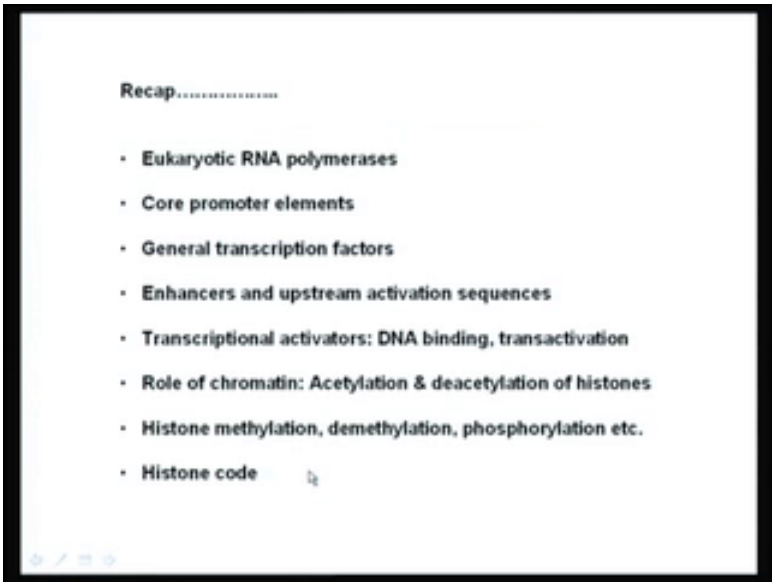
Eukaryotic Gene Expression: Basic& Benefits
Prof. P N Rangarajan
Department of biochemistry
Indian Institute of Science, Bangalore

Lecture No. # 09
Eukaryotic gene regulation
DNA Methylation

We have discussed this for the various post translation modification that takes place for histones and how all these post translation modifications affect chromatin structure and also gene expression. We have studied how histones acetylated, how they are then deacetylated, then histone acetylation and then, histone methylation, histone phosphorylation, histone demethylation and so far so forth.

What we will discuss today is also see in addition to the post translation modifications of histones, the DNA itself is also subjected to certain modification and the one particular modification I will be going to discuss today is about methylation of DNA. So, today's topic we are going to discuss about how methylation of cytosine reduces in DNA is going to affect gene expression in eukaryotes.

(Refer Slide Time: 01:09)



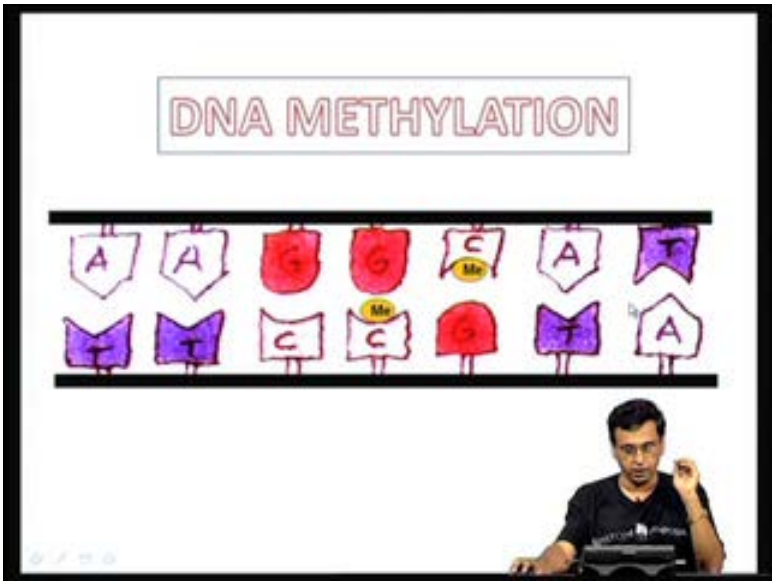
So, before I begin, I will just want to, I just want to briefly summarize what we have learnt so far. As I have discussed in the previous classes, we began with eukaryotic RNA polymerases. Then, we discussed about the various core promoter elements and how

their variations can influence in regulation. Then, we discussed about various general transcription factors starting from TF2, A, B, D, E, F and so on so forth.

How they affect gene regulation and how variations within general transcription factors itself can bring about differences in the levels of gene expression. Then, we moved further up in the promoter, discussed about upstream activation sequences, distil promoter sequences, proximal promoter sequences and how various transcription activators through the DNA binding domain bind to the upstream activation sequences and to the trans activation domains can interact with general transcription machinery and bring about transcription activation.

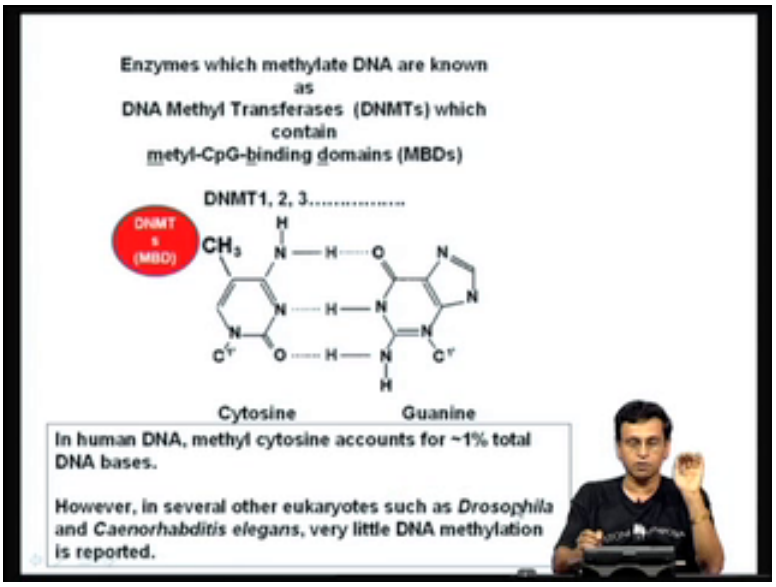
Then, we discussed a very important aspect of gene regulation, wherein we said DNA is not present as naked DNA inside the cells; DNA is actually present as chromatin. Therefore, we discussed how chromatin templates are transcribed in vivo as well as in vitro. Then, we brought in role of histones and we discussed about how histones act as negative regulators of gene expression and unless you modify the histones, either add an acetyl group or remove the acetyl group or methylate or phosphorylate, histones cannot be moved around and depending upon the histones bind DNA tightly or bind DNA loosely, activation or repression of transcription time can take place and the whole bunch of regulators which do these modifications of histones and then, we brought about the very important concept called histone code in addition to the genetic code.

(Refer Slide Time: 02:51)



Today, we will discuss about a very new concept in regulation of gene expression in eukaryotes namely how methylation of DNA influences gene regulation. We all know the DNA consists of four different bases like adenine, guanine, cytosine and thymine. Now, among these four bases, it turns out the cytosine actually can undergo methylation in DNA of eukaryotes in exception in higher eukaryotes.

(Refer Slide Time: 03:16)

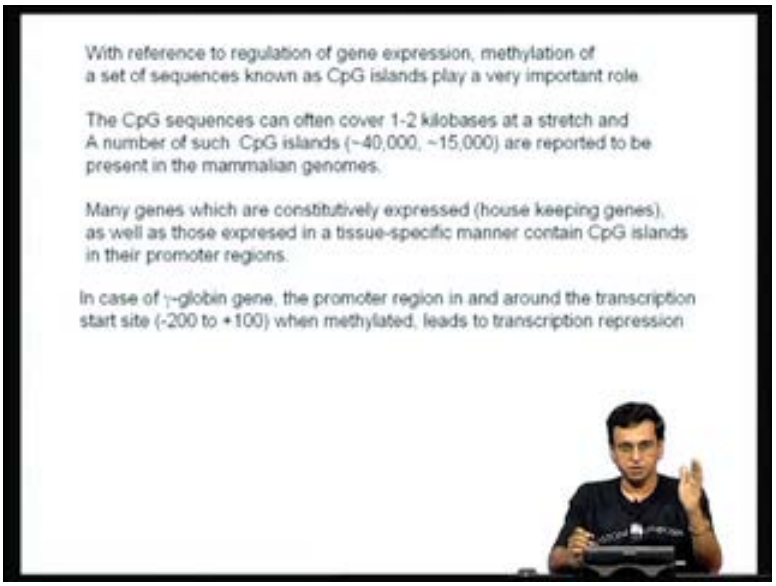


So, you can see this is the usual GC base pair that exist in double stranded DNA. The guanine base pairs with cytosine and there are 3 hydrogen bonds between the guanine and the cytosine in eukaryotic DNA. What we are going to discuss is that there are bunch of enzymes in eukaryotic cells and these are known as DNA Methyl Transferases or DNMTs and it turns out that these DNA Methyl Transferases can actually add methyl group at the 5 position of the cytosine induce. Therefore, you can convert a cytosine into a 5 methyl cytosine.

So, this is a very important modification which has a very profound role in regulation of gene expression. This is what we are going to discuss in class today. It turns out these enzymes which are called as the DNA methyl transferases, they contain what is called as a Methyl-CpG binding domains. There has to be H here. I apologize for the mistakes Met HIl. So, it is methyl-CpG binding domains are MBDs and through is methyl binding domains that these DNA Methyl Transferases add a methyl group to the DE up to cytosine residues of DNA.

In fact, it turns out in case of humans, if you take the genomic DNA in humans; almost one percent of the total DNA in human undergoes well accounts for methyl cytosine. Whereas, very interestingly in other eukaryotes, especially lower eukaryotes like drosophila and caenorhabditis elegans which is a nematode, you do not see DNA methylation. So, this is a very important difference. Unlike the histone modifications which occur in all most all the eukaryotes, it was starting from fly or east to man, the DNA methylation has not been reported to take place in many of the lower eukaryotes like drosophila and celigants, but it is very well reported in mammals, especially mouse and humans analyze also well studied.

(Refer Slide Time: 05:05)



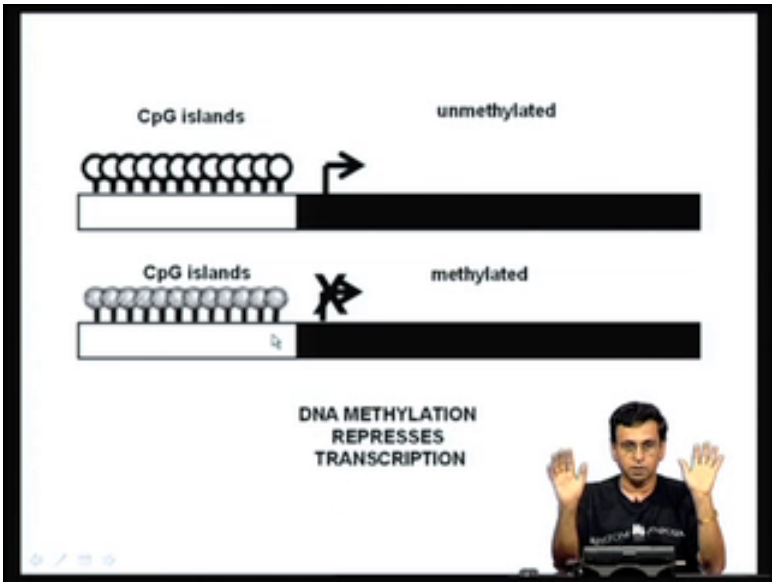
So, when you come to regulation of gene expression, what is the focus of our lecture series? It turns out methylation of set of sequences known as CpG islands play a very important role in the regulation of gene expression. So, this di-nucleotide CGCG is what the target is for DNA methylation transferases and this di-nucleated repeats the cytosine gets methylated in selected regions.

The CpG sequences can often be present up to 1 to 2 kilobase stretches. Huge stretches are these di-nucleotides. CG repeats may be present in certain regions of promoters and such stretches are actually called as CpG islands. For example, if you look at the mouse genomes, there are about 15,000 such CpG islands have been estimated to be present,

whereas in the human genome, the number may be as high as 40,000 such CpG islands may be present.

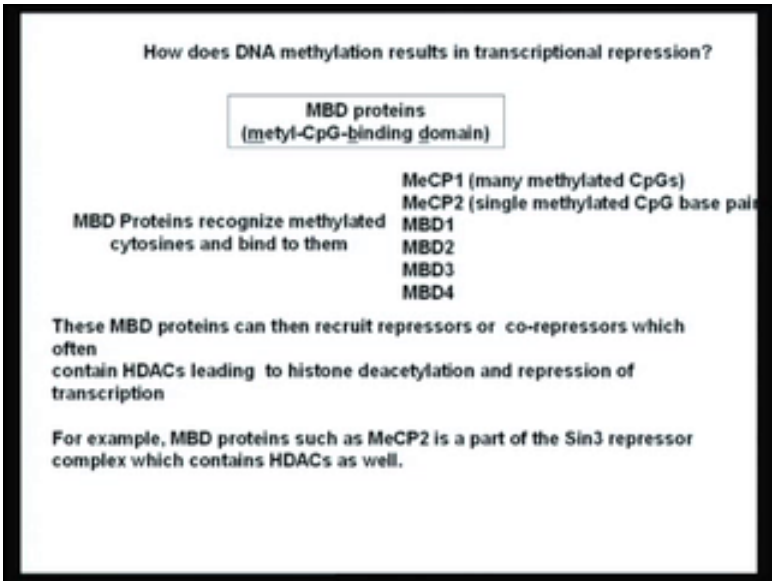
So, many of the genes which are constitutively expressed which are actually called as housekeeping genes. These housekeeping genes as well as those genes which are expressed in a tissue specific manner, they contain the CpG islands in their promoter regions. For example, if you take the gamma globins gene, the promoter region in and around the transcription start side, that is 200 base for sub-stream and 100 base for the downstream of the transcription start side when methylated leads to transcription repression. So, in general, the basic rule before we go into the details of this talk, remember methylation of the CpG, the cytosine in the CpG results in repression of transcription whereas, de-methylation results in activation of transcription.

(Refer Slide Time: 06:41)



This is what has been shown. So, a number of promoters in eukaryotic genes contain what are called as Cp G islands and then, these CpG islands are present in un-methylated state, you get transcription activation, where as in the same CpG islands, when they are present in the methylated form, this results in transcription repression. So, these enzymes which add a methyl group to the cytosine residue CpG islands act as a negative regulator of gene expression in the case of eukaryotes. Remember this point. So, the important principle that we have discussed is that DNA methylation actually represses transcription in eukaryotes.

(Refer Slide Time: 07:20)



How does this DNA methylation repress transcription? As I told you, these MBD proteins which contain this methyl CpG binding domain, they recognize what are called as the recognized methylated cytosine's and bind to them. Sorry I made mistake in the previous slide. The MBD proteins which recognize this methylated cytosine's contain the MBD domain. DNA Methyl Transferases do not contain MBD domains. DNA Methyl Transferases add a methyl group to cytosine and then, once the cytosine is methylated, this methyl cytosine is actually recognized by proteins called MBD proteins which actually contain methyl CpG binding domain. That is why they are called MBD proteins.

So, these MBD proteins contain a methyl CpG binding domain and through these domains, they actually recognize this specific methyl group of cytosine residues in the DNA. There are number of such methyl CpG binding domain contain proteins have been identified so far in different eukaryotes like they are called as for example, MeCp1, MeCp2, MBD1, 2, 3, 4 so on and so forth and there these enzymes or these proteins vary in a number of different properties as well as in there structure. For example, the MeCp1 actually mainly methylate many methylated CpG's, whereas the MeCp2 actually methylate only single CpG base pairs.

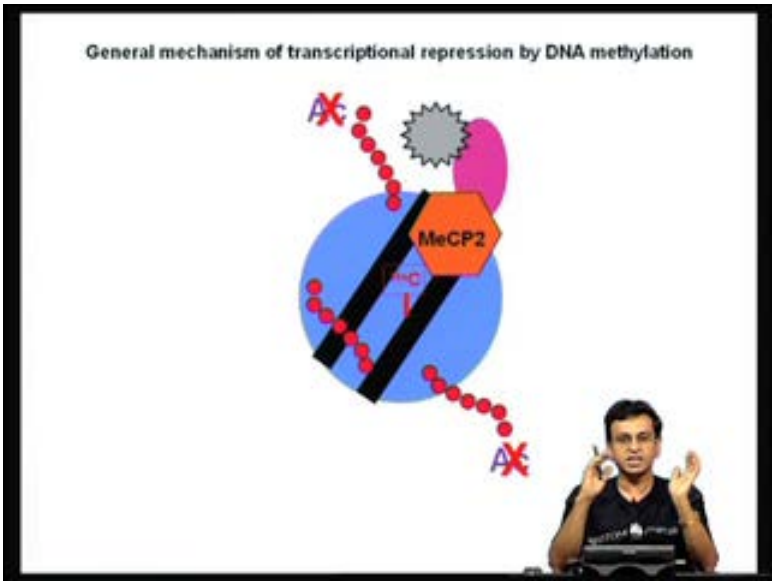
So, there are many minor differences between these enzymes and also on their ability to methylate various CpG islands. So, these MBD proteins, when they recognize these

cytosine residues in the DNA and when they bind, then they go and recruit repressors or co-repressors. These co-repressors suppressor complex usually contain either, histone deacytelizers or histone methyl transferases and when these are recruited to the promoter that leads to transcription repression.

So, the message I am going to convey you is that DNA methylation acts in conjunction with histone modifications to bring about transcription repression. So, once DNA is cytosines, the DNA is methylated, these methyl cytosines are recognized by specific enzymes which contain these MBD domains and once they recognize this methyl CpG islands, cytosine's residues some bind the DNA. Then, they recruit enzymes which negatively modulate gene expressions such as the histones deastylizers or histone methyl transferases. They then modify the histones in the vicinity of the promoter leading to tightening the histones and resetting repression of transcription. For example, MBD proteins such as MeCp2 is actually one of the components of a big repressor called sin3 repressor which is a multi-protein complex and one of the components of this repressor is MBD containing protein namely MeCp2 and these repressor complex also contains a histone deacetylates.

So, you can see the methyl binding proteins or the methyl cytosine binding proteins actually attract the histone modifying enzymes, such as histone deciliters and histone methyl transferases and that is how they repress gene expression. So, the major mechanism by which DNA methylation represses transcription is by actually recruiting histone deacetylases or histone methyl transferases to the vicinity of the promoter through the interaction of the MBD proteins with these histone modifiers.

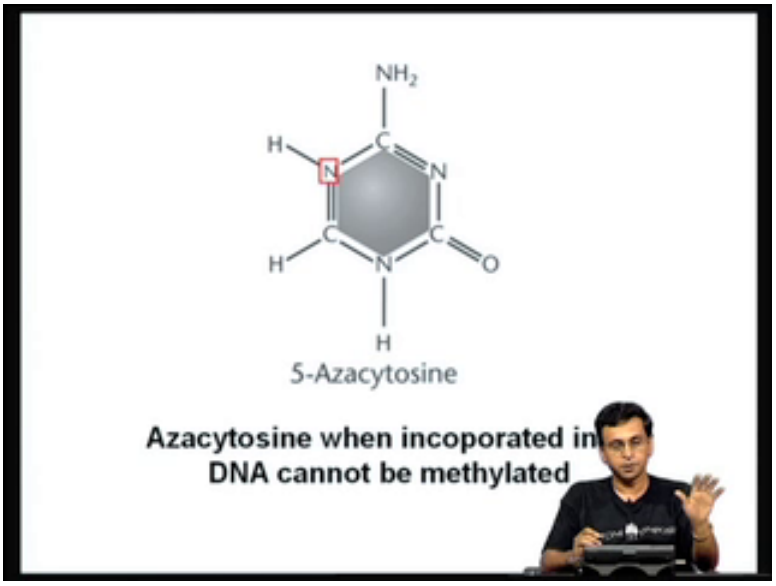
(Refer Slide Time: 10:32)



So, this is the general mechanism of transcription repression by DNA methylation have a depicting form of the cartoon. Let us say for example, there is a transcriptionally active chromatin where specific residues in the H3 or H4 and histones are actually acetylated. That is what I have shown here and once the DNA in this nuclear zone is methylated, this methylated cytosine is now recognized by the methyl binding protein or the MED containing protein. In this case for example, the MeCp2.

Once this MeCp2 binds to this methyl cytosine residue of the DNA, this now attracts a co-repressor complex and this co-repressor complexes usually contain negative regulators of gene expression, such as histone deastylizers or histone methyl transferases. In this case for example, if this contains histone deacytalize, then this removes acetyl groups of the histones. Therefore, histones now bind the DNA very tightly. There we are preventing the assembly of prenishes complex resulting in the repression of gene expression. So, these are very general mechanism by which DNA methylation brings about repression of transcription eukaryotes

(Refer Slide Time: 11:38)

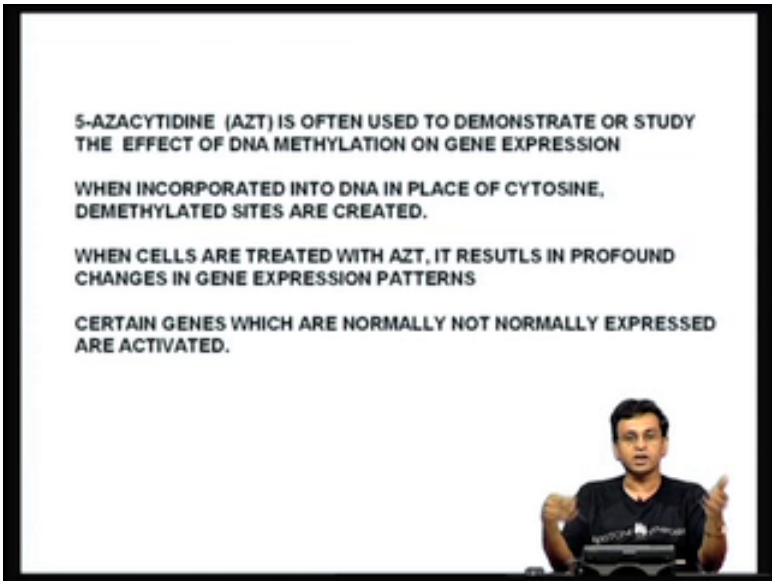


Now very interestingly, there is a compound. This is the normal structure of cytosine, where I have shown and this is the 5 carbon of the cytosine and if you now replace this 5 carbon of cytosine with a nitrogen, then this becomes 5-Azacytosine and it turns out this 5-Azacytosine cannot be methylated because there is no carbon there. When this 5-Azacytosine is incorporated into the DNA, it cannot be methylated and therefore, this can actually inhibit DNA metallization.

So, 5-Azacytosine is a very important repressor of DNA methylation. So, when you treat cells with 5-Azacytosine, DNA methylation is limited and we will now go add, then see how inhibitors of DNA methylation can be used. Not only understand the importance of DNA methylation inside the cells, but they also have very important applications in treatment of cancer and many other diseases and so on and so forth.

So, just remember there is an important compound called 5-Azacytosine which contains nitrogen in the place of carbon in the 5 position and therefore, this 5-Azacytosine cannot be methylated. When you incorporate, when 5-Azacytidine gets incorporated into DNA, it cannot be methylated and therefore, it serves as an inhibitor of DNA methylation.

(Refer Slide Time: 12:56)



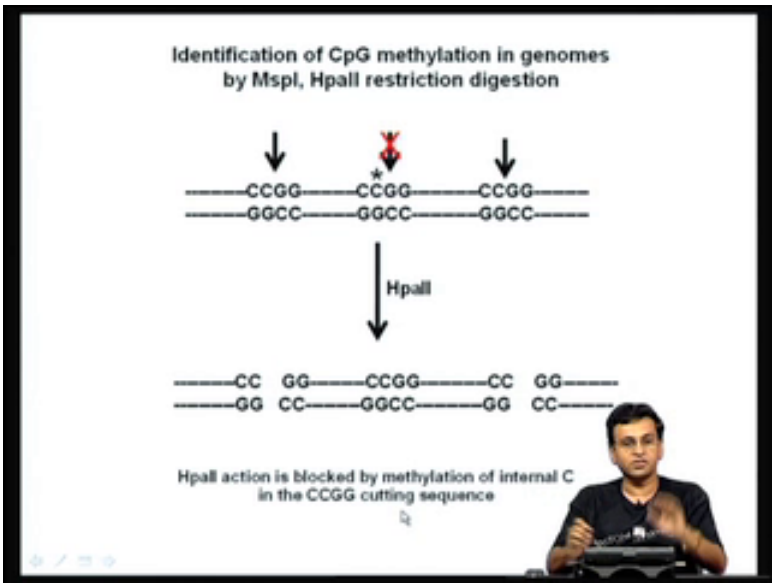
5-Azacytidine is often of abbreviated AZT and it is often used to demonstrate or study the effect of DNA methylation on gene expression. So, we take cells which are not treated with Azacytidine, you take cells which are treated with Azacytidine. Now, what happens is this Azacytidine is now get, AZT now gets incorporated into the DNA and therefore, this DNA cannot be methylated. Therefore, if you now take a look at the gene expression status of this DNA which is not treated with the Azacytidine and cells which are treated with the Azacytidine and compare gene expression pattern, you can clearly say whether DNA methylation has a role in gene expression or not.

So, when incorporated in the DNA in the place of cytosine, de-methylated sites are created because this Azacytidine cannot be methylated by DNA methyl transferases and it turns out when cells are treated with Azacytidine, it results in profound changes in gene expression patterns. This was demonstrated way back in 1980s and 90s, in fact even earlier and in fact, this clearly told that somehow DNA methylation seems to be have playing very important role in regulation of gene expression.

What happens when you treat these cells with this kind of DNA methylation inhibitors? Certain genes which are not normally expressed, remember DNA methylation is a negative regulator of transcription. So, what happens when you incorporate Azacytidine as an inhibitor of DNA methylation? The DNA methylation is prevented and therefore, many genes which are normally whose promoters are normally methylated and not

expressed, when you treat cells with these compounds, these genes are now begin into get expressed because methylation or negative regulation is lost.

(Refer Slide Time: 14:29)

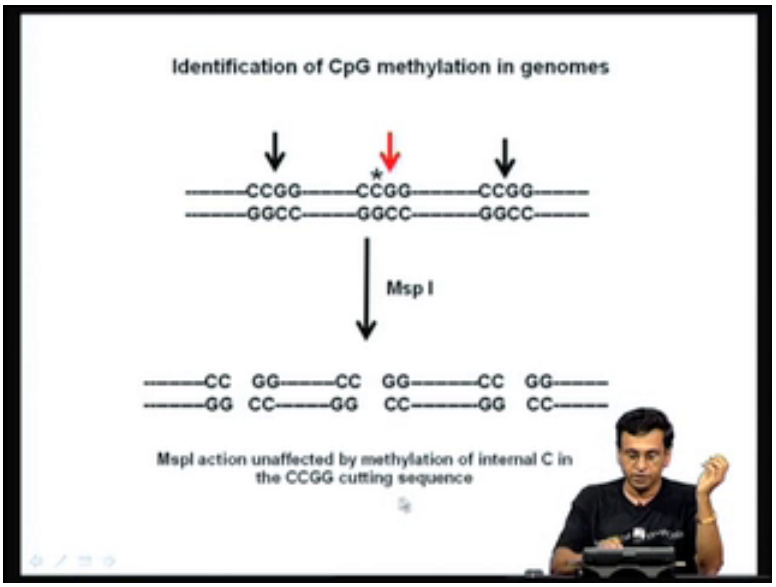


Now, how do you identify that a specific CpG sequences in a specific DNA fragment is getting methylated or not? There is a very simple methodology where this actually used and it is a very well-known methodology. This is actually by using a set of restriction enzymes called Mspl and Hpa2. These are called Isoschizomers that is Msp1 and Hpa2, they recognize the same sequence, both of them recognize CCGG, but there is a small difference in their activity.

Let us see, how exactly they do it. Let us say, we have a DNA fragment here which contains 3CCGG motives. So, there are three sides for three restrictions for the recognition of Msp1 or Hpa2 and let us assume that one of the cytosine is actually methylated among the three. When you now take this DNA and when continued it with Hpa2, when you digest with HpaI2, now Hpa2 does not distinguish whether the cytosine is, sorry Hpa2 does not cleave if the cytosine is methylated within the CCGG. Therefore, all those recognition sequence for Msp and Hpa2 are the same. Hpa2 will not cleave, if the cytosine here is methylated within CCGG whereas, Msp1 does not care whether cytosine is methylated or not. So, when you take a DNA in which one among the 3 CCGG motives is methylated and cut them with Hpa2, Hpa2 will be cleaved only here

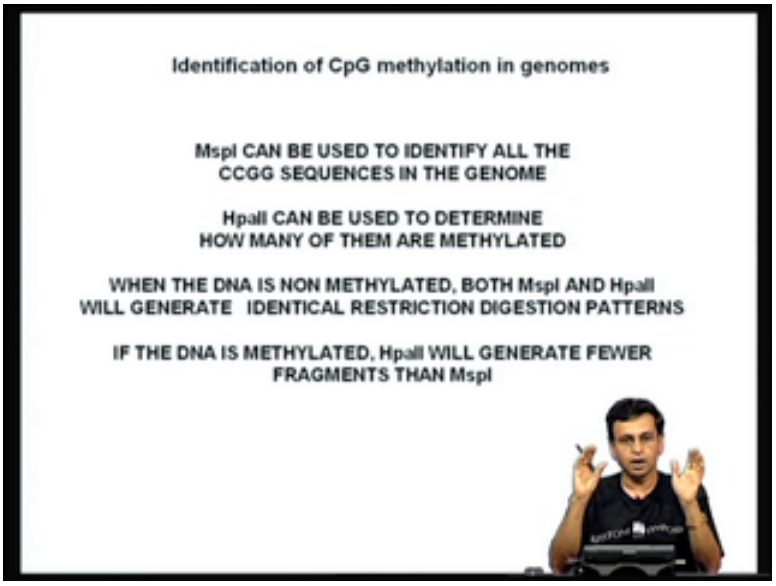
and here, but not here. So, this side does not get cleaved. So, Hpa2 action is blocked by methylation of an internal C in the CCGG cutting sequence.

(Refer Slide Time: 16:02)



Now, let see what happens when you do the same digestion with instead of Hpa2. If you treat the same DNA with Msp1, what happens now is Msp1, if you cut the same DNA with Msp1, Msp1 does not care whether the C is methylated or non-methylated as long as the recognition sequence CCGG. Therefore, it cuts with equal efficiency and therefore, all the three sides will get cleaved by Msp1. So, Msp1 action is unaffected by methylation of the internal C residues in the CCGG cutting sequence.

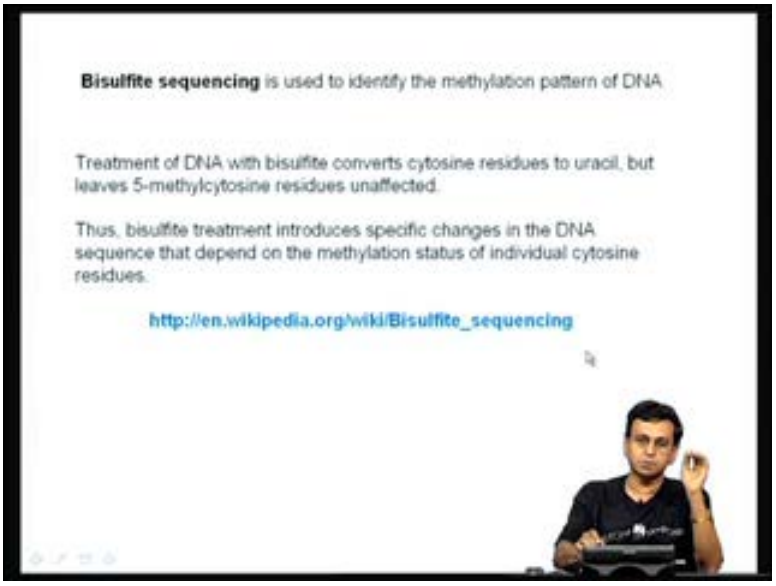
(Refer Slide Time: 16:28)



So, what is the net outcome of all these things? So, when you take the specific DNA in which certain C are methylated within the CCGG sequence and do a duplicate digestion, one with Msp1 and another with Hpa2 and then, compare the restriction digestion pattern by doing a sudden brought or something. For example, in the genome, I want to say a specific region, a specific promoter region is getting methylated or not. So, what you do is you take the genomic DNA, cut them with Msp1 or Hpa2 and then, run it on hydrogen and transfer to nitro cellulous and do a sudden blot with a promoter specific probe.

So, what happens if there is a difference, if there is no methylation in this promoter, you will get an identical restriction pattern for Msp1 and Hpa2, but if there is a difference in the restriction pattern between the Msp1 and Hpa2; that means, there is some methylation and that is why, these two enzymes are behaving differently. So, Hpa2 can actually be used to determine how many CCGG sequence in a promoter region are actually methylated and when the DNA is non-methylated, both Msp1 and Hpa2 will generate identical restriction digestion patterns whereas, the DNA is methylated, Hpa2 will generate fewer fragments than Msp1.

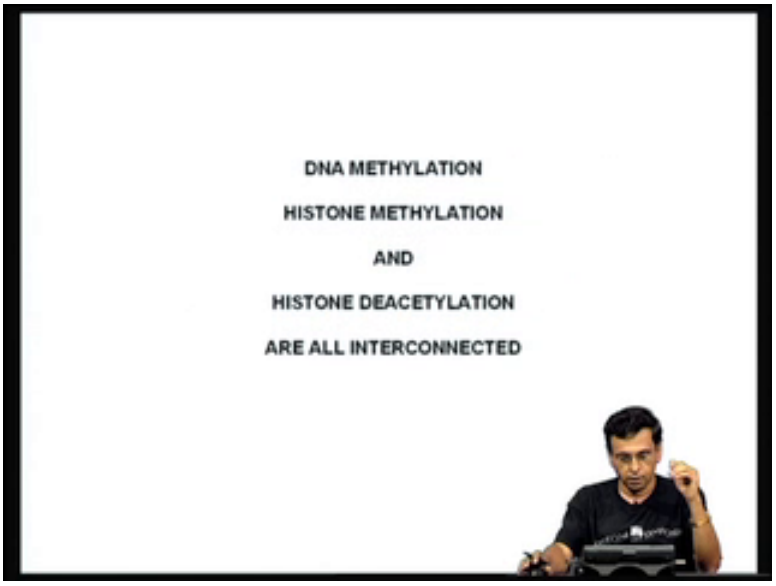
(Refer Slide Time: 17:49)



So, by using this simple restriction digestion using Isoschizomers Msp1 and Hpa2, you can actually find out whether a specific promoter region has been methylated or not. It is very simple, but very efficient and very widely used technique for demonstrating methylation of CpG sequences in a promoter region. There is also much more recent technique which is actually used to determine from methylation of DNA that is actually called as bisulfate sequencing and this is also used to identify methylation patterns of DNA. Now, treatment of DNA with bisulphate converts cytosine residues to uracil, but it does not affect, does not do anything for the 5 methyl cytosine. So, only cytosine is converted into uracil when you treat DNA with bisulfate whereas, methyl cytosine is not affected. Therefore, bisulfate treatment introduces specific changes in the DNA sequence. That depends on the methyl methylation status of individual cytosine residues.

Now, the bisulphate sequencing is a little bit elaborate procedure. I am not going to go and discuss in detail how exactly the bisulfate sequencing is actually carried out to identify whether DNA is methylated or not. This bisulphate sequencing has become a very important rule. In fact, we are going to discuss this in detail in the later classes when you are going to talk about gene silencing heterochromatization so on and so forth. So, I will not discuss bisulphate sequencing in detail at this stage, but you are interested right now, you can actually go to this web site in Wikipedia, where there is a very nice write up or very nice description of how bisulfate sequencing is actually done.

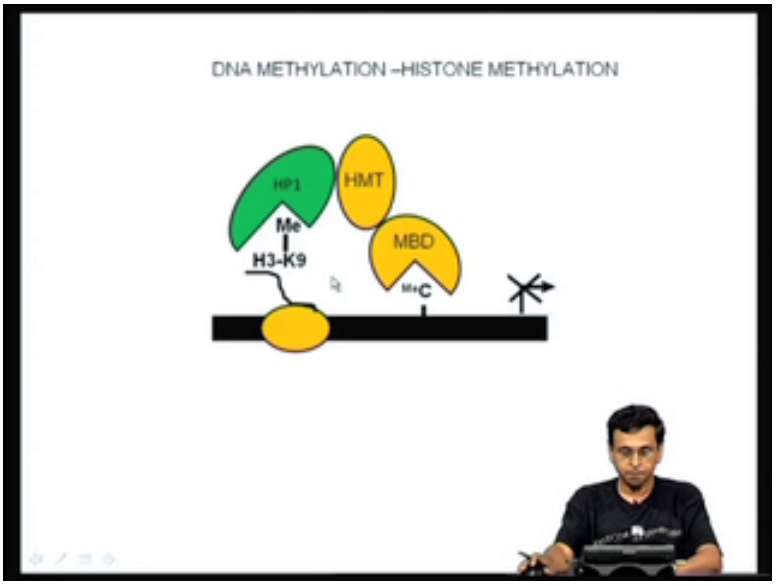
(Refer Slide Time: 19:05)



Now, what I would like to discuss is that the DNA methylation. So, before DNA methylation, we discussed about histone acetylation, histone deacetylation, histone methylation, histone de-methylation, histone phosphorylation. Now, I will talk about DNA methylation, but I am going to tell you that all these are inter-related because I told you just now the mechanism by which DNA methylation inhibits transcription is actually by reputing histone deastylizers to the vicinity of the promoter.

This is how DNA methylation actually regulates gene expression. So, all these mechanisms are actually interconnected. So, do not assume that just because we are discussing a different lecture series histone for acetylation separately, histone deacetylation separately, DNA methylation separately, they all occur in a particular sequence. At any given promoter, all these mechanisms have to actually act to together in order to either turn on or turn off their expression of a particular gene. So, let us spend some time to see how these histone modifications and DNA methylation work in consult with each other to regulate the expression of a particular gene.

(Refer Slide Time: 20:07)

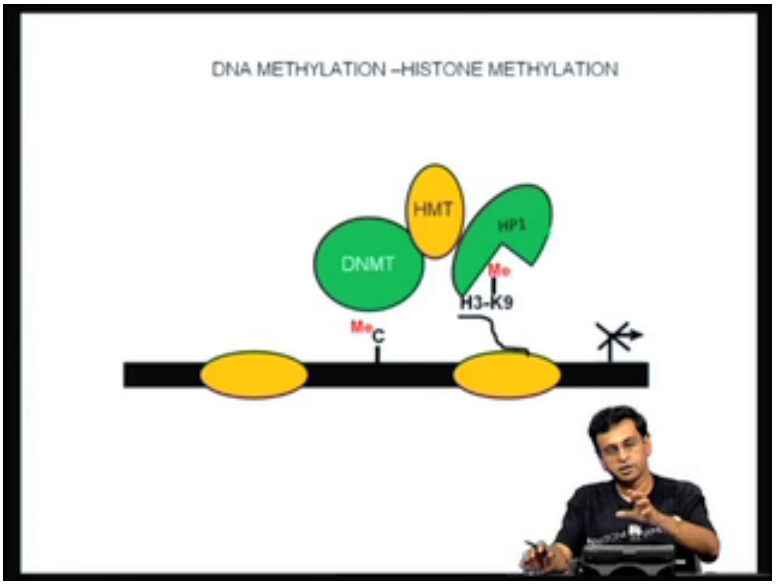


Let us now assume for let us take one particular example. What is the length between histone methylation and DNA methylation? Let us assume we have a promoter in which you have a histone or you have histones here and now we are focusing our attention on histone and the 9 residues of histones. Let us say in this promoter region, there is a particular CpG islands, cytosine is methylated here.

Now, what happens in this methyl cytosine is now recognized by a methyl MBD domain containing protein, that is a methyl CpG recognition protein and these methyl CpG protein can actually recognize, can actually recruit a histone methyl transferase and this histone methylation transferase can now recruit then Hp1. If you remember, earlier the histone methyl transfer actually we discussed very intensely about Hp1 and histone methyl transferase and so on and so forth.

This HP1 now actually can recognize the methyl lysine 8 residues and they can bring about heterochromatization. So, you can see the binding of a methyl MBD containing protein to a methylated promoter can result in the recruitment of histone methyl transferase, which methylates and nine residues of a histone history 3. That is now recognized by the methylated lacing recognizing protein like the Hp1 which will then initiate heterochromatization leading to repression of transcription. So, there is a link between histone methylation and DNA methylation here transcription repression.

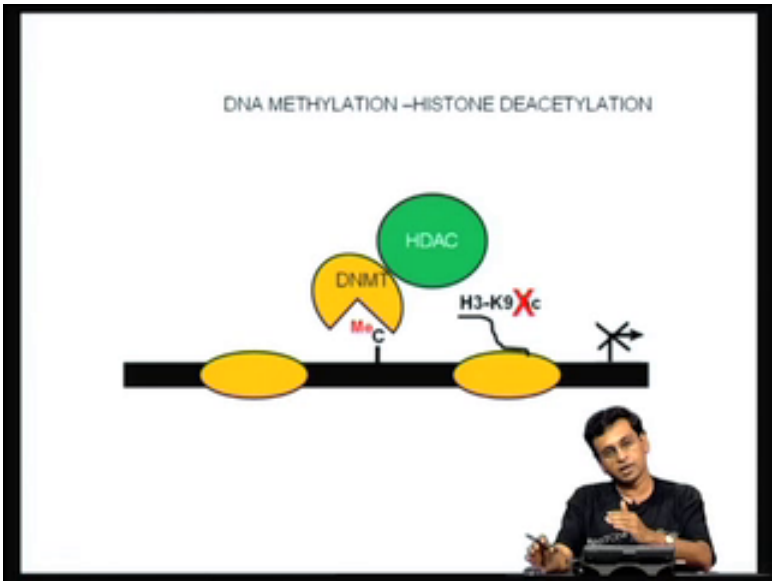
(Refer Slide Time: 21:37)



There is another example. For example, here where instead of, let say for example, a DNA methyl transferase actually methylates a cytosine residue here and then, this DNA methyl transferase can now recruit a histone methyl transferase and then, histone methyl transferase now methylates the lysine residue and these lysine residue methylated lysine is not recognized by Hp1. Therefore, it will be resulting in negative regulation of gene expression.

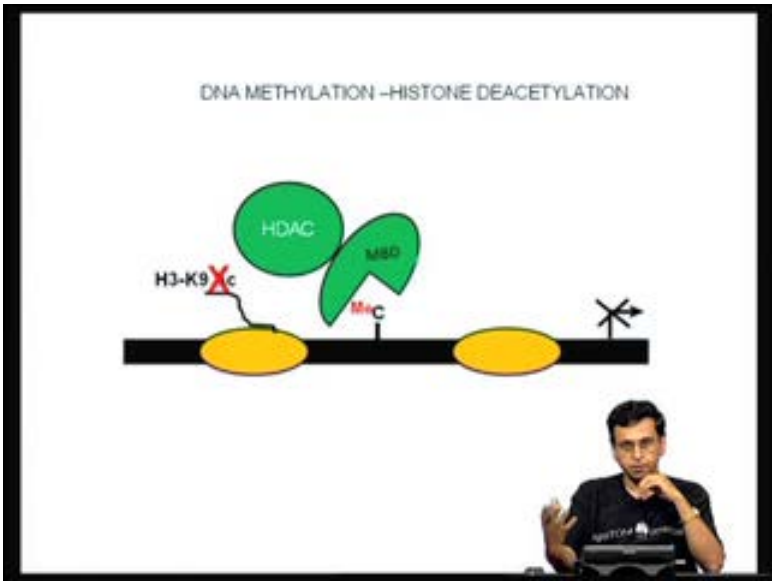
So, in one case, the MBD contain protein interact with histone methyl transferase. Here, the DNA methyl transferase is actually is interacting with histone methyl transferase. Both of them resulting in the repression of gene expression and the examples I am giving you because if you now go into text books, if you going to literature, there are number of examples for each one of these situations. There are some genes where a DNA methyl transferase recognizes a methyl CpG containing DNA and recruits a histone methyl transferase or in another case, an MBD contain protein interacts the histone methyl transferase and brings about the regulation or repression of gene expression.

(Refer Slide Time: 22:39)



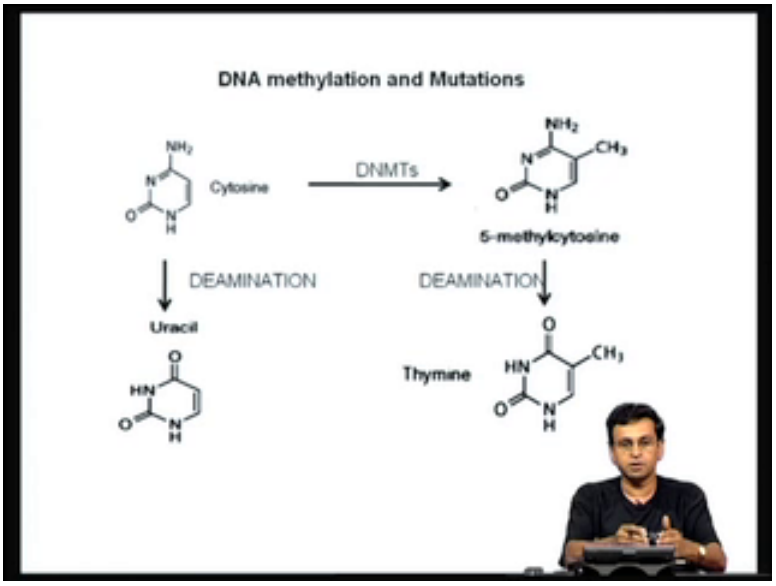
There is another example. Again, methylation of a cytosine by a DNA methyl transferase may result in the recruitment of e Hdac, I histone, deacetylates and these histone de-acetylates now will remove the histone of a specific lysine residue in the H3 and as a result, it can result in negative regulation of transcription. So, in the previous two examples, I have told you the interaction between a DNA methyl transferase and a histone methyl transferase or an MBD contained protein and a histone methyl transferase. Here, we are showing a DNA methyl transferase can also interact with a multi protein complex containing a histone residue de-acetylates and as a result, DNA methylation can lead to histone de-acetylation that can result in the repression of transcription.

(Refer Slide Time: 23:18)



Another example. Here, MBD containing protein can actually recognize a CpG methylated cytosine residue and this MBD protein can now interact with the histone de-acetylates and this can again bring about de-acetylation of specific lysine residues in H 3 or H 4. Therefore, result in repression of transcription. So, I just gave you four examples just to give you an idea that DNA methylation actually serves as a kind of a signal for the recruitment of other histone modifiers, especially negative regulators of histone transcription and either by histone methylation or by histone de-acetylation transcription patterns is portable. So, DNA methylation actually works in conjunction with some of the histone modifying activities.

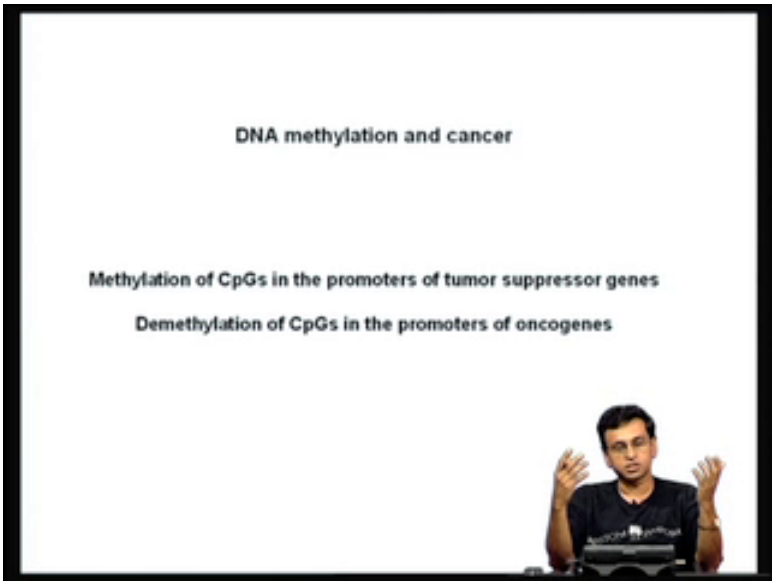
(Refer Slide Time: 24:02)



Now, I just want to take couple of minutes just to explain. The DNA methylation has a very profound effect not only just on gene expression, but also on a number of other physiological processes. In fact, here is an example where I want to show, discuss with you that methylation of cytosine can also lead to mutations, create a mutation. For example, there is what is called as a spontaneous deamination. Now, there is a spontaneous deamination of cytosine and which spontaneous deamination of cytosine can actually convert cytosine into uracil, whereas if this cytosine by the action of these DNA methyl transferases is converted into methyl cytosine and this methyl cytosine undergoes spontaneous deamination, you can see methyl cytosine deamination will result in the formation of thymine.

Now, there are many enzymes. For example, there are something like a uracil deglycosiline, many other enzymes around this DNA repair path way. We will not be going into those details because our focus is primarily on gene regulation, but just suffice to know at this point that cytosine methylation by DNA methyl transferase can also lead to specific mutations, wherein a cytosine normally on a deamination becomes uracil, whereas a methyl cytosine on deamination gives rise to thymine.

(Refer Slide Time: 25:17)



Now, DNA methylation has a very profound influence in cancer. It turns out, I told you very clearly when you have the CpG islands in the promoter region of a gene and when these CpG islands are de-methylated or present in a non-methylated form, then it is very likely that this promoter is active. This gene is actively transcribed whereas, the CpG residues in the promoter of this gene if it is in methylated and it is very likely that gene is not active, that gene is repressed. It turns out in many of the cases methylation of the CpG's in the promoter regions of tumor suppressor genes, if these genes turn out to be tumor suppressor genes.

As you know cancer is the result of two groups of genes called as oncogenes and tumor suppressor genes. Now, oncogenes are those if they are there actually, it is like an automobile having a break and an accelerator. So, you can regulate the speed of an automobile, either by pressing a break or pressing an accelerator. So, oncogene is like an accelerator. So, they promote self-proliferation. So, if you over expressed oncogenes, the cells derived faster. Just like if you press an accelerator, a car drives much faster, if you express oncogenes very highly, the cells will derive. So, they promote cell proliferation whereas, the tumor suppressor genes act like breaks. So, they actually control, decrease the speed of an automobile or they decrease the speed of this cell proliferation. So, you remember. So, what happens if you now methylate the promoters of a tumor suppressor gene? So, the tumor suppressor genes will not be expressed. That means, the breaks would not function properly and therefore, cells start dividing uncontrollably now.

On the other hand, if you now de-methylate the CpG residues of oncogenes, then the oncogenes will start expressing at very high levels. That means, you are pressing the accelerator very fast and therefore, the automobile goes very fast. That means, cells will start dividing very fast and resulting in uncontrolled cell proliferation. In many case, it has turned out in specific cancers. When they found out in particular, in many tumor types the tumors cell line either has methylation of CpG residue is tumor suppressor genes or de-methylation of promoters of the oncogenes.

In fact, this is where the compounds like Azacytidine became very important anti-cancer agents because in cases where the tumor suppressor genes are actually methylated. If you now take such kind of cancer cells and treat them as Azacytidine, what happens is Azacytidine will get now incorporated into the DNA. Therefore, the Azacytidine cannot be methylated and therefore, these tumor suppressor promoters will now be active. Therefore, the tumor suppressor genes are expressed and therefore, you can press the breaks very fast and cell-proliferation can be inhibited.

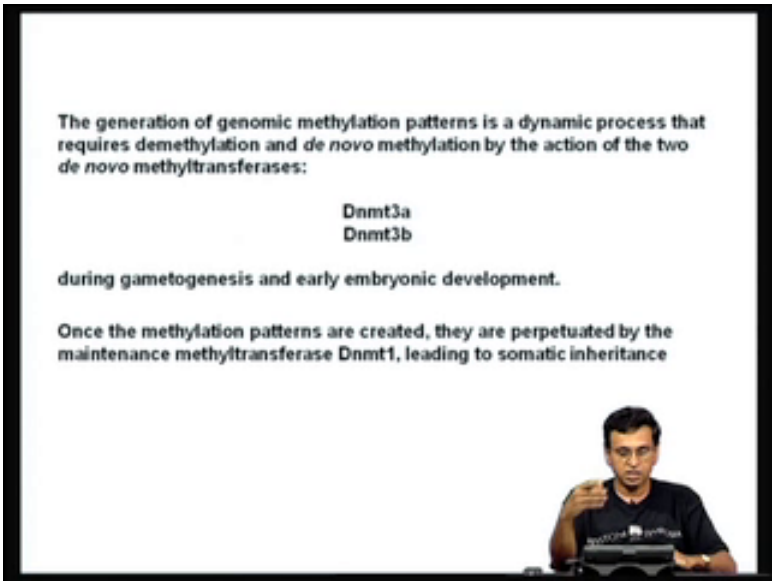
So, by simply treating certain tumor cells with inhibitors of DNA methylation, such as Azacytidine, you can actually inhibit the proliferation of cancer cells. So, this is one important application that came out of the studies on DNA methylation, where certain inhibitors of DNA methylation can actually used for anti-cancer therapy.

(Refer Slide Time: 28:15)

Now, DNA methylation also has a very profound influence on a number of other processes. I have listed some of them here. For example, genomic imprinting, X chromosome inactivation, gene silencing, regulation of tissue-specific expression, maintenance of heterochromatin.

Now, we are going to discuss each one of these topics in detail in the few classes which come in the next series of lectures, but today, we are going to focus only on the regulation of gene expression on DNA methylation, but remember the DNA methylation has a very important role not only in regulation of gene expression, but also in a number of cellular processes.

(Refer Slide Time: 28:52)

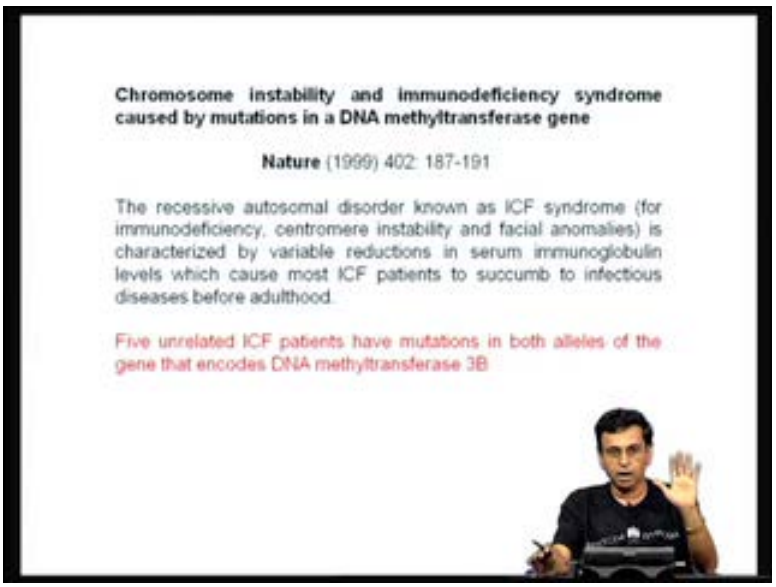


For example, the generation of genomic DNA methylation pattern is a very dynamic process and it requires de-methylation and de novo methylation by the action of two de novo methyl transferases. So, if you take for example at any given time, either from development to the adult, there is a nice homeostasis between methylation of DNA in de-methylation. So, certain regions of a chromatin are kept in a methylation state and certain genes of a chromatin kept in a de-methylated state. So, the generation of those genomic methyl patterns is a very dynamic process and depends on the interaction between methyl transferases and demethylizes.

There are actually two DNA methyl transferases called DnMt3a and DnMt3b which are actually involved for these maintenances of this methylated chromatin in bio, especially

during gametogenesis and early embryonic development. Now, one specific methylation pattern we have created, they are perpetuated by maintenance methyl transferase called DnMt1 leading to their somatic inheritance. So, basically, what I am trying to say that there are number of DNA chemical transferases. Some of them actually involved in the maintenance of the DNA methylation transferase and some of them are involved in specific methylation of DNA at specific stages of development.

(Refer Slide Time: 30:05)

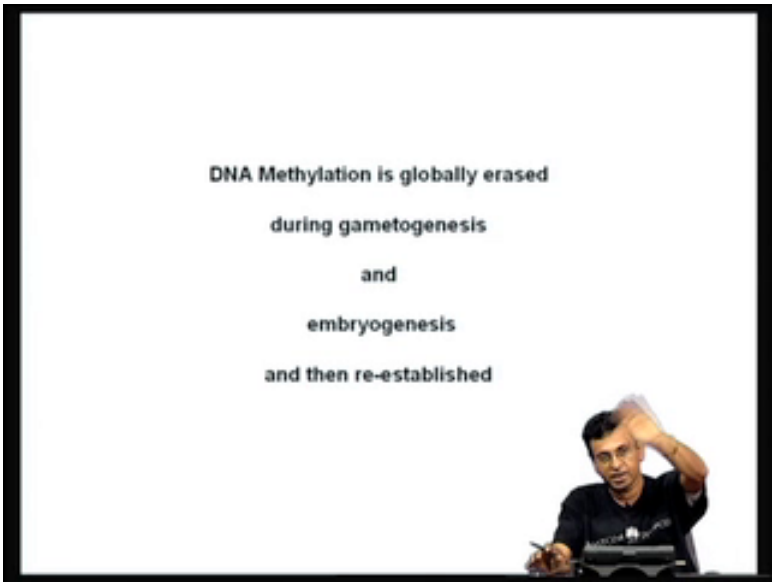


In fact, people have found out when you have mutations in specific DNA methyl transferases, it can actually lead to certain; it can result in the manifestation of certain genetic disorders clearly indicating that DNA methylation is a very important phenomenon. Mutations in enzymes which are catalyzing DNA methylation can result in certain genetic disorders.

Here is one example, where chromosomal instability and immuno-deficiency syndrome caused by mutations in a DNA methyl transferase gene. In this paper, the authors have clearly shown that there is recessive disorder called as ICF syndrome, which stands for Immunodeficiency Centromere Instability and Facial Anomalies. It is kind of a genetic disorder and these patients which there clinical manifestation is, they contain varying, they contain a drastic reduction in serum immunoglobulin levels and they also succumb to infection diseases before they reach adulthood.

They are highly susceptible to number of infection diseases and if they took this ICF patients and then, looked at where the mutation is and they found that the mutations are actually in the gene that encodes for a DNA methyl transferase 3B clearly indicating that if you tamper with DNA methylation enzymes, it can manifest as genetic disorders cleaving that DNA methylation regulates the expression of very important genes that are required for normal functioning of or normal homeostasis during the life span of an adult.

(Refer Slide Time: 31:32)



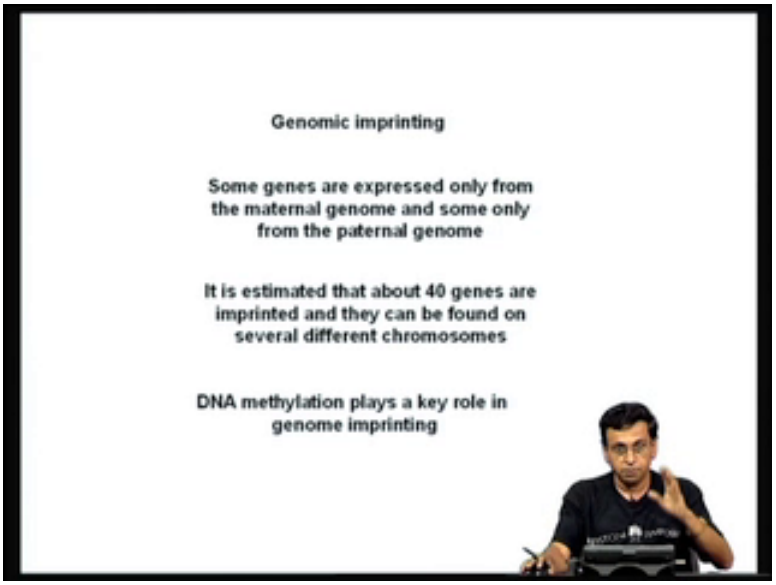
Very importantly, DNA methylation is globally erased during gametogenesis and embryogenesis and is then, re-established. This I am mentioning here, although it is not directly related to gene expression because today, we are talking about what are called as stem cells, right. Now, people are talking about IPS cells, people are talking about a well differentiated skin cell or well differentiated fibroblast and converting them into an embryonic stem cell. Then, in all these cases what is happening is that many times, it is not been successful all these differentiation if you take a differentiated cell and want to convert them into a total dependent embryonic stem cell.

Although, people are trying variety systems, somewhat it is not been very highly successful mainly because the methylation pattern of DNA in a highly differentiated cell has to be completely erased if this differentiated cells has to become a stem cell because during normal development at the time of gametogenesis, the de-methylation map is

completely erased and then, when the embryonic development, a DNA methylation pattern is totally re-established.

So, in addition to the genetic code, this DNA methylation pattern also plays a very important role during development and differentiation. That is why, these conversion of these adult cell into stem cells, all those things have lot of problems because the signature, the DNA methyl signature of an adult differentiated cell is very different from the DNA methylation signature of an embryonic stem cells. So, if you have to convert in adults differentiated cells into embryonic stem cell, you have to make sure that this DNA methylation signature which is there in the differential completely erased or has been completely modified to become to that of a stem cell and then only, it becomes successful and here are what is called as a very important phenomenon which is actually called as genomic imprinting.

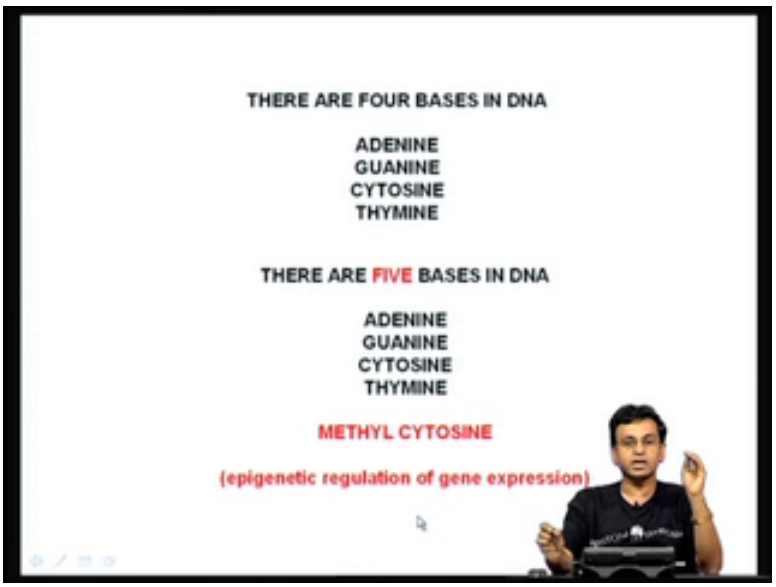
(Refer Slide Time: 33:16)



When in the case of genomic imprinting some genes is expressed only from the maternal genome whereas, certain are expressed only from a paternal genome and it is estimated that about 40 genes are imprinted. They can be found on several different chromosomes. So, this is actually called a genomic imprinting. That means, you have both the maternal allele and paternal allele, but when they fuse and form as a zygote, certain genes, only the maternal allele is imprinted by impressed, whereas in case of certain genes, only the paternal allele is impressed. This has implication especially when you have a XX or XY

phenotype. We will not go into the details, but the point I want to tell you is that the DNA methylation actually plays a very important role in this kind of a genomic imprinting. So, DNA methylation is very important for this kind of a maternal and paternal inheritance of genes.

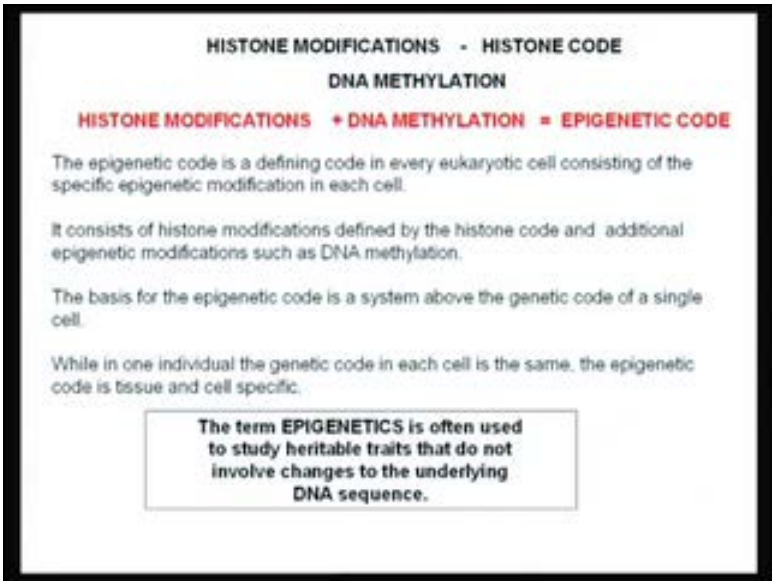
(Refer Slide Time: 34:13)



So, the reason why I am emphasizing on the role of DNA methylation in all these experiments is that, so far we have studied that there are only about four bases in DNA, namely adenine, guanine, cytosine and thymine, but people realize the importance of cytosine or importance of methyl cytosine and a number of physiological processes. People have now realized that. In fact, methyl cytosine actually now becomes a fifth base.

Because of its physiological importance in a number of cellular process, methyl cytosine is actually now recognized as a fifth important base for a number of physiological processes. So, in addition to adenine, guanine, cytosine, thymine, methyl, cytosine is also now become very important component of DNA because of it various variable role in number of cellular processes. Because of this DNA methylation, a new term or a new terminology has been coined to understand gene expression changes that take place without actually change in the genetic sequence of an organism and that is what is known as the epigenetic regulation of gene expression.

(Refer Slide Time: 35:22)



So far we discussed histone modification that has resulted in what is called as a histone code. In addition to that, we are all very familiar with the genetic code which is actually determined by the sequence of bases in a DNA and which is inherited from generation after generation. After discussing all the histone post translation modifications and then, understanding how post translation modifications of histones play very important role in the regulation expression, we brought in what is called the histone code in the last few classes, but now after studying the role of DNA methylation and how DNA methylation regulates a number of physiological processes, DNA methylation in combine.

Combination of histone modifications now constitute what is called as a epigenetic code. They have genetic code, histone code. Now, we are talking about epigenetic code. Now, what is an epigenetic code? The epigenetic code is a defining code in every eukaryotic cell consisting of the specific epigenetic modification. In the each cell like just I told you, the epigenetic modifications present in. For example, in ear cell may be different from the epigenetic modification and a brain cell, although the genetic code is the same.

Similarly, the epigenetic footprint or of a zygote is very different from the epigenetic modification. Let us say after may be 50 or 60 divisions, cell divisions, it is all because there is an active programming involving DNA methylation or histone modifications. This epigenetic map changes in different cell types and also in different stages during

development. So, this study of inheritance of certain trades which is not exactly is dependent on the DNA sequence is known as epigenetics.

So, this epigenetic modifications actually consists of histone modifications as defined by the histone code and in addition, epigenetic modifications involving DNA methylation. So, DNA methylation in conjunction histone modifications, such as histone acetylation, de-acetylation, phosphorylation and so on and so forth can actually bring over what is called as a epigenetic regulation of gene expression. These epigenetic maps differ in different cell types and also during different stages of development.

So, the basis for a epigenetic code in the system above the genetic code of a single cell, this is very important for it. That is why I am emphasizing. Very importantly in addition to genetic code, you have several new codes that are coming up. One is called a histone code based on the post translation modifications of histones. Although, where many people argue that the histone code is not really perfect.

Remember genetic code by actually look at the exact sequence of DNA; you can exactly predict what kind of a protein will be made. The genetic you can actually predict what kind of a protein will be made, what kind of phenotype you will get depending upon the genetic code, but the histone code, we still have not understood this histone code epigenetic code. So, well perfectly that we cannot really predict a specific phenotype depending upon a specific histone code or specific epigenetic code. So, we still have a long way to go, understand. So, many people think that it is too early or too premature to actually use these terms like histone code and epigenetic code because you cannot really predict what a phenotype is going to be depending upon a specific code.

So, when one individual genetic code in each cell is the same, the epigenetic code is tissue and cell specific. So, this is very important part I am emphasizing. In fact, this is the outcome of all the 10 lectures we have done so far. Based on all these 10 lectures, we initially started with describing how transcription regulation takes place in the naked DNA. Remember, if you had been studying gene regulation, only a naked DNA, this new concept of histone code and epigenetic code would not have come at all, but the moment people realized that gene regulation actually happens on chromatin templates and the chromatin structure is intricate a link to the regulation of gene expression and one has to modify histones and one has to modify DNA by modifying cytosine residues, you can

actually regulate gene expression. These are very profound observations. They have actually now brought on the new concepts of what is called as epigenetics or epigenetic regulation of gene expression.

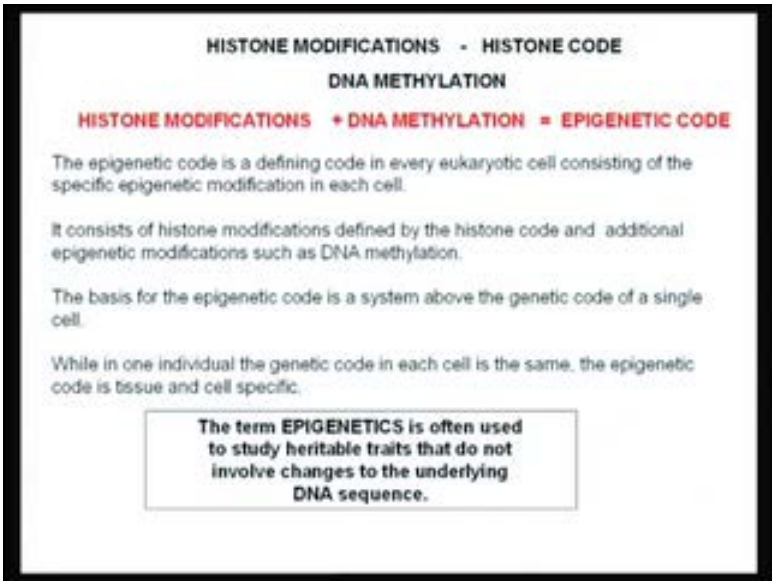
It is now very becoming clear that the methylation pattern, the cytosol methylation patterns or histone modification patterns, they vary from different cells in adult organism and also varies very profoundly during embryonic development. Therefore, although all the cells of our body has the same genetic code, the histone code and epigenetic code of these different cell types may vary from cell to cell and also from different stages of development.

So, the term epigenetic is often used to study heritable trades that do not involve changes in the underlined DNA sequence. It is very important for you to know this. So, what I have told you so far is, let us just take a very brief review of all the topics we have studied so far. The important concept that I would like to convey up to this stage of our lecture series because this marks a very important part of this entire lecture series. We have discussed very briefly how transcription takes place on a DNA template in the absence of, in the absence of histones. That is how we started with.

We discussed that there is an RNA polymerize enzyme and this RNA polymerize enzyme in order to go and bind to core promoter region requires the help of general transcription factors. Then, these general transcription factors include tf2, atf2, bde and f and is with the help of the general transcription factor that the RNA polymerize 2 is brought to the core promoter region. That is how transcription essentially takes place. Then, we discussed very importantly the core promoter region.

Many of you might have studied. The core promoter actually means only Tata box, but we discussed it is not just the Tata box. There are many variations of the core promoter limits like the initiator, BRE and so on and so forth and these variation in the core promoter elements can itself bring about differential gene regulation. Then, we discussed although people talk about in textbooks about only general transcription factor means tf2, abde and f, there are variations within the general transcription factors there. For example, there are many tiffs, the TBP associated factors which are expressed in a cell type specific manners and they actually bring about differential gene regulation.

(Refer Slide Time: 35:22)



So, variations within the general transcription factors itself can bring about differential gene regulation. Then, we discussed that it is not a core promoter region. There are other upstream sequences which play a very important role in the regulation of gene expression and variations in this upstream elements, which actually serve as binding sets for various transcription factors that actually contribute to different levels of activation or repression of gene expression.

Then, we discussed how specific eukaryotic transcriptional activators or transcriptional repressors which basically contain two major domains called as a DNA binding domains and transcription activation domains and through this DNA binding domains, they interact with specific sis acting elements and through that transactivation domains, interacts with the components of the pre-recession complex and then, facilitate faster recruitment or slower recruitment of RNA polymerizes leading to either activation or repression transcription variations in the rate of transcription initiation.

We also discussed what the major kind of DNA binding domains that are present in eukaryotic transcription factors are. For example, what are called as a leucine zipper, sorry helix stone, helix motifs, you have helix luke helix motifs, you have zinc finger motifs, you have leucine zipper motifs and through some of the major DNA binding domain motifs, these transcription factors bind to specific sequence of DNA. We also discussed a number of experimental techniques like genability shift acids and DNA and

foot printing etcetera which are normally used for identifying transcription factor binding sets in various promoters, promoter elements.

We also then discussed the functional analysis which are actually used for studying eukaryotic gene regulation. We discussed about cell transfection acids, wherein you can take promoters link to reporter genes and you can make the different deletions of the promoter regions. Then, by asking the reporter gene activity, you can actually measure the rate of the strength of a promoter by transfecting them into cell lines and then, we discussed in very detail about what is called as cell free transcription studies, wherein you can prepare nuclear extracts which are transcriptionally competent. A lot of effect as such gone to describe for the development of this self-transcription asses and using these kinds of nuclear extracts, people who have put naked DNA templates into this nuclear extracts and then, started identifying transcription factors and promoter regions which are essential for transcription situation.

Then, we came into a very important fascist of gene regulation, where we discussed DNA is not actually naked. DNA is actually started with histones in the form of nucleosomes and therefore, studying transcription regulation on naked DNA templates is of no use. We actually have to study gene regulation in the context of chromatin templates and that is why, you brought in the concept of histone modifications. We discussed very briefly how there are number of enzymes which either had an acetyl group. These are called as hats or histone acetyl transferases. There are enzymes called H dax, histone, deacetylizes, which remove histones from chromatin templates leading to negative regulation of gene expression and there is a very fine equilibrium.

You had some H dax and these dynamic equilibrium between the two actual result describes whether gene has to be activated or a gene has to be repressed. Then, we talked about histone methylation and histone de-methylises. Histone methylation generally results in the repression of transcription whereas, de-methylation results in the activation of transcription. There we brought in the concept of histone phosphorylation and with all the studies of this histones and their role in gene regulation, we brought in what is called as a histone code and we discussed how important is the histone code for the regulation of gene regulation.

Then, now recently we have brought in the concept of DNA methylation in this class, wherein we have discussed how there are enzymes called as DNA methyl transferases which are involved in methylation of specific DNA sequences, especially the CpG motives in the promoter DNA. When this CpG islands or the CpG motives are methylated by DNA methyl transferases, this methylated cytosines are recognized by specific enzymes which contain called as the methyl binding domains. Through this methyl binding domains, these proteins are recognized. This methyl cytosines, they then go and recruit specific histone as diastalsis or histone methyl transferases. So, DNA methylation in conjunction with histone modifying enzymes can actually lead to repression of transcription.

The most important thing that we have learnt in this class is that DNA methylation has a very profound role in a number of cellular processes in addition to regulation of gene expression and that is why, the methyl cytosine is actually now recognized as the fifth base in addition to adenine, guanine, cytosine and thymine. In fact, a number of research is now going on a very fast phase to identify small molecules which can actually inhibit DNA methylation. In many cases, people have actually shown many specific cancer types can be the proliferation of this cancer cell can be inhibited by treating these tumors with specific inhibitors of DNA methylation. So, understanding the mechanism of DNA methylation has had a very profound effect on cancer research.

A number of these compounds which are actually DNA methylation inhibitors are being tested in human clinical trials to see whether, they can be developed as therapeutic agents for anti-cancer therapy. In the last part, we have gone through, discussed to introduce a new concept or a new component called as an epigenetic code. DNA methylation in conjunction with the histone modifications or histone post trans modifications together cost what is cost epigenetic code.

An important take home message that I have given from these lectures is that the genetic code of an organism is probably the same in all most all the cells of a body. You take a liver cell or a brain cell or a muscle cell or during development, nothing happens. The gene sequence is the same, but the differential gene expression that we see is actually because of epigenetic modifications that are taking place the reason why certain genes are expressing. Liver cell are not being expressed in a muscle cell, but a different set of genes are getting expressed during development and during various tissue types is

mainly because these kinds of epigenetic changes that are taking place in these different cell types are different.

Therefore, understanding this epigenetic modifications have a very profound influence in understanding gene regulation. In fact, people are actually started now identifying how even nutrition affects these kind of a epigenetic regulation of gene expression. For example, DNA methylation. Now, the important thing is, there are many of our footsteps actually if you now take certain methyl donors. If methyl donors are present in many of our footsteps that we take in, then it actually provokes histone methylation or DNA methylation. Therefore, the gene expression tangles actually changes. So, if you take certain mice and then, feed them with a diet which are actually methane methyl donors. For example, certain kind of folded derivatives. So, if you remember many of this methylation reaction involves tetra hydro-folate whereas, the methyl tetra hydro-folate and tetra hydra-florate plays a very important role in many of this methylation reactions.

So, if you know to treat animals or if you take animals and feed them with a food which is very rich in this methyl donors and compare them with this mice, with those which are not fed, this methyl donors there are profound changes in the gene regulation. So, people are now advocating the kind of diet that we are taking. When you actually smoke, you are actually taking of many methyl donors and these methyl donors now when you take in, they actually promote methylation of histones or methylation of DNA. Therefore, it can have a profound effect on gene expression.

So, how environmental factors are affecting our gene regulation or affecting our gene expression is primarily determined by epigenetic regulation of gene expression. So, I think I will conclude here primarily to tell you that DNA methylation has emerged as a very important mechanism by which regulation of gene expression takes place in eukaryotes. Methyl cytosine has emerged as very important base in addition to the four bases that are very well recognized and these studies on DNA methylation in combination with histone modifications has signaled a new era of research and this is what is known as epigenetic regulation of gene expression.

This epigenetic has now become very important, not only to understand regulation of gene expression, but also to chromatin structure. In fact, we have discussed now that if you look at for example, a hetero chromatin, where the chromatin is very tightly bound

to DNA whether the histones are very tightly bound to DNA and you look at the euchromatin and you know a specific methyl, specific methylation or specific deacetylation of histones. There are signatures which are very characters of euchromatin and there are very characteristic signatures of histones which are characters of hetero chromatin and the same way, chromatin regions which are not very highly expressed actually contain methyl cytosines or DNA methylated. DNA is highly methylated and DNA methylation in conjunction histone modification seems to be playing a very important role in determining whether a chromatin in that particular region is going to be transcriptionally active or chromatin in that region is going to be transcriptionally inactive.

There is a very dynamic equilibrium between enzymes which modify histones as well as enzymes which modify DNA. All these things in concert bring about what is called as an epigenetic regulation of gene expression. So, I think with this lecture series, I have actually concluded one part of regulation of gene expression, wherein we brought in the concept of histones and the concept about DNA methylation and how all these factors together are contributing to regulation of gene expression eukaryotes.

What we will do in the next class is to bring about another important component of gene regulation, namely how chromatin structure is modified or chromatin structure can also be modified by what is called as a chromatin remodeling enzymes ATP dependent chromatin remodeling and its effects on gene regulation. So, what I am going to tell you in the next class is that in addition to modification of histone tales, you can also, there are also enzymes which can move nucleosomes or which can move histones on the nucleosomes by ATP driven process and this ATP dependent remodeling of chromatin also plays a very important role in the regulation of gene expression. So, chromatin remodeling and these effects on gene regulation will be discussed in the next class. Thanks