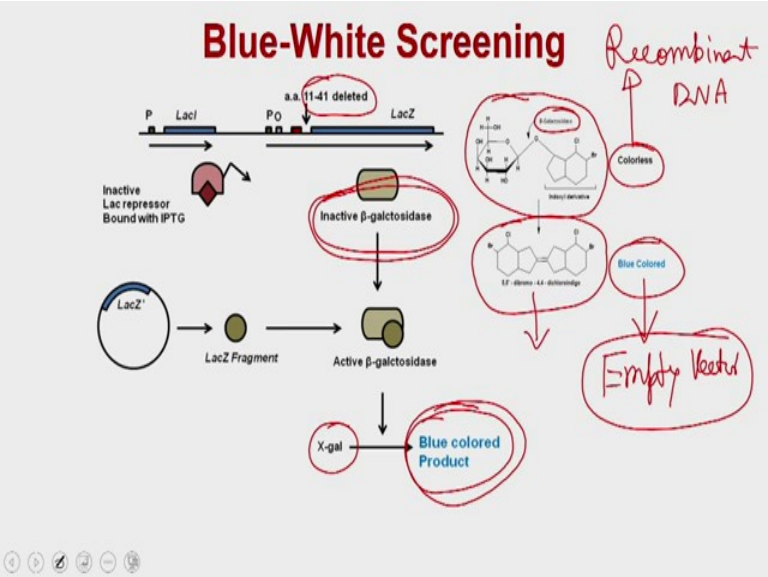
So if this is the protein the initial region is important for activity of this particular protein. So what you have to do is, in this particular system, what you have to do is you take the host which lacks the lac Z gene without the initial region, whereas the vector contains…. So this particular region, which is 11 to 41 is called as the alpha peptide or the alpha region.

So if you remove this particular region, so if you have a bacterial strain which does not contain this particular region, then the lac Z what you have is not going to be functional. So in that case, what you do is you take, you design a vector which actually going to express this missing portion, which is called as the alpha peptide and once you transform this vector containing alpha peptide into the host containing remaining lac Z the two segment are going to come together to form the active enzyme.

Which means you are re-naturing the, recovering the activity of the enzyme simply by having one component onto the vector the rest component onto the host cells. Now once the two fragments are going to form and will come together the enzyme is going to be functionally active which means it is going to process the colorless product X-gal into a blue colored product.

Now, in addition the alpha peptide region is a multiple cloning site and as a result the insertion of gene, the alpha peptide will not be able to synthesize to give fully active the enzyme beta-galactosidase oxidizes X-gal to form 5-bromo-4-chloro indoxyl and galactosidase. The indoxyl derivative is oxidized in air to give the blue color dibromo-dichloro derivative. Hence the blue color containing indicate an active enzyme and absence of insert whereas the colorless colonies indicate presence of an insert.

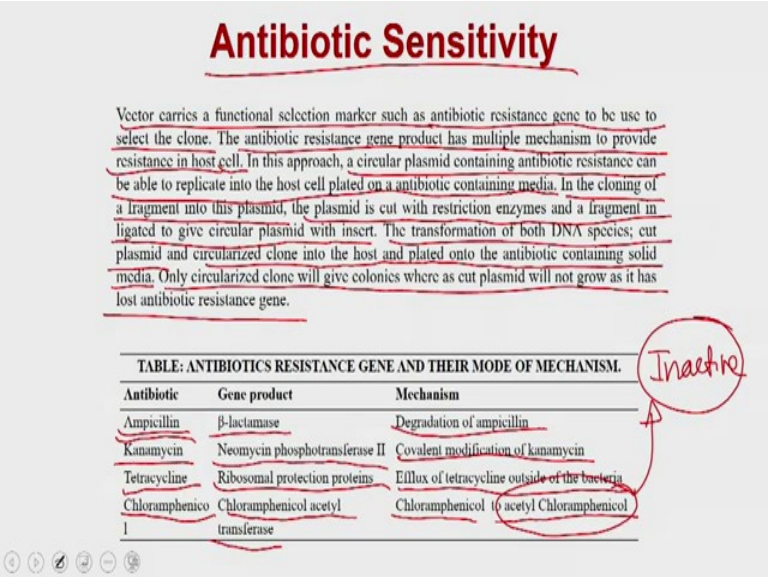
(Refer Slide Time: 08:24)



So what it is saying is that the, you have this 11 to 41 part missing and this 11 to 41 part can be supplied by the vector. So the in a typical scenario when you, if you transform the vector alone the vector is going to give you the functionally active protein and that is how the X-gal is going to be processed to a blue color compound. Whereas if you clone a gene into the alpha peptide in that case what will happen is, the alpha peptide is not going to complement the lac Z and because of that if the recombinant DNA is been formed and it is been transformed into the host the lac Z is not going to be lac Z is going to be, going to produce a inactive beta-galactosidase and once it is producing the inactive beta-galactosidase, it is not going to process the X-gal to a blue color compound and as a result of that you are going to see the colorless colonies.

So the colorless colonies are the colonies which are going to have the recombinant DNA. Whereas the blue colored colonies are the colonies which are going to have only the empty vector because the empty vector has the functionally active alpha peptide and if you stain them. if you stain the cells along with the X-gal you will see the blue as well as the white color colonies. The white color colonies are the (recomm) the colonies which contains the recombinant DNA whereas the blue color colonies are containing the empty vector or non-recombinant DNA. How the blue color is formed? This is the beta gal when it is been processed by the beta galactosidase, it is been converted into a product which is called as the 5, 5 dibromo 4, 4 dichloro indigo and this is a blue color product.

(Refer Slide Time: 10:52)



Now let us move on to the next thing, next is called Antibiotic Sensitivity. So in the antibiotic sensitivity what you have is that the vector carries a functional is functional selection markers such as the antibiotic resistance genes and that actually can be used to select the clone. The antibiotic resistance genes normally produced, has the product which has a multiple mechanism to provide the resistance in a host cells. You can see you have the kanamycin.

The kanamycin resistance gene is going to produce a protein which is called as the neomycin phosphotransferase-2 and what the neomycin phosphotransferase-2 is going to do, it is going to make the covalent modifications of kanamycin which means it is going to phosphorylate the kanamycin and once it is going to phosphorylate the kanamycin, the kanamycin will not going to perform its anti-microbial activity.

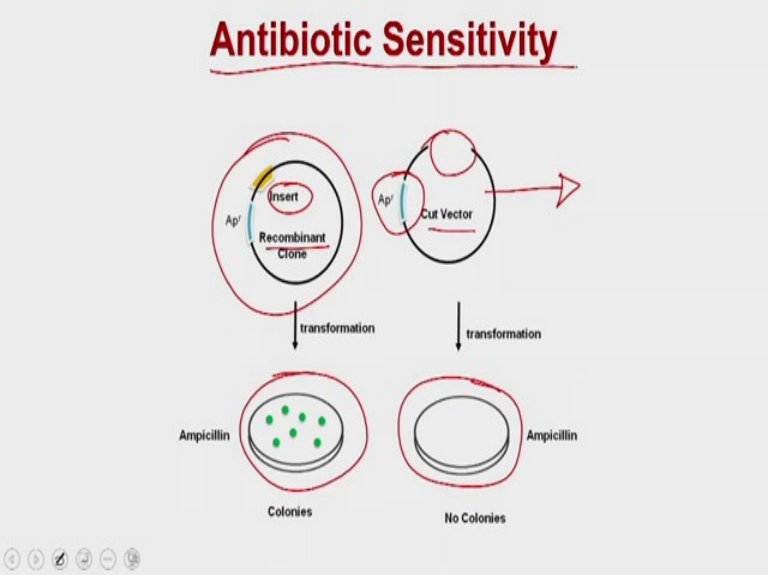
Similarly, you have the tetracycline, the tetracycline resistance gene is going to produce the several ribosomal protection proteins and that actually will contribute into the flux of tetracycline, which means it is going to throw the tetracycline out of the bacteria and that is how the tetracycline level is going to be very-very low into the bacteria and that is how the bacteria is going to acquire the resistance against the tetracycline.

Similarly you have the chloramphenicol, the chloramphenicol is chloramphenicol resistance gene is going to produce a protein which is called as the chloramphenicol acyl transferase and that

actually is converting the chloramphenicol to acetyl chloramphenicol and acetyl chloramphenicol is inactive compared to the chloramphenicol and that is how the it is going to console the resistance to the that particular bacteria. So in this particular approach a circular plasmid containing antibiotic resistance gene can be able to replicate into the host cells plated onto a antibiotic containing media.

Which means suppose you have the ampicillin resistance gene onto the plasmid, you can plate these cells which contains the ampicillin resistance gene onto a ampicillin containing media in the cloning of a fragment into the plasmid, the plasmid is cut with the restriction enzyme and a fragment in the is ligated to give a circular plasmid with insert, the transformation of both the DNA species cut plasmid and the circular clone into the host and plated onto the antibiotic containing solution media. Only the circularized clone will give you the colonies whereas the cut plasmid will not grow as it has lost the antibiotic resistance genes.

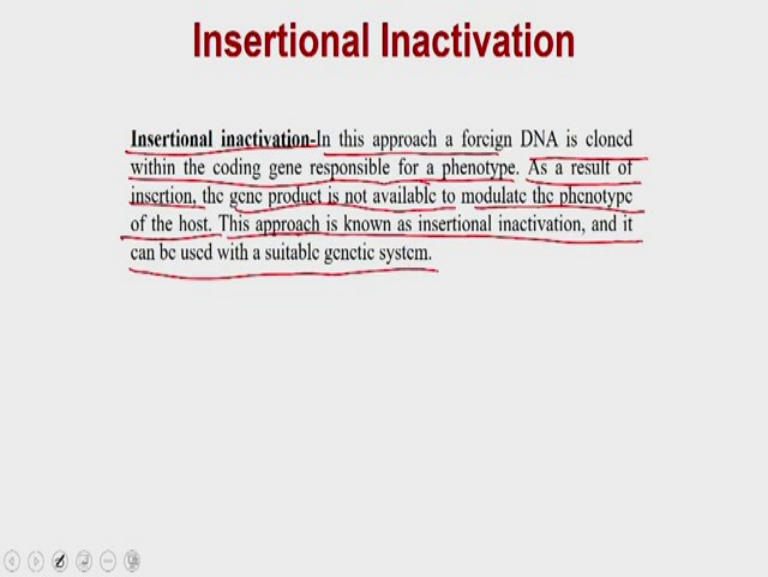
(Refer Slide Time: 13:48)



Which means if you have the insert, so what you do is you put the insert into the plasmid and go for ligation reactions. It will give you the circular plasmid whereas in the other case you are going to have the linear plasmids which still contain the ampicillin resistance because it is a linear plasmid. It is not going to express the ampicillin resistance gene and that is how this will not going to confer any resistance.

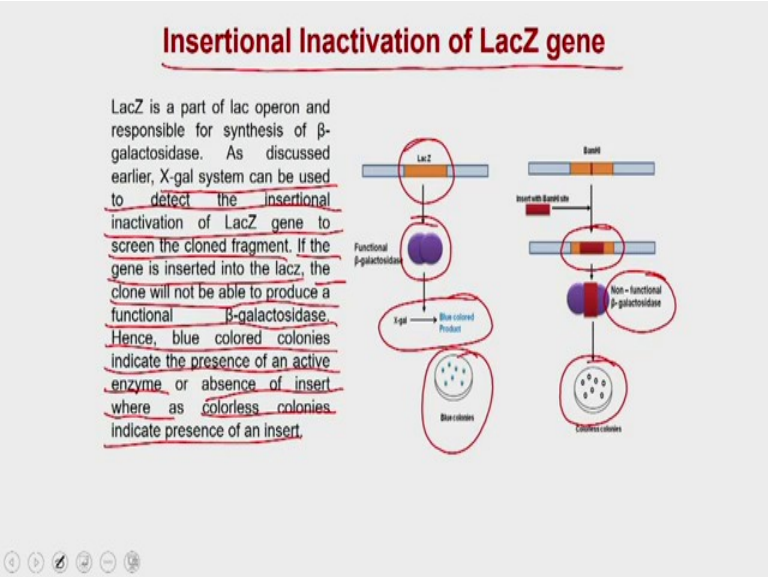
So when you do the transformation of the cut vector as well as the recombinant clone, the recombinant clone will give you the colonies because on the ampicillin containing media whereas the cut vector is going to give you the no colonies and that is how you can use the antibiotic sensitivity as a criteria to select the recombinant clones.

(Refer Slide Time: 14:43)



Now the second approach is Insertional Inactivation. What is mean by insertional activation is that you have a functional protein or functional product and what you can do is you can just clone within that functional gene and because of that you are going to disrupt the function of that particular protein. So as it said in this approach a foreign DNA is cloned within the coding gene (see) responsible for a particular type of phenotype. As a result of the insertion the gene product is not available to modulate the phenotype of the host and that is how in the control conditions the host will going to show you the modified phenotype whereas if the insertional inactivation is going to happen, it is not going to show you that particular phenotype and as a result you can be able to select the molecules. This approach is known as the insertional inactivation as it can be used with a suitable genetic system.

(Refer Slide Time: 15:50)

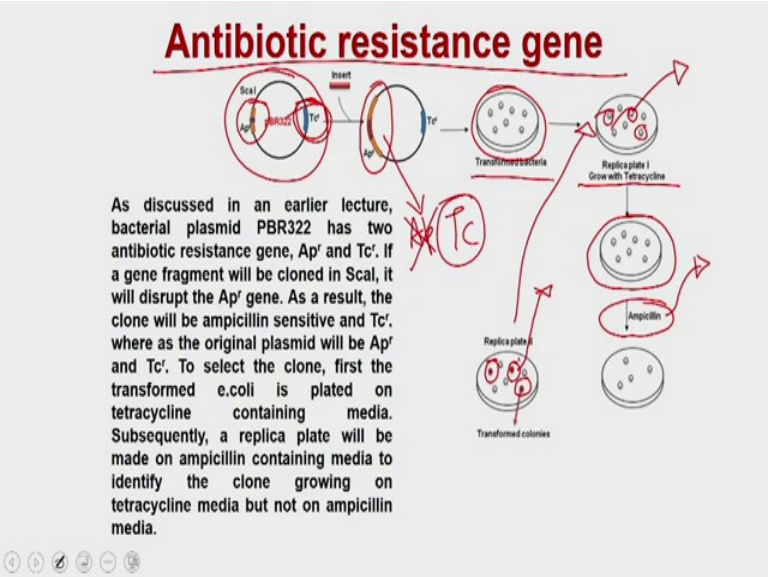


So the first is insertional inactivation of a lacZ. So this approach is exactly the opposite what we have discussed just now where we were using the lacZ complementation. So we were regenerating the compliment lacZ. In this case you have a functional lacZ onto the vector and that actually produces the functional beta-galactosidase and that actually form processes the X gal to blue color product. Whereas if you are actually using a BamH1 side and you are inserting your gene within the lacZ side, so what will happen is that you are going to produce a recombinant DNA but what in this process you have destroyed the lac Z gene and as a result, you are not going to produce the functional beta-galactosidase.

You will going to produce a non-functional galactosidase and that is how it is not going to be able to process the X-gal to give you the blue-colored colonies and what you are going to get is the colorless colonies and these colorless colonies indicate the recombination the delivery of recombinant DNA into the host. So the x-gal system can be used to detect the insertional inactivation of lac Z gene to screen the cloned fragment.

If the gene is inserted to the lac Z, the clone will not be able to produce a functional beta-galactosidase. Hence the blue color colonies indicate the presence of an active enzyme whereas the absence of an insert whereas the colorless colonies indicate the presence of an insert. Okay, so this is the one of the example where insertional inactivations in the lac Z gene will give you the criteria to screen the transformation of the clone into the host.

(Refer Slide Time: 18:04)



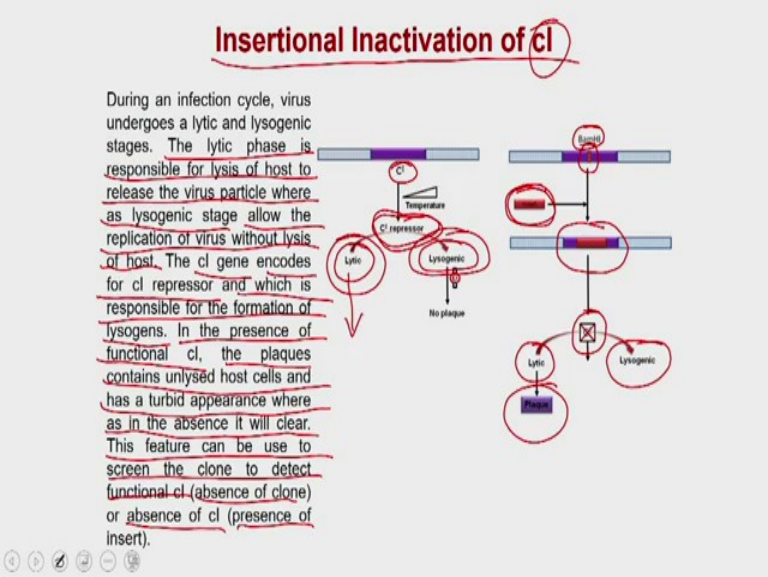
Now let us talk about the antibiotic resistance gene. So if you remember we have discussed about the PBR 3-2-2 the one of the classical and the first vector which is been developed by the molecular biologist and the PBR 3-2-2 has the (two ampicillin) two antibiotic resistance genes, one is called ampicillin resistance gene and the other one is called as the tetracycline resistance genes and so suppose you have used the ampicillin resistance gene and you have cloned insert within the ampicillin resistance gene, which means if you use this fragment, this fragment is going to have no resistance for ampicillin but it is going to have the resistance for tetracycline. So once you transform this into a transform bacteria, you are going to get the colonies.

Now what you have to do is make a replica of these colonies and grow them onto the tetracycline. So all the bacteria will grow whether your insert is present or not because the tetracycline is intact in both the conditions whether the vector has the insert or whether the vector does not have the insert. So these are the colonies which are actually tetracycline races. Now what you do is take these colonies and grow them onto the ampicillin. So what will happen? The some of the colonies will die because they are not going to contain the ampicillin resistance. So what are the colonies are being died?

These are the three colonies which are been died while you were screening them onto the ampicillin. Okay, so these are the colonies which are actually received the recombinant DNA or so now you can take this and go to the original master plate and extract them, so you can take out

these colonies and you can use them for downstream applications. So this is actually called as the insertional inactivation approach where you are using the insertional inactivation of ampicillin resistance gene or one of the antibiotic resistance genes and then you are using the two antibiotics to screen your desirable clones.

(Refer Slide Time: 20:31)



Now the third approach is where you are doing the insertional inactivation of C I or the factor. The CI is a crucial factor and response to the temperature. So CI is a repressor which actually modulate the viruses activity between a lytic cycle to lysogenic cycle. So within the lytic cycle lysis of the host to release the virus particle, whereas the lysogenic stage allows the replication of the viruses without lysis of the host and the CI gene or the CI repressor encodes a CI repressor and which is actually responsible for the formation of the lysogens.

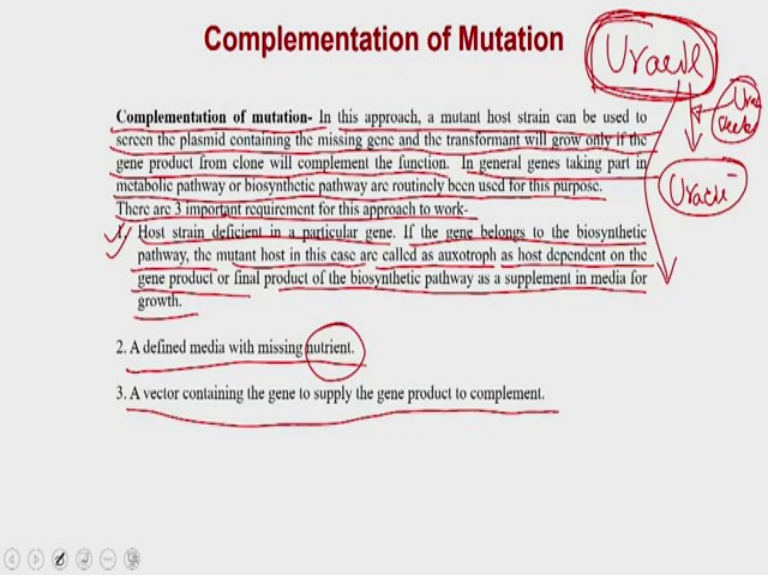
In the presence of functional CI protein, the plaques contains unlysed host cells and has a turbid appearance whereas in the absence of it, it will be clear. This feature can be used to screen the clone to detect functional CI which means absence of clones. So if you have the functional CI it is means you have your clone is not been inserted. But if the CI is been insertionally inactivated, then the CI is not going to perform the function and that is how you are going to confirm the presence of insert.

So you can imagine that CI is actually a CI repressor which actually controls the activity of the virus between the lytic phase to lysogenic phase and the lytic phase is actually responsible for the lysis of the host cells and that is how it is going to produce the plaques whereas the lysogenic phase is allowing the viruses to replicate within the host cells without forming the plaque.

So once you insert into once you take your recombinant gene and insert into the (CI represent) CI genes with the help of the BamH1 what will happen is you are not going to generate the functionally active CI protein and once you cannot generate the functionally active CI protein the virus cannot go to the lysogenic cycle.

It has to go through the lytic cycle and that is how it is actually going to form the plaques. So if you see the formation of plaques after your cloning as well as the transfection using this particular type of vector the viral vector you will be sure that your insert is within the CI gene and it has formed the plaques.

(Refer Slide Time: 23:25)



Now the last approach is that the Complementation of Mutations. In this approach, a mutant strain, so you use the mutant strain and can be used to screen the plasmid containing the missing gene and the transformants will go only if the gene product from the clone will complement the functions. In general the genes taking part in metabolic pathway or biosynthetic pathway are routinely been used in this process.

So in the complementation of mutation, what you have is you have the host strain which cannot synthesize a particular gene product. For example, you may have the mutations within the uracil biosynthesis pathway or you may have some of the essential amino acids biosynthesis pathway. In those cases as long until you won't supply those particular functional gene from the externally supplied vector. It is not going to survive into a deficient media. So the host strain will grow onto the normal media which contains that particular essential item.

For example, if you are talking about uracil biosynthesis pathway, okay and some of the genes of these uracil biosynthesis pathways are mutated into the host strain then this particular host will not grow into a uracil minus media, as long as you will not supply the uracil synthesizing gene by the vector. So if you supply the gene by the vector this particular recombinant transformed host will grow onto the uracil minus media, but onto a normal media which contains a uracil the host strain will grow normally.

So in these cases you will going to have the host strain which is deficient in that particular gene or well that particular enzyme and that enzyme you will supply by the externally supplied vector. There are three important requirement of this approach to work, one which is very important is that you should have a host strain which is deficient in a particular gene. If the gene belongs to the biosynthetic pathway the mutant host in this case are called as the auxotroph as the host depends on the gene product or the final product of the biosynthetic pathway as a supplement in media.

Which means the first requirement to use this particular screening method is that you should have a host strain which is deficient in that particular enzyme or that particular factor. If this factor is a part of the biosynthetic pathway then the host strains slowly depends on the either that particular product or the final product of the biosynthetic pathway and in those cases the host strain is been called as the auxotroph on and that particular media. For example, in this case if you are using the yeast and if you are using the uracil, then the yeast is going to be the auxotroph for uracil.

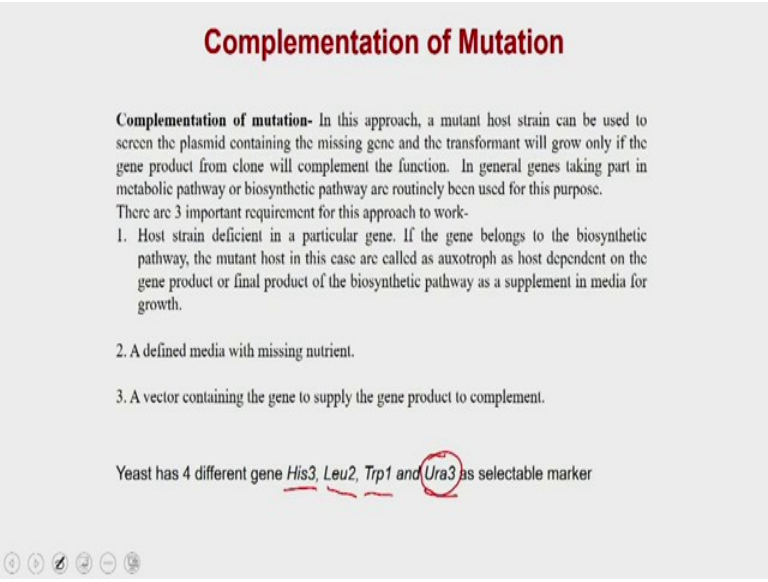
The third is you should have a defined media which should not contains that particular nutrient which you are using for screening purposes, which means it is going to have the missing nutrients and then the third is you should have a vector which actually could be able to supply

that particular gene product which can supply this particular nutrients so that the you will go in to do a complementation and that is how the host which are going to receive this particular type of vector is going to survive into the media deficient into this particular important nutrients and let us continue our discussion about the screening of clones in the case of eukaryotic host.

So what we were discussing, we were discussing about the complementation of a particular mutation and in this particular approach the host is having the mutation in the essential genes and if these essential genes are part of the biosynthetic pathway, then the host is dependent solely on providing this particular nutrients or this particular downstream final product for surviving and running is metabolism, in those cases the host is considered to be an auxotroph for that particular nutrients and as we discussed in this particular approach, there are three requirement.

One, a where the host should have the mutation in one of the crucial gene or the pathway so that it should be dependent on the providing this particular factor from the exogenously provided the vector. Then you should have a suitable vector so that it should be able to provide that particular factor and the third is that the you should have a defined media which should not have this particular nutrients.

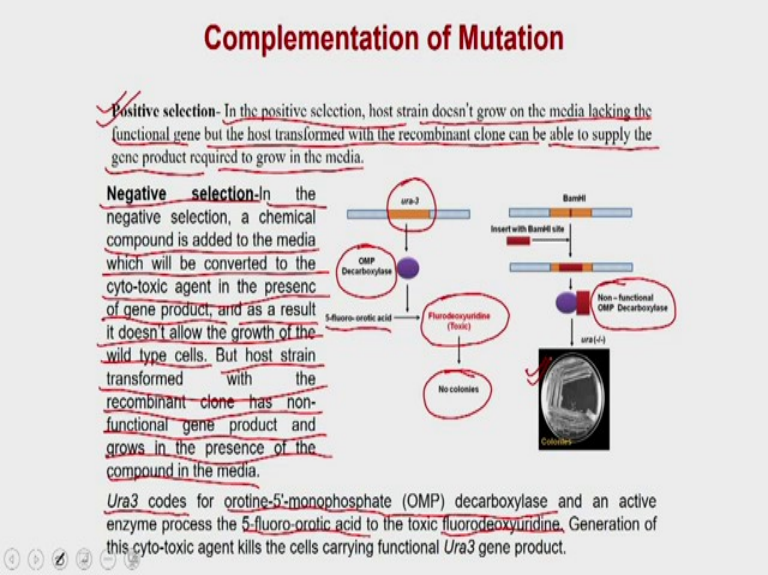
(Refer Slide Time: 29:02)



So in the case of yeast, yeast has 4 different genes which are part of this particular approach. These are the 4 genes of part of the 4 different biosynthetic pathways whether it is for the

histidine, leucine, tryptophan or the uracil and all these 4 genes and their mutations is been found in the yeast different yeast strains and that is why you can use these particular genes for performing this approach which is called as the complementation of the mutations and this approach can be done in two different ways. One is called as the positive selection, the other one is called as the negative selection.

(Refer Slide Time: 29:53)



What is mean by the positive selection? The positive selection is that in the positive selection the host strain does not grow on the media lacking the functional gene, which means the host has the some kind of particular mutations that is why the host will not grow on the media as long as it does not contain that particular metabolite or that particular nutrient but the host transform with the recombinant clone, which is actually going to provide or going to supply these gene product will allow this particular host to grow on the media in the absence of that particular nutrients.

So this is called as the positive selection which means if your gene is present, the clone is going to grow on the media deficient in that particular nutrient but the other approach is that you can have the negative selection. In the negative selection what you can do is, a chemical compound can be added to the media which will be converted to the cytotoxic agent in the presence of gene product and as a result, it does not allow the growth of the wild-type cell which means the wild-type cell are actually producing a active enzyme and this active enzyme you can add some

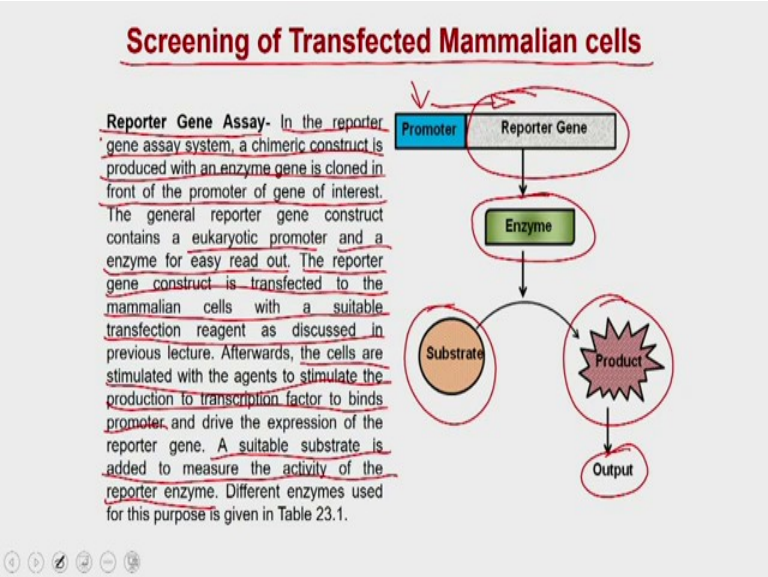
chemical and that chemical is going to be get converted into a cytotoxic agent and then because the enzyme is active the wild type is going to get killed.

But if the host strain is going to be transformed with the recombinant clone has non-functional gene product and grow in the presence of the compound in the media. So if you supply the host strain and it is transform a recombinant clone which will going to probe which will be going to disrupt this particular gene product then even in the presence of that particular chemical the transformed host cell is going to grow.

One of the example is URA-3 gene, which is actually coding for the orotine-5 monophosphate decarboxylase. So this is an enzyme which actually processing a compound which is called as the 5-fluoro-orotic acid to a very-very toxic fluorodeoxyuridine. So what you do is you add the 5-fluoro-orotic acid and in the presence of (OMP) decarboxylase which is going to be produced by the Ura-3 and to a very-very toxic fluorodeoxyuridine and the fluorodeoxyuridine is very toxic, so it will not allow the host to grow but if you clone your insert within the Ura-3 gene.

What you are going to do is, you are going to produce a non-functional OMP decarboxylase and in that pressure and in the absence of the functional OMP decarboxylase the 5-fluoro-orotic acid will not going to be get converted into the fluorodeoxyuridine which means the compound will not going to exert its toxic effect and as a result you are going to see the colonies onto the plate. So this is called as the negative selections. Now, let us move on to the higher eukaryotic organisms such as the mammalian cells and let us see how you can select the mammalian cells when they are being transformed or when they are being transfected with the extra with the recombinant clones.

(Refer Slide Time: 33:39)

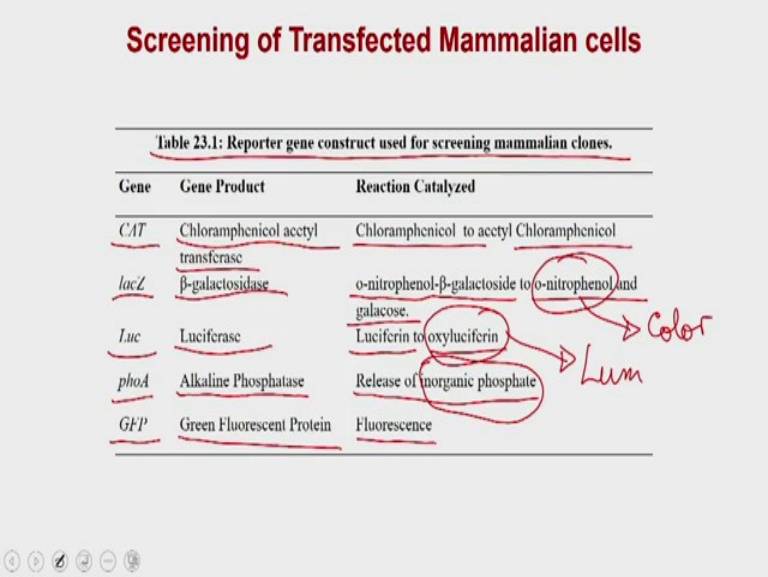


So the screening of transformed or transfected mammalian cell is being done in a utilizing the different types of approach. One of the classical approach is the reporter gene assay. In the reporter gene assay what you have is that in the reporter gene assay a chimeric construct is produced with an enzyme gene product in the front of the promoter of the gene of interest. So (general report) general gene reporter construct contains a eukaryotic promoter so that you will be able to produce that particular gene utilizing the cellular machinery of the mammalian cells and an enzyme which is going to give you a readouts which means these enzyme is going to process a substrate to a product and this product may give you some kind of fluorescence or luminescence or some kind of readable product.

The reporter gene construct is going to be transfected to the mammalian cells with a suitable transfection agent such as the lipofectamineor other kind of transfection reagent, which we have discussed. The cells are now is stimulated with the agents to stimulate the production of transcription factors to bind the promoter. So what will happen is the when you stimulate the cells to produce the transcription factor this transcription factor will come and bind to the promoter and that is how they will drive the transcription as well as the translation of this particular reporter gene and once the translation of the reporter gene is being done, it is going to produce the enzyme and this enzyme is going to process this particular substrate to form the product and that product can be measured by many mean.

For example, it can be used by fluorescence method or luminescence method or simple the chronometry method. A suitable substrate is required, is added to measure the activity of the reporter enzyme. There are several examples of reporter gene system.

(Refer Slide Time: 35:54)

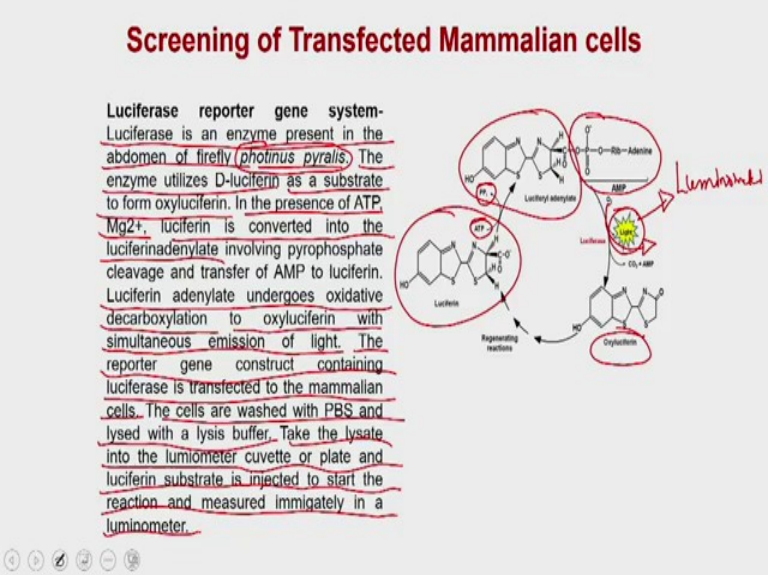


So in a different system, you can use the different types of reporter gene constructs. For example, the CAT, CAT is a gene which is reported which is expressing for the chloramphenicol acyl transferase and that actually converts the chloramphenicol to acetyl chloramphenicol so you can catalyze this reaction and know. LacZ which is actually expressing the beta galactosidase and the beta galactosidase can be used to monitor this reaction where the ortho-nitrophenyl beta-galactosidase will be get converted into the ortho-nitrophenol and galactose and ortho-Nitrophenol is going to give you the color and that can be used to monitor the activity of beta galactosidase.

Simply you have the lac gene which is going to produce the luciferase and the luciferase enzyme is going to convert the luciferin to oxyluciferin and the oxyluciferin is going to give you the luminescence and the luminescence can be measured in a luminometer. Similarly you have pho-A, pho-A gene is actually producing the alkaline phosphatase and the alkaline phosphatase is releasing the inorganic phosphate from the substrates. So you can use the different types of substrate to monitor the release of inorganic phosphate and at the end you have the green

fluorescent protein. The green fluorescent protein or the GFP, which is actually going to give you the fluorescence in the green channel.

(Refer Slide Time: 37:32)

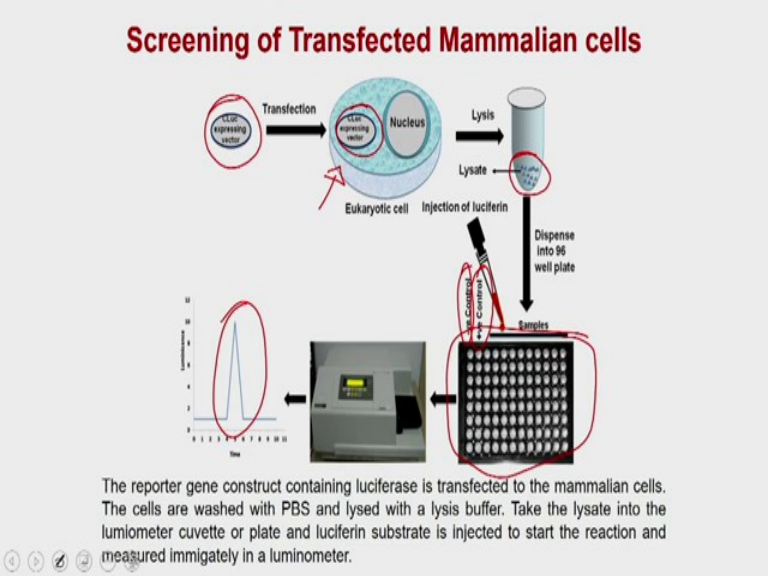


So let us see few examples, one of the classical example is the luciferase based reporter gene construct. The luciferase based reporter gene constructs is that where the luciferase which is an enzyme present in the abdomen of firefly photinus pyralis. So this is a commonly found firefly which is producing this enzyme in their abdomen and because it catalyzes this reaction the abdomen of this particular fly is glowing in the night. The enzyme utilizes D-luciferin as a substrate to form the oxyluciferin and in the presence of ATP and magnesium, the luciferin is converted into the luciferinadenylate which means the enzyme is converting the luciferin to the luciferinadenylate in the presence of ATP.

So in this process it is actually utilizing the ATP to form PPI and the AMP is getting conjugated to the luciferin to form the luciferinadenylate and the luciferinadenylate undergoes the oxidative decarboxylation, which means this particular product is getting oxidative decarboxylation to form the oxyluciferin and once it is going through this oxidative decarboxylation, it actually emits the light from this. So the oxygen is being utilized and that actually produces the light and this is a short-lived signal. So as soon as the luciferinadenylate is getting oxidatively decarboxylated to form the oxyluciferin, it actually produces the small amount of light and that light can be measured simply by an into the luminometer.

The reporter gene construct containing the luciferase is transfected to the mammalian cells and the cells are washed with PBS and lysed with the lysis buffer. Take the lysate into the luminometer or cuvette and the luciferase substrate is injected to start the reaction and measured immediately in a luminometer. So as I said, this signal is very short-lived so you what you have to do is you first transfect the construct into the mammalian cells then once the reaction is or once the stimulation is over, then you can lyse the cells, collect the lysate and put it into the cuvette or the plate, the luminescence plate and then you add the substrate to measure the activity.

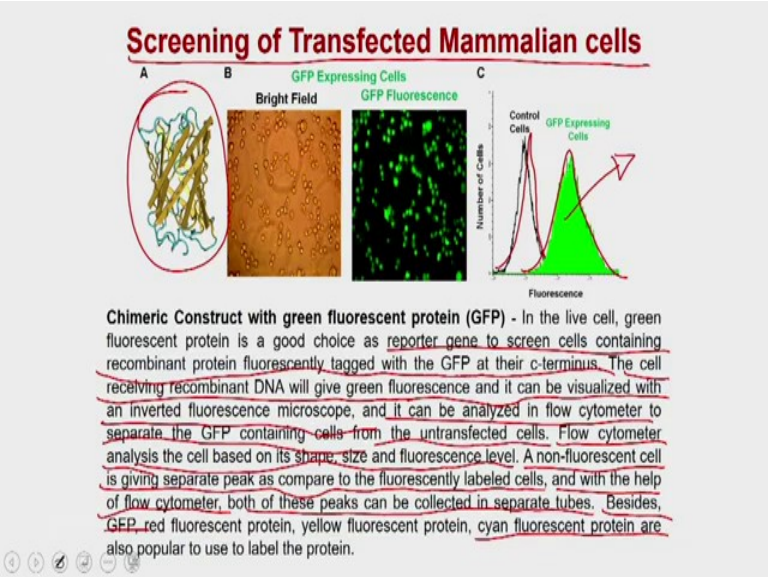
(Refer Slide Time: 40:33)



So these are the things are been depicted here in a schematic diagram that you take the luciferase expressing vector, transfect that to the mammalian expression system, it will go and express the luciferase and then once the stimulation is over which means you can stimulate these cells because the luciferase will be in control to the eukaryotic promoter.

The luciferase will be produced inside the cell according to the level of stimulation what is being produced from this eukaryotic cell, then you can collect this lysate and put it this lysate into the black plate which is actually the ideal plate for doing the luciferase reactions. You can include some negative control as well as the positive control and then you add the substrate into this and once you add the substrate you put it into the multimode reader or the luminometer and once you do that you are going to see a short-lived signal from the enzyme.

(Refer Slide Time: 41:39)



Now let us move on to the next approach, the next approach is that you can actually express the green fluorescent protein into the cell and the green fluorescent protein is a reporter gene to screen the cells containing the recombinant protein fluorescently tagged with the GFP at their c-terminus. The cells receiving recombinant DNA will give green fluorescence and it can be visualized with an inverted fluorescent microscope or it can be analyzed in a flow cytometer to separate the GFP containing cells from the untransfected cells.

Which means the GFP is a green fluorescent protein. It gives the green fluorescence so in the first step, what you do is you take the cells, you transfect them with the GFP. If the construct is getting into the cell, it is actually going to give the green fluorescence. So all the cells which are actually giving the green fluorescence are transfected cells and all these transfected cells can be separated from the non-transected cell which are not going to give you the green fluorescence simply by using the flow cytometer.

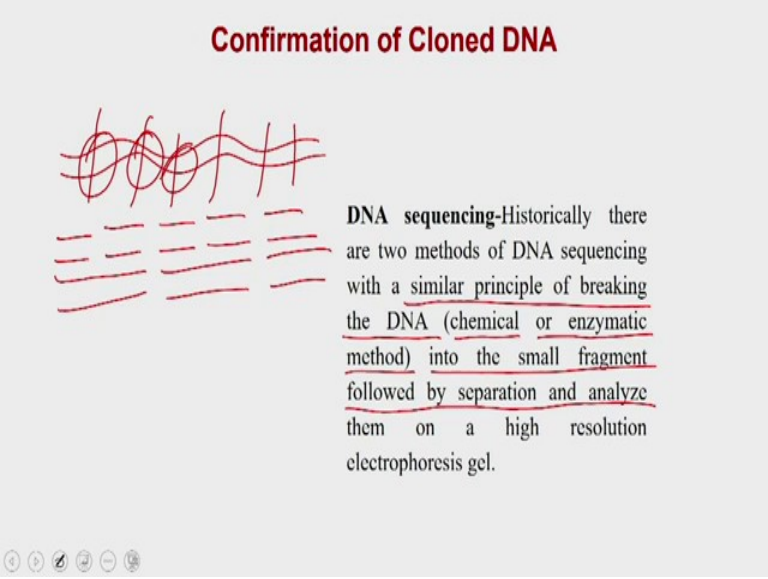
So flow cytometer is an instrument which actually separates the molecule based on the fluorescence, based on the size and shape also so you can select the cells based on the size and shape and then you can measure the fluorescence. If the cells are showing the green fluorescence, you can ask the instrument to collect these cells, a typical fluorescence curve will look like something like this where these are the cells the black colored line are the control cells

or the non-transfected cell whereas the green color cells are considered to be the GFP expressing cells.

So you can ask the instrument to collect these cells into the separate wells and that is how you can select the transfected cells from the non-transfected cells. So the flow cytometer analyzes the cell based on shape, size and the fluorescence level. A non-fluorescent cell is giving the separate peak as compared to the fluorescent labeled cells and with the help of flow cytometer both of these peaks can be collected in a separate tubes besides GFP.

Because the GFP is giving green fluorescence, you can also have that option of red fluorescent proteins, you can use the yellow fluorescent protein, you can use the cyan fluorescent proteins and so on and all these fluorescent protein are very popular to label the protein and to separate the transfected cell from the non-transfected cell.

(Refer Slide Time: 44:32)



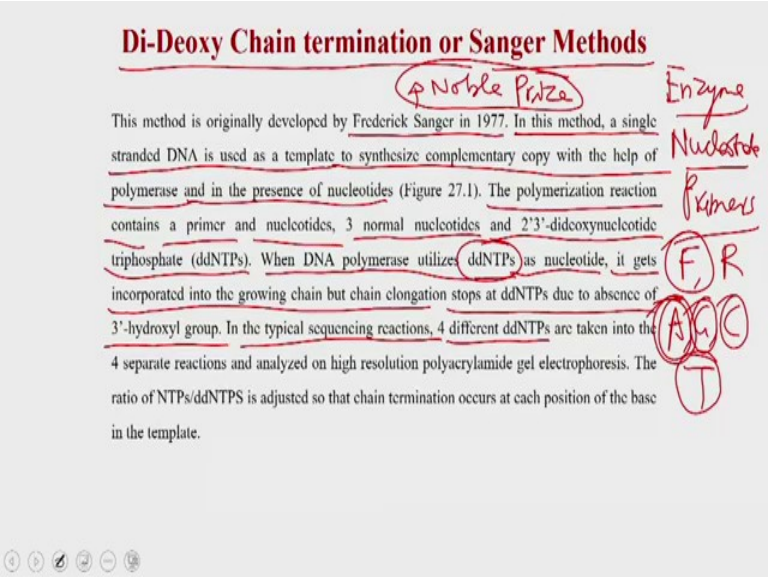
So this is all about the isolation or to detection of the transfection into the cell or to screening of the clone based on the sum of the selection markers which are available on to the vectors and once you are sure that the host or you have selected the desirable clones, the next step is that you have to confirm that your host is containing the clone which also contains the right gene. So for confirming the clone, what you are supposed to do is that you are going to sequence the complete

gene to know that whatever you have cloned, the foreign gene is actually intact, it does not have any mutations and it is actually in the right frame.

So for that purpose you have to do a DNA sequencing to confirm the clones or to confirm the DNA isolated from these clones. So for that purpose we have the, so there are only two methods or there are two methods which we are going to discuss in this particular lecture. So in both of these methods are having the similar principle that they are actually breaking the DNA either by using the chemical or the enzymatic method into the smaller fragments and then they are been, so what you are doing is suppose this is your full gene length, what you are doing is you are actually cutting this DNA into the smaller fragments and once you do that you are going to get these fragments.

Small fragments from this particular DNA but apart from this you are also going to get these fragments as well. Which means the fragments which are overlapping this region or this region or this region, which means so if you analyze all these fragments you will be able to put them into a right order and that is how you can be able to sequence them using the these techniques. So we are going to discuss the two methods, one is called as the Sanger method. The other one is called as the Maxam Gilbert method.

(Refer Slide Time: 47:05)



So in the Dideoxy chain termination method or the Sanger method, which is actually been developed by the Frederick Sanger in 1977 for which the Fredrick Sanger got the Nobel prize and so Frederick Sanger was the only person who got the Nobel prize first for discovering the method of protein sequencing and then he has developed another method for doing the DNA sequencing and for both of these discoveries the Frederick Sanger got the Nobel prize.

So in this method what we are doing is, we are generating a single stranded DNA as a template. We are using single standard DNA as a template to synthesize the complimentary copy with the help of a polymerase, which means typically we are just doing the PCR in the presence of nucleotides the polymerization reactions, you know the polymerization reaction requires the enzyme which means the DNA polymerase it requires the nucleotides and it requires the primers if you remember, when we were discussing about the PCR. So these are the three things we require.

For the primers you need the forward primer as well as the reverse primer. So in this case, we do not use the forward or the reverse primer. We use only one of the primers and because we have to just generate the single-stranded DNA. So what you do is in the polymerization reaction, you have the primer your nucleotides so, then the nucleotide what you are going to add, you know that the we have 4 different types of nucleotide adenine, guanine, cytosine, and thymine.

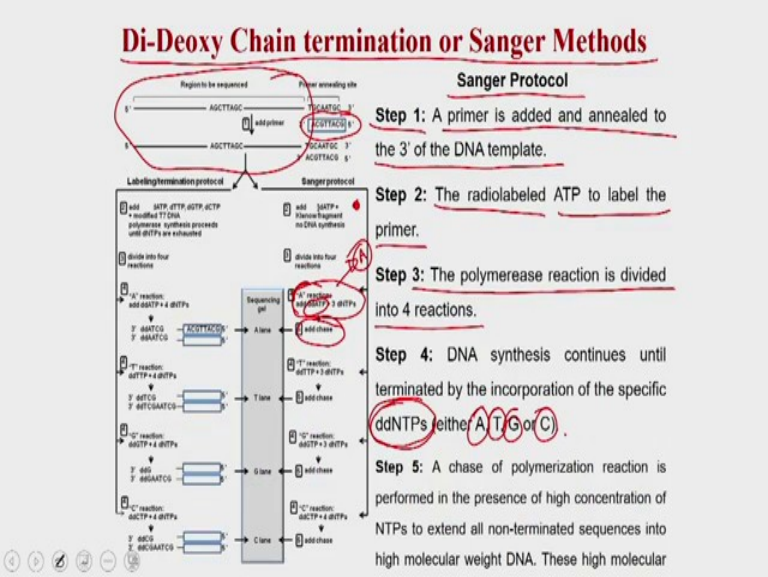
So when you do these reactions you add 3 nucleotides, normal nucleotides which means either adenine, guanine, cytosine and the instead of fourth nucleotide you add the 2-3 dideoxynucleotide triphosphate, which means suppose if this is a reaction for adenine then you add the guanine, cytosine and thymine as a normal nucleotide and you add the two prime, three prime di-deoxynucleotide adenine triphosphate as the fourth nucleotide.

So what will happen is when the DNA polymerase utilizes the di-deoxy nucleotide as a nucleotide the di-deoxy nucleotide is does not have the, is just blocks the growing chain. So that is how as soon as the DNA polymerase utilizes the di-deoxy nucleotides it actually blocks the chain terminations due to the absence of the three prime hydroxyl group.

In the typical sequencing reaction, you use the 4 different types of di-deoxy nucleotides, which means you are going to do the 4 reaction, one for adenine, one for guanine, one for cytosine and

one for thymine and then you are going to run these 4 separate reactions on to the high-resolution electrophoresis gels and then you separate these four reactions and then you analyze them and that actually. Once you interpret these you will be able to get the sequence of that particular DNA.

(Refer Slide Time: 50:25)

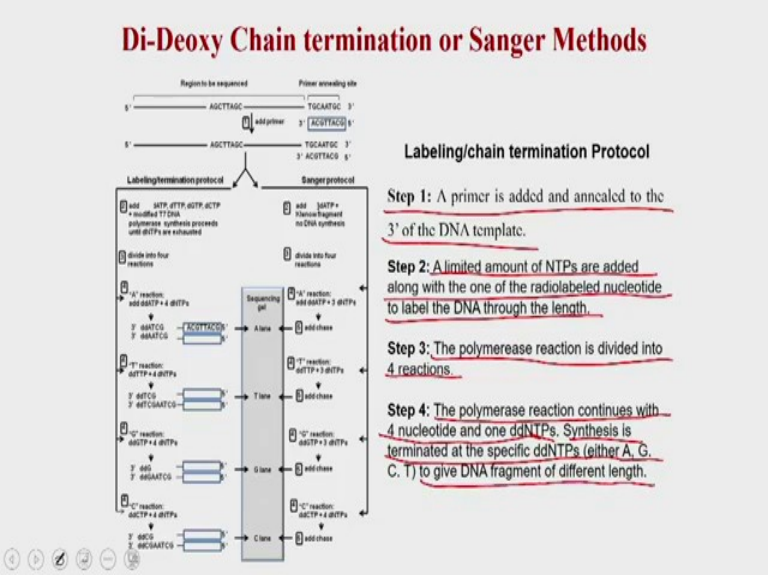


So in a typical Dideoxy chain termination method you can actually do this method in two, utilizing the two different protocols. One is called as the Sangers original protocol which Sanger has been developed. So in this what you have is in the step 1 you add a primer and let it be annealed to the three prime end of the DNA polymerase, a DNA template. Which means suppose this is the region which you have to sequence in the Sangers method, what you add is first you add the primer okay, which was going to anneal on to the three prime end of the DNA template then in the second step you use the radio labeled ATP to label the primers which means you are actually tagging the primer with a particular type of radioactivity.

Now after this in the step 3, what you do is, you take this particular reaction and divide that into the four aliquots or four reactions and all the four reactions you are going to add the either adenine or thymine, guanine or cytosine. So you are going to do the reactions for all the four nucleotides by using the different types of Dideoxy nucleotides, which means in if you are doing reaction for adenine, you will add all other three nucleotides which means you are going to add guanine, cytosine and thymine but what you are going to add?

You are going to add the Dideoxy adenine triphosphate and then once you do that, you add the chase which means you are going to allow this reaction to happen. So what will happen is the places where you are going to add the ddNTP's for adenine, thymine or guanine or cytosine it is actually going to terminate the reactions on that particular nucleotides. Which means if you are doing A reactions, the termination would be for A. If you do the T reactions then the termination would be on thymines. If you do the (termination) reactions for guanine then it is termination on the guanosine and if you do it for cytosine, then the termination will be on the cytosine.

(Refer time slide: 52:49)

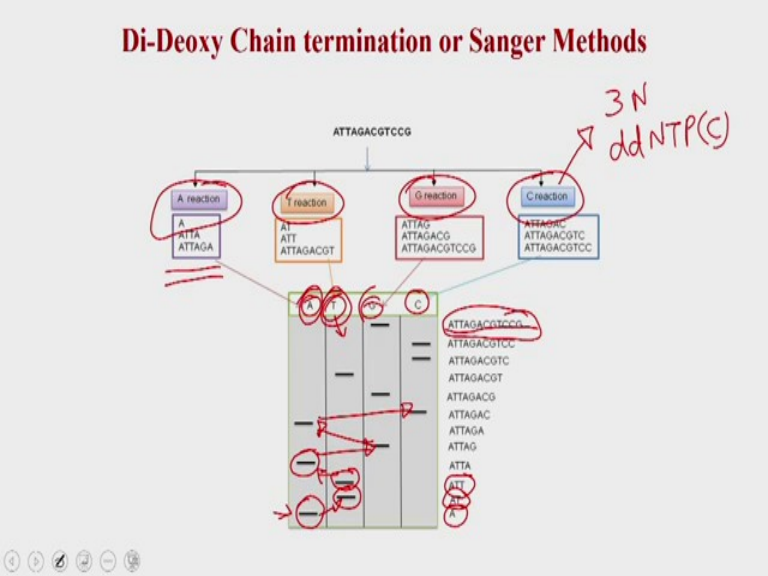


Whereas in the alternate reaction which is called as the (label) labeling or the termination protocol, the step one remains the same where you add the primers and let it be anneal on to the three prime over end. Then what you add is, you add the limited amount of NTPs along with one of the radio labeled nucleotide to label the DNA throughout the length. Then you add the polymerase reaction into the four reactions and then you do exactly the same what is being discussed for the Sangers protocol and the polymerase reactions continue with the four nucleotide and one of the nucleotide would be the di-deoxy nucleotides and synthesis is terminated at the specific di-deoxy nucleotides to give the DNA fragments of different lengths.

Now once you analyze, once you follow the protocols whether it is Sangers protocol or the termination protocol, you are going to get the DNA fragments which means you are going to get the four reactions. Now you run these four reactions on to the high sensitivity sequencing gels,

which are actually the polyacrylamide gels and the once you run them you are going to separate these small stretches of DNA and then what you are going to do? You are going to do the auto radiograms because the primers are labeled so you can do autoradiogram and that actually will going to give you the bands.

(Refer Slide Time: 54:14)



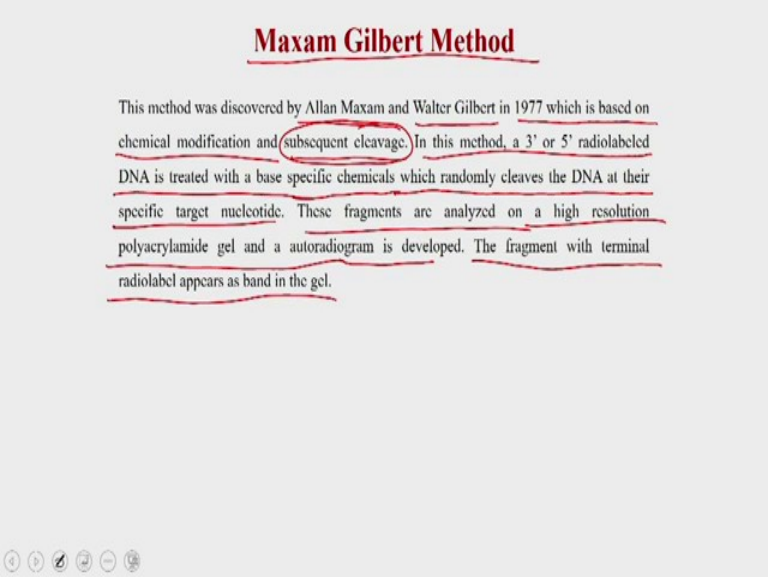
Now let us see how you are going to interpret. So imagine that you have done the A reactions, T reactions, G reactions and C reaction which means in the C reactions you have the three nucleotides which are normal and the fourth one is ddNTP which is actually the cytosine. Similarly is for the guanine (similarly) for thymine and adenine. So once you have performed the separations you are going to get the autoradiogram like this. So this is the lane for A, this is for T, this is for G, this is for C.

Now what you are supposed to do is, you are going to read in a reverse order which means you will first go to the lowermost band. So in this case the lowermost band is in the A lane, so you are going to write A. Now the next band is in the T so you are going to write T. Now the third band is also again in the T so you will write ATT. Now the fourth band is again in the A so you will actually going in the reverse direction, which you are going in the reverse direction, like this.

Like this, okay, so you will be keep moving, keep moving and ultimately it will going to tell you the sequence of DNA. In case, the places where you have the bands on the same height then you will be able to distinguish simply by looking at the other nucleotides. So that is how you, so you can imagine you can understand this that in the case of A all the termination happens on the A, in the T all the terminations are happening.

So this means the terminal nucleotide is going to be the T in the case of D reactions, terminal nucleotides will be A in the case of A nucleotide and that is what is given here, okay. So that is how you actually if you go by the in the reverse order you will be able to interpret the sequence and you will be able to deduce the complete sequence of the DNA.

(Refer Slide Time: 56:31)

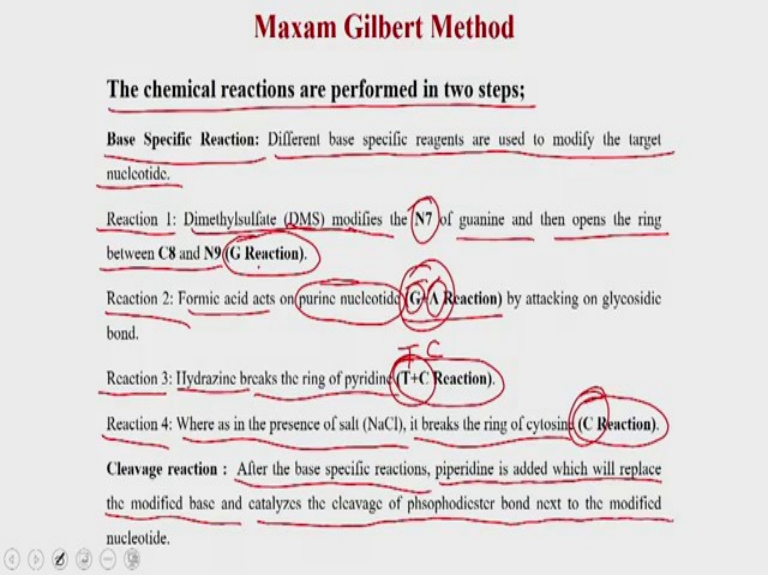


Now let us move on to the next step, so the next method is called as the Maxam Gilbert method. The Maxam Gilbert method was discovered by the Maxam as well as the Gilbert in 1977, which is actually based on the chemical modification and subsequent cleavage. So if you remember the Sanger method is based on the PCR mediated reactions, which means where the enzymatic enzymes are also been involved, DNA polymerase and other kinds of enzymes and once you incorporate the enzymes, actually you are also could make the mistakes because you know that all these DNA polymerases have the some level of processtivity. So if the DNA polymerase incorporate A instead of T, because even if the template is saying that you have to incorporate A and it is incorporates T (you) your sequencing may not be correct.

Because the ultimately what you are going to get is the reactions or the bands and that is how the Sanger method was dependent on the enzymatic or the enzymes to give you the final DNA sequence, which means the whole method depends on the sensitivity and the accuracy of the DNA polymerase which you are using for these reaction whereas in the case of Maxam Gilbert method, the Maxam Gilbert method utilizes the base modification as well as the cleavage. So in this method you actually first modify the basis and then you catalyze a cleavage reactions.

In this method also, what you what you are supposed to do is either you label the 3 prime or the 5 prime of the DNA and then you treat the DNA with a base specific chemicals which are going to modify the bases and then you add the cleavage reaction and that actually is going to cleave the DNA at the random stuffs or random places. These fragments are then you analyze on a high-resolution polyacrylamide gel and subsequently you develop them in an autoradiogram. The fragment with the terminal radiolabel appear as band in the gel.

(Refer Slide Time: 58:49)



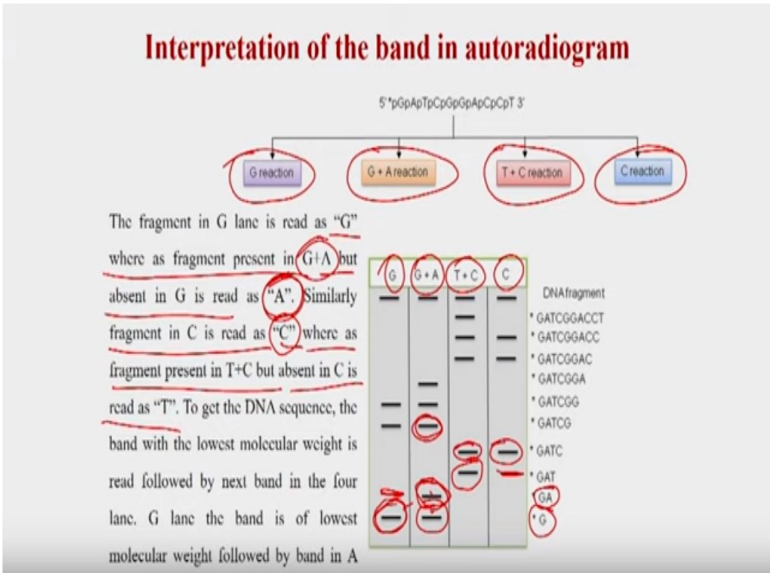
For Maxam Gilbert method the chemical reaction has to be done in two different (process) steps. In the step 1, so the chemical reactions are being performed in two steps, the base specific reactions which means you are going to perform the base specific reactions to modify the nucleotides; these you have to do for guanine and all other nucleotides. So the reaction 1, you have to do for the, with the dimethyl sulphate and (di) what the dimethyl sulphate is going to do? It is going to modify the N7 of guanine and it is going to open the ring between the C8 and N9

and these reactions are called as the G reaction because it is modifying the guanine into the DNA.

In the reaction 2, you are going to add the formic acid and that actually going to modify the purine nucleotide which means it is going to modify the guanine as well as the adenine and that is why, this reaction is going to be called as G plus A reaction. Now in the reaction 3, you are going to use the hydrazine and that actually will going to modify the all pyridines and that is why it is called as the thymine and cytosine because it is going to modify the all the thymine as well as the cytosine and that is called as the C reaction.

So you can see that we have used the 4 different reagents and all these 4 reagents are specifically modifying either the guanine, or guanine or adenine, or thymine and cytosine, or the cytosine reaction. Once you have done all these modifications, then you will perform a cleavage reactions and after the base specific reactions are over then you add the piperidine which is actually going to replace the modified basis and catalyze the cleavage of phosphodiester bond and that is how it is going to generate the DNA fragments of random sizes and these (generally) these DNA of the random sizes are then going to be separate onto the high resolution polyacrylamide gels and subsequently you will do the autoradiogram and you will get the bands.

(Refer Slide Time: 61:22)



Now what you are going to do? You are actually getting the 4 reactions G, G plus A reactions, T plus C reactions and C reactions. So this is the pattern in the G reaction, G plus A reaction, T plus C reaction and C reaction. So how the interpretation goes on? The fragment in G lane is read as G whereas the fragment present in G plus A but absent in G is read as A okay, which means if you are getting the two bands, one is in the G lane another one is the G plus A lane which means you are going to have G and A together.

But suppose you have the bands in the, you do not have the bands in the G lane then the band which is present in the A (lan) G plus A lane is going to be read as the A. Similarly the fragments in C is read as C which is which means all these fragments are having the terminal C. Whereas the fragment present in T plus C but it is absent in C is termed as the T. Now let us see how it is done. You have here, the G here you have G plus A okay, which means you are actually having this as G.

In the second one you see that here you have a G plus A lane, but the band into the corresponding position within the G reaction is absent which means this is actually a band for A reactions. So that is why it becomes GA. Now you go to the T plus C, you will see the T plus C has a band but the corresponding band in the C reaction is absent that is why it is (consider) this is going to be T, so you write the T. Now in the fourth lane if it is reaction in, if you have band in both the lanes, which means it is going to be C. Same is true here.

So in this case also, you will go in the reverse orientation, which is which means this is going to be the last nucleotide then the second last and second last and so on. So these, so that is how you are actually going to interpret the DNA sequencing results from the Maxam Gilbert method where you are going to read from the bottom and if you have the bands both in the G as well as G plus A then it is considered to be G if you do not have the bands in G lane, but you have the bands in G plus A, then it is going to be considered as A and same is true for the T plus C and the C also.

So this is what we have discussed so far. Apart from these two methods which are actually the older method, which are more important in terms of the understanding how the process been evolved. But currently we do not use the Sanger or the Maxam Gilbert method but we use is

some of the High throughput sequencing technologies, which actually gives more accurate result as well as they also give you the result in a very-very small period of time.

So with this I would like to conclude our lecture here and in the subsequent lectures now what we are going to discuss, we are going to discuss about how to exploit this transformed host colonies or transformed host to produce the proteins for downstream biotechnology applications and in that particular kind of discussion we are going to discuss how to, what are the different strategies you can use or (you) what are the different strategies you can adopt for protein production into the prokaryotic system as well as the eukaryotic system. So with this I would like to conclude our lecture here. Thank you.