Eukaryotic Gene Expression: Basic and Benefits

Prof. P N Rangarajan
Department of Biochemistry
Indian Institute of Science, Bangalore

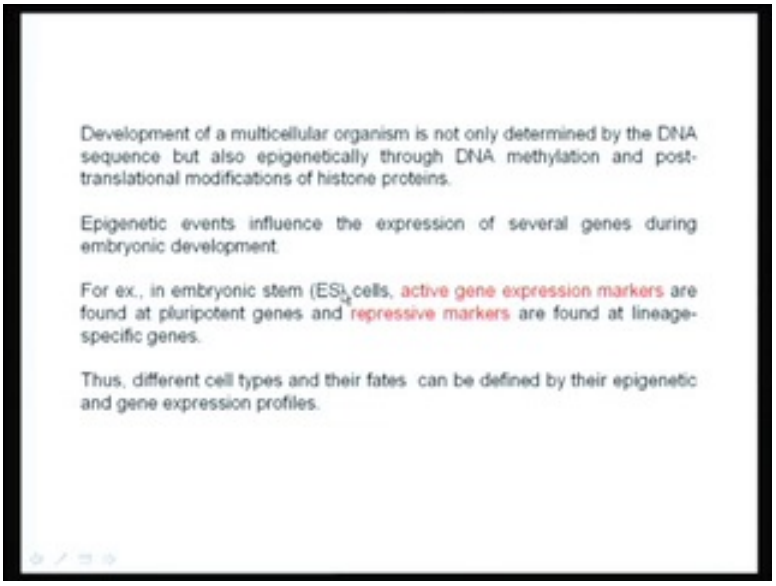
Lecture No. # 27
Epigenetic regulation of gene expression during development

Welcome to this lecture series on eukaryotic gene expression basics and benefits. This is the 27th lecture in this series. The last few lectures we have been discussing the role of gene expression during embryonic development; we have been using Drosophila as a model system. I have tried to explain how maternal genes as well as zygotic genes and their differential expression, ultimately is responsible for the development of the single cell zygote into a multi cellular organism. How spatial and temporal expressions of genes play a very important role in the regulation of gene expression at different stages during development? And we discussed things, like, maternal, in maternal effect genes, then segmentation genes. Then we talked about homeotic genes and how all the expression of these various genes during various stages and development, ultimately results in the development of an adult organism.

The point I had made in the last 3 lectures in understand, trying to understand the role of gene expression during development is the DNA sequence, as well as the spatial and temporal expression of the various genes, is plays a very important role during embryonic development. But I do not want to give an impression, that it is only the DNA sequence and it is only the transcription of this sequence information from the gene into messenger-RNA and ultimately, translate in to the protein that alone is responsible for the development of an adult organism from a fertilized egg.

What I would like to tell you today is that in addition to this genetic basis, in addition to the DNA sequence information, epigenetic also plays a very, very important role in the regulation of embryonic development. So, let us try to understand what is this epigenetic regulation of gene expression during embryonic development and how important is for the proper development of embryo?

(Refer Slide Time: 02:21)



Now, development of a multicellular organism is not only determined by the DNA sequence, but also epigenetically through DNA methylation and post-translational modifications of histones. And in the last, the some of the earlier classes, when we discussed about the role of chromatin in during the regulation of gene expression, I had extensively told you, how the histone code, histone post-translational modifications, as well as, DNA methylation play a very, very important role in the regulation of gene expression. And in fact, based on this, we had discussed about what is called as histone code and epigenetic code and so on, so forth. So, what we will discuss today is that how is this modification of histones as well as DNA methylation play a very, very important role in the regulation of embryonic development as well.

Epigenetic events influence the expression of several genes during embryonic development. If you take an example, the embryonic stem cells, that is, these are the cells in the stage of blastula, which are pluripotent and which have, which have the capability to develop into full-fledged organisms, the embryonic stem cells, if you take the embryonic stem cells, active gene expression markers are usually found in pluripotent genes, whereas repressive markers are found in the lineage specific genes, clearly indicating, that the epigenetic markers are pretty different in different cells of the embryo.

Thus different cell types and their fates can be defined by their epigenetic and gene expression profiles.

(Refer Slide Time: 3:46)

Now, as I just mentioned, we had had extensive discussions about the various post-translational modification histones undergo and how these post-translation modifications play a very, very important role in the regulation of gene expression.

For example, we have discussed how acetylation of histones by group of enzymes, called histone acetyl transferases, can result in the activation of gene expression or is a kind of a fingerprint for actively transcribed genes, whereas well, histones are deacetylated, it usually signals repression of gene expression in addition to this acetylation. Deacetylation histones also undergo several modifications, like phosphorylation, methylation, ubiquitination and so on, so forth.

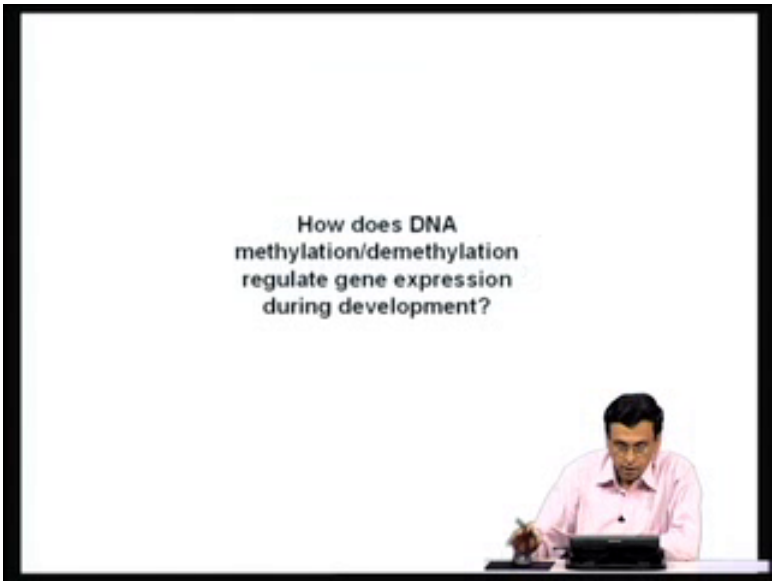
And based on all these studies, people have actually identified, what are called as signature motifs, that is, today I can say, if in, in the promoter region of gene, if I find a histone H4 whose lies in residues 6 is acetylated, then I can say, this is likely to be actively transcribed. So, this is what is called the histone code hypothesis. So, based on studying a number of genes and the kind of histone modification, that are associated with the expression study of genes, people are now, what is called as histone hypothesis.

Some of the examples have shown here, for example, if you have acetylation of histone H3 or H4, usually it signifies, that these genes are likely to be described very actively. Similarly, deactivation of histones usually signals inactive genes and in the nucleus, if we look at the heterochromatin region, which basically contains inactive genes, usually histones remain deactivated in the heterochromatin region of the nucleus. Similarly, methylation of the lysine 4 residue of H3 usually associated with actively transcribed genes. Whereas, methylation of the lysine 9 residue of the histone H3 usually signifies, that this is involved in genes silencing. And genes in this region are not likely to be expressed very actively. So, this, based on these kinds of studies, we have discussed what is called as histone code hypothesis and after studying the histone code hypothesis, we would also discuss, in addition to histone modifications, DNA modification, especially the methylation of cytosine also plays a very, very important role in the regulation of gene expression.

We have discussed how 5-methyl cytosine, because of its importance and regulation of gene expression, has now come to be known as the 5th base in addition to adrenal, thymine, guanine, cytosine and together the histone code, and the DNA methylation patterns constitute, what is called as a epigenetic code.

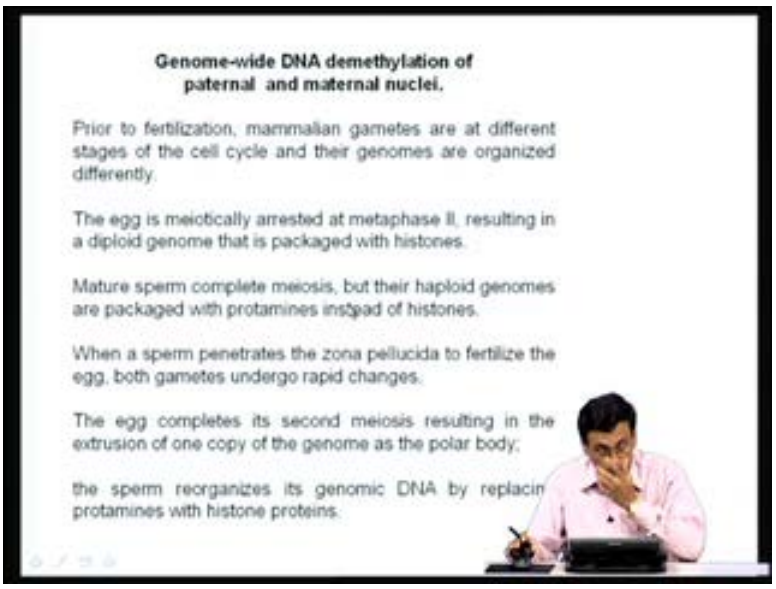
So, what we have learnt so far in our earlier lectures is that even without altering the DNA sequence information, you can alter gene expression by either bringing about post-translation modification of histones or by altering the methylation pattern of the DNA. This is what is called as the epigenetic regulation of gene expression.

(Refer Slide Time: 06:37)



Now, let us try to understand, how does this DNA methylation or DNA, methyl, demethylation of cytosine residues of DNA regulate gene expression during development?

(Refer Slide Time: 06:49)



Now, let us try to understand, what kind of methylation changes takes place during development, both in the maternal, as well as the paternal genome. Now, let us try to start our discussion. Just before fertilization, what is the status of this methylation status, as well as the genome status in the, for the sperm as well as oocyte. The prior to

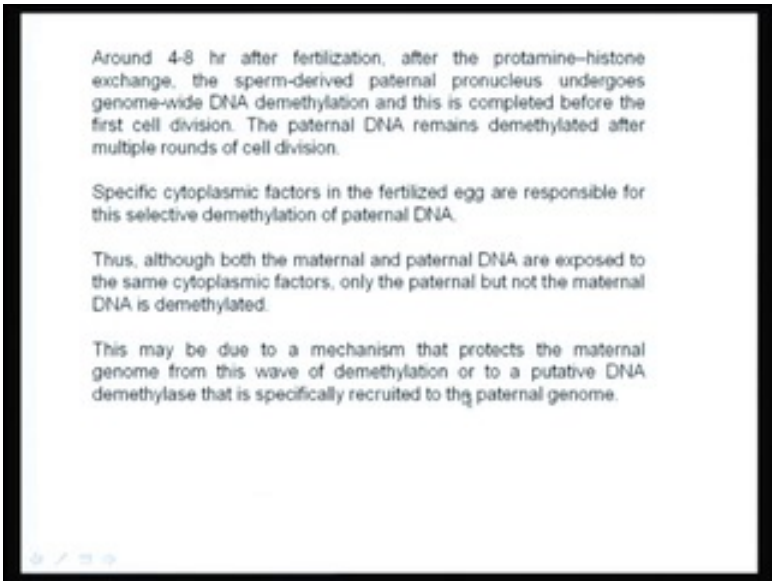
fertilization, mammalian gametes are at different stages of the cell cycle and their genomes are organized very differently. If you look at the egg for example, the egg is meiotically arrested at metaphase-2 resulting in a diploid genome that is packaged with histones, so that the egg does not undergo complete meiosis, the meiosis arrested the metaphase-2 stage.

The mature sperm, on the other hand, completes the meiosis, that their haploid genomes are packaged with protamines instead of histones. So, by, during spermatogenesis, you can clearly see most of the histones are replaced by another (( )) DNA binding proteins, called as a protamines. So, histones are no longer present in the sperms. When a sperm penetrates the zona pellucida to fertilize the egg, there is the, in time of fertilization both the gametes undergo very rapid changes.

So, you had the oocyte in which the meiosis was not completed, it was stage you had a sperm in which the DNA is highly condensed, but most of the histones are replaced by what are called as protamines. Also, they, what are called as some testis specific histone variants, replace the normal somatic histones in the present in the other somatic cells. But the moment sperm penetrates the egg rapid changes take place, not only in the sperm nucleus, but also in the oocyte nucleus.

What happens to oocyte? The egg completes its 2nd meiosis, resulting in the extrusion of 1 copy of the genome as polar body; the sperm reorganizes its genomic DNA by replacing proteins with histones. So, you can see, during spermatogenesis the histones are replaced by protamines and this cyclic condensed, highly condensed chromatin in the sperm DNA. Now, after fertilization, the protamines are again replaced by histones in the fertilized egg in the case of the sperm.

(Refer Slide Time: 09:02)



Now, around 4 to 8 hours after fertilization, especially after the protamine-histone exchange, that is, when the proteins in the sperm nucleus are replaced by the histones, the sperm-derived sperm paternal pronucleus undergoes, what is called as genome-wide DNA demethylation and this is completed before the 1st cell division. So, the methylated cytosines in the sperm DNA are removed very actively and a sperm DNA undergoes very extensive demethylation within about 4 to 5 hours after fertilization. So, 1st thing that happens after the entry of the sperm into the egg nucleus, into the egg, is the protamines are replaced by histones and then the sperm DNA undergoes extensive demethylation. The paternal DNA, this demethylated DNA remains demethylated even after multiple rounds of cell division, after fertilizer divides several times.

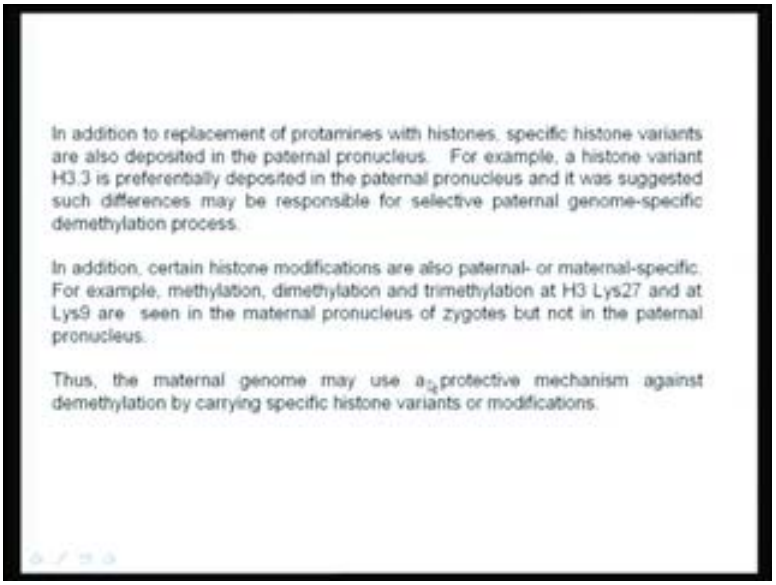
So, specific cytoplasm factors, which are present in the fertilized egg, are responsible for selective demethylation of the paternal DNA. Remember, you have the oocyte DNA and you have the sperm DNA. Only the sperm DNA undergoes selectively demethylation within hours after fertilization.

Thus, although both the maternal, paternal DNA is exposed to the same cytoplasm factors, only the paternal, but not the maternal DNA is demethylated. So, you can see, already the maternal nucleus in the paternal nucleus are, already divide the DNA inherited from the father. The DNA inherited from the mother, they are already very different; one of them undergoes selectively demethylation, other does not. One of them,

the protamines is replaced by the histones, but others remain the same. So, the paternal, maternal nuclei are not the same. So, this may be due to a mechanism that protects the maternal genome from this wave of demethylation or to a putative DNA demethylase that specifically recruited to the paternal genome.

So, what is the mechanism? You have 2 nuclei, you have DNA in both of them, but only one of them, the DNA is extensively demethylated, but other is not. So, there is some mechanism by which there are certain factors present in the cytoplasm, they recognize only the paternal DNA and demethylate only the paternal DNA and not the maternal DNA.

(Refer Slide Time: 11:21)



So, in addition to replacement of protamines with histones in case of sperm DNA, specific histone variants are also deposited in the paternal pronucleus. So, not only the pronucleus is replaced by histones immediately after fertilization, certain specific variants of histones are also deposited in the sperm DNA. One of the examples is histone H3.3. This particular variant of the histone is preferentially deposited in the paternal pronucleus and it is suggested, its differences, such as this, that is, replacement of protamine by histones and selective addition of during the histone variants like H3.3 only to the paternal DNA, but not in the maternal. These may serve as actually signals for this demethylases present in oocyte cytoplasm to selectively demethylate only the paternal genome, but not the maternal genome.

In addition to certain histone modifications, paternal or maternal specific, in addition certain histone DNA modifications are also paternal-maternal specific. For example, methylation, dimethylation and trimethylation of the histone H3 at lysine 27, as well as at lysine 9 are seen only in the maternal pronucleus of zygotes, but not in the paternal nucleus. The point I am trying to make is that although you have the DNA inherited from sperm and the DNA of the oocyte, although the genetic components in terms of DNA sequence is same, more or less same, in both, but you can see, some of this epigenetic changes are quite different in the case of the sperm DNA. The DNA undergoes extensive demethylation, whereas the octet nucleus does not undergo that the same way.

Some specific variants of histones are added only to the paternal nucleus, but not the maternal genome. In addition, specific histone modification takes place only in one of them, but not the other. So, already you can see, the epigenetic signatures of the maternal-paternal genomes are pretty altered very early during embryonic development, in fact, immediately after the fertilization.

So, the maternal genome may use a very protective mechanism against demethylation by carrying certain specific histone variants or histone modifications. So, because of these differences in epigenetic nature of these genomes, one of them undergoes since demethylation and another do not, so people started looking and see, what kind of factors actually regulate this differential demethylation of these 2 genomes.

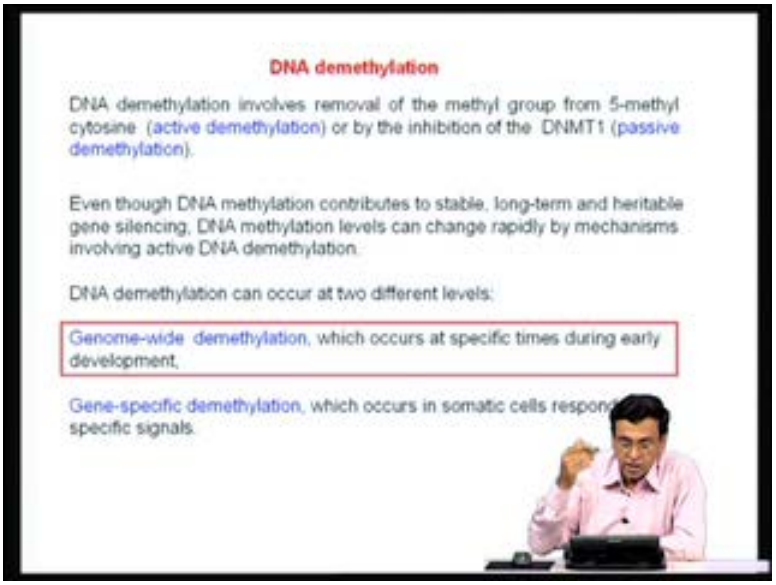
(Refer Slide Time: 13:52)

And extensive research is going on and one of the interesting components has come out recently is, that it has been reported, that certain non-histone factors present in the oocyte might protect the maternal genome from demethylation. For example, zygotes lacking a maternal effect gene known as Stella DPPA3 or PGC7 exhibited demethylation of both pronuclei, although how Stella protects the maternal genome from demethylation remains to be determined.

Some of the studies, in fact, this, this is a very nice paper, which has published in Nature Cell Biology in 2006 and titled PGC7 or Stella protects against DNA demethylation in early embryogenesis. That means, this is a maternal effect gene, that means, this gene is already transcribed and RNA is already stored in the oocyte (( )) fertilization, but the product, the protein arising out of this maternal RNA is known as Stella and if you have this Stella, the maternal genome does not undergo demethylation. But if we now delete or if you make a mutation or if you have a mouse in which this particular gene is not (( )), then both maternal as well as paternal genomes undergo demethylation indicating, that one of the cytoplasm factor, that may be protect, that may protect the maternal genome from demethylation is this protein, called as Stella.

And the exact mechanism, how this protein is acting is still not very well understood. So, let us now try to understand what exactly, how, what is exactly significant of this DNA demethylation and DNA methylation during embryonic development?

(Refer Slide Time: 15:19)



DNA demethylation involves removal of the methyl group from the 5-methyl cytosine. As I said, immediately after the fertilization, the particular genome undergoes extensive demethylation. This demethylation can take place with 2 mechanisms: active demethylation or passive demethylation.

In the case of active demethylation, it actually involves removal of the methyl group from the 5-methyl cytosine. That means, demethylase has to be expressed in the oocyte cytoplasm fertilization and somehow, this demethylase will remove the methyl group from the cytosine of only the paternal genome, but not the maternal genome. The other alternative is that there can be actually inhibition of the DNA methyl transferase-1; therefore this methylation cannot take place. So, this is called as passive demethylation.

So, when a DNA methyl transfer is not expressed, so methylation is not possible and therefore, there is no methyl, methyl (( )) is present. Now, even though DNA methylation contributes to stable long term and heritable genes silencing, but already discussed in the previous classes, how once the DNA is methylated, the methyl cytosine attract proteins of NVCP2 and this NVCP2, in turn recognize specific histone deacetylases and then histone deacetylases come and deacetylase histones and this is how demethylation leading to histone deacetylation, ultimately results in the repression of transcription, that results in activation of gene expression. So, the DNA methyl, DNA methylation actually can lead to gene silencing, usually involving histone deacetylation.

So, the DNA demethylation can occur at 2 different levels, one is genome wide demethylation, which occurs at specific times during early embryonic development. As I had just now mentioned, the entire paternal genome, most of the cytosine gets demethylated. On the other hand, it can also be gene-specific demethylation, which usually occurs in somatic cells, responded to specific signals.

So, many of the epigenetic changes of regulation of gene expression, we have studied in earlier classes, belong to the 2nd category, where we have discussed how specific transcription factors, when they go and bind to promoter region, by either recruiting DNA methyl transferase can methylate DNA and this, in turn, can attract histone deacetylases and therefore, those genes can get inactivated and this is called as genes specific methylation or demethylation. But what we are discussing in earlier embryonic development is that entire genome most of the cytosine is getting demethylated.

(Refer Slide Time: 17:48)

So, during embryonic development some genomic regions are resistant to demethylation. So, when I said the entire paternal genome undergoes demethylation, it is not, that every methyl cytosine in the methyl group is removed, there are certain regions in these genomes, either paternal or maternal, they are protected from undergoing this kind of demethylation. So, there is a global wave of demethylation going on. You can see, how interesting the whole scenario is, immediately after fertilization, selectively, the parental genome gets demethylated. But now, I am telling, not all regions of the paternal genome get demethylated; there are certain regions within the paternal genome, which are protected from undergoing this demethylation.

These regions include, what are called as, imprinting control regions, intracisternal A particle, these are nothing but some transposons, and centric and pericentric hetero chromatin, that is, the chromatin or the DNA in and around the center nucleic region are also protected from undergoing this kind of demethylation. So, there are certain regions, which are protected from this global wave of demethylation that happens immediately after fertilization.

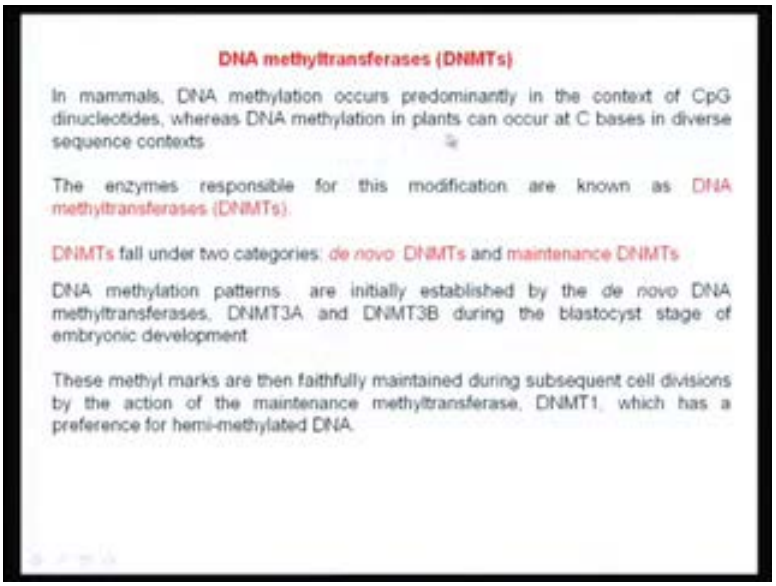
(Refer Slide Time: 18:49)

Now, let us, so now paternal genome gets demethylated, how does it get undergo demethylation? Now, so egg must start dividing and as they did what kind of enzymes are involved in the demethylation, in the methylation of this demethylated DNA? Specific patterns of DNA methylation are reestablished during early embryo development by the action of DNA methyltransferase and demethyl enzymes. This is a very, very important aspect; I would like to spend some more time to emphasize this point because it has lot of implications.

What I am trying to tell you now is, that as soon as the fertilization is over, the genomes of paternal undergoes extensive epigenetic changes. It could be histone modification, it could be replacement of protamines with histone, it could be addition of histone variants, in addition demethylation of selected, demethylation of the paternal genome, all these things takes place. Now, once the demethylation has taken place, a new methylation pattern is established during the embryonic development. This has lot of implications, as the goal of the lecture, it will become very clear. Today people are trying under, do what is called as nuclear cloning, wherein you take the nucleus from a skin cell, that is, adult diploid cell and put this nucleus inside an oocyte which has been enucleated and see, whether this adult nucleus can now trigger embryonic development and can you produce an embryo, and can an adult develop from such kind of a…? This is what is called as cloning, like the dolly sheep was produced by these kinds of methods.

But what I am trying to tell you is that the epigenetic program of an adult differentiated nucleus is not exactly the same as that of the, either the sperm nucleus or the oocyte nuclei. So, this is the reason, all in the DNA sequence information may be same because of this difference in epigenetic signatures of this. That is why, the failure, that of putting this kind of offspring for this kind of somatic nuclear transplantation are very, very low because the epigenetic reprogramming, that takes place in early embryonic development, may not be the same when we take an adult differentiated nucleus and put inside oocyte. That is very different from that of a sperm versus egg fusing together and forming a diploid nucleus.

(Refer Slide Time: 21:11)



Now, the enzymes, that do this DNA methylation is called DNA methyltransferases and in mammals, DNA methylation occurs predominantly in the CpG residues. We have this, discussed these things quite a bit in our earlier classes, but interestingly, in the case of plants DNA methylation can occur at cytosine bases in many other sequence contexts. Whereas, in animals predominantly is a dilute receptor CpG, which is the substrate for DNA methylation. The cytosine in the CpG gets undermethylated and the enzymes responsible for this DNA methylation are known as DNA methyltransferases.

There are 2 types of DNA methyltransferases, de nevo, de novo DNA methyltransferases and maintenance DNA methyltransferases. Now, the DNA methylation patterns are

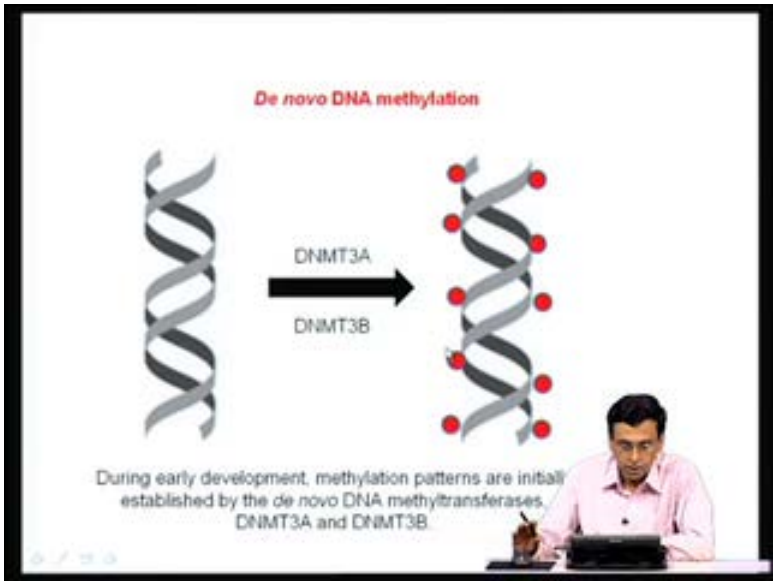
initially established by the de novo DNA methyltransferases. There are 2 such enzymes, called DNMT3A and DNMT3B during the blastocyst stage of embryonic development.

As I said, the genome undergoes extensively demethylation, immediately after fertilization and it is this DNMT3A and DNMT3B, which can again reestablish the methyl pattern of the zygotic nucleus. These methyl markers are then, faithfully maintained during subsequent cell divisions by the action of maintenance methyltransferases called d n, DNMT1, which has a preference for hemi-methylated DNA. Let me try to explain these 2 aspects, choosing specific cartoons, what exactly you mean by DNA demethyltrnsferases and maintenance method transferases.

(Refer Slide Time: 22:39)

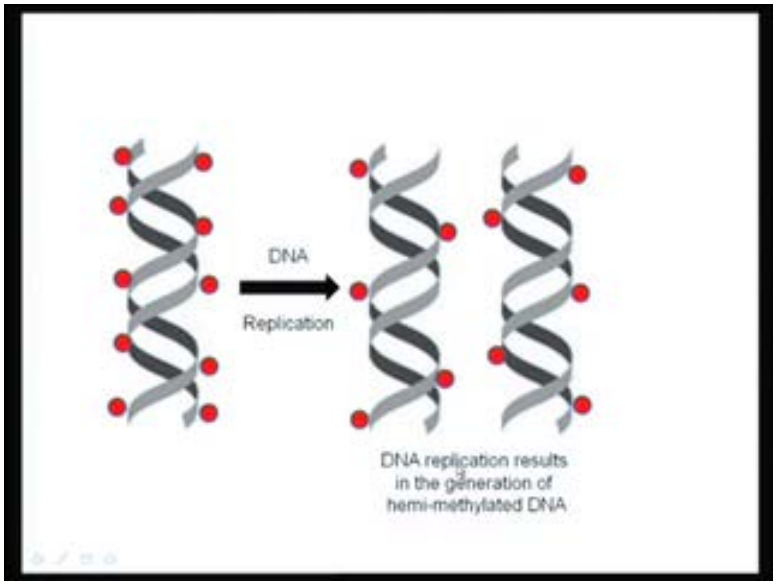
So, there are 3 DNA methyl transferases, one is called DNMT1, which is a maintenance methyltransferase and DNMT3A, DNMT3B, which are called as de novo methyltransferases. And 3A and 3B are highly expressed at embryo implantation, then waves of de novo methylation occur in the genome. Let us try to understand, what exactly happens using the cartoon.

(Refer Slide Time: 22:51)



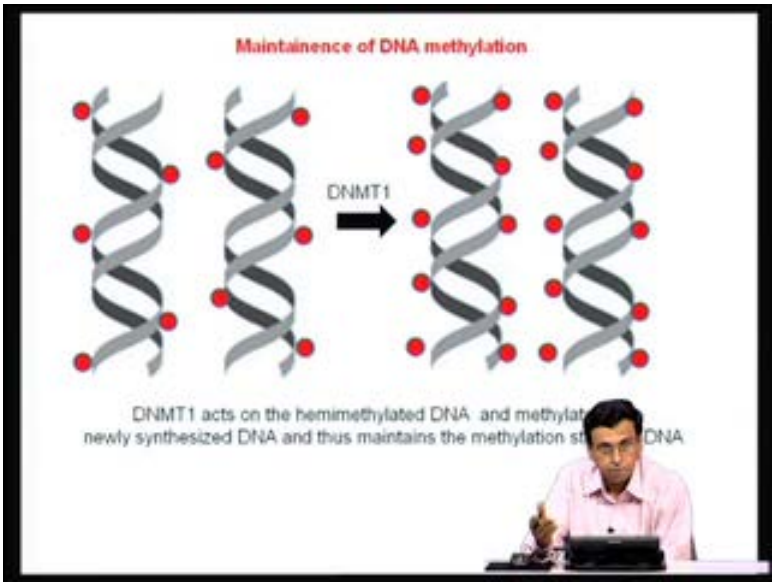
This is what happens, as I said, let us assume, that this is the DNA from a sperm and you can see, the DNA has now been completed demethylated by the action of the demethylases immediately after fertilization. And this demethylated DNA now acts a substrate for 2 enzymes, DNMT3A and DNMT3B, these are called as de novo, demethylases and they now add methyl group cytokine residues of this, let us say for example, is sperm DNA. So, during early development, methylation patterns are initially reestablished by the de nova, de novo DNA, methyl DNA methyltransferases DNMT3A and DNMT3B.

(Refer Slide Time: 23:27)



Now, when as a zygotic nucleus divides, the DNA replicates and DNA also divides, so 1, the dotted stands are formed and therefore, what do you get is a hemi-methylated DNA. So, the methylated DNA, the 2 strands, which are in the row, the one goes to each one of the dotted strand, strand and therefore, the newly formed dotted stands are remain unmetyhylated. So what you get is a hemi-metheylated DNA after 1st round of replication. This hemi-methylated DNA now becomes a substrate for the next enzyme DNMT1.

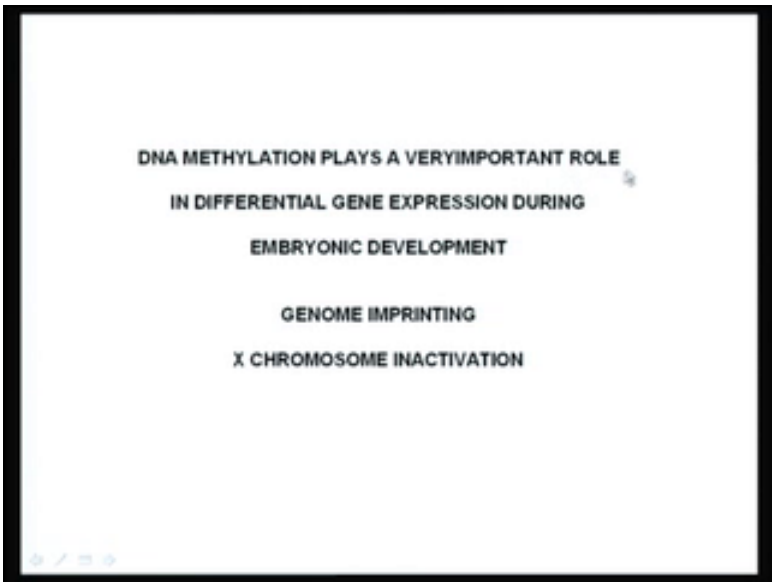
(Refer Slide Time: 23:57)



So, the methylation is first reestablished by the de novo methyltransferases 3A and 3B. Once the specific methylation pattern is established by them, this pattern is maintained during successive cell division by this maintenance methyltransferases. For example, in a promoter region for a gene, for example, let us, there are 10 cytosine residues, but only 5 are methylated by the DNMT3A and DNMT3B. The same 5 residues continue to be methylated by DNMT1 through the subsequent cell division. So, the methylation signature is maintained by the maintenance methyl, DNA methyltransferases. So, the hemi-methylate DNA is a better substrate for this enzyme than the completely methylated DNA. And therefore, the other dotted stands get methylated and therefore, the hemi-zygotic, the hemi-methylated DNA now becomes fully methylated. So, DNMT1 acts on the hemi-methylated DNA and methylate the newly synthesized DNA, and thus maintains the methylation signature of DNA during embryonic development.

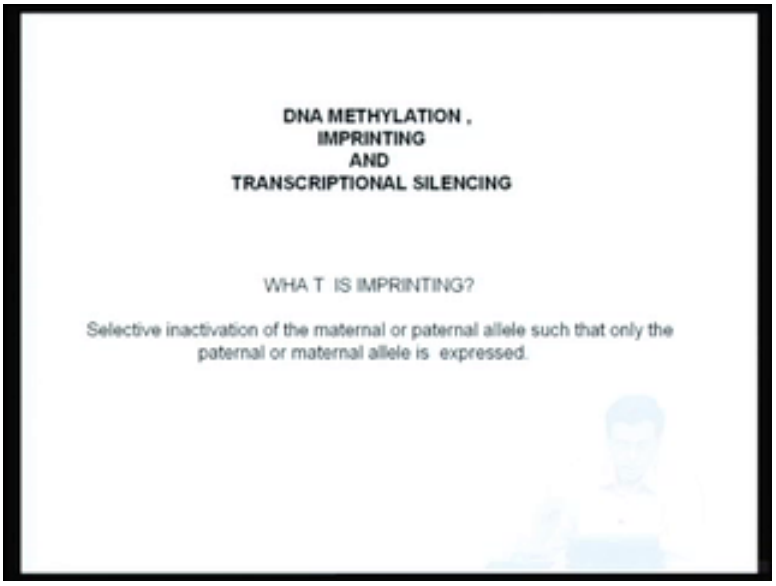
Now, DNA methylation, so what I told you so far is that first there is a demethylation and then there is a remethylation by both, de novo methyltransferases, as well as maintenance methyltransferases. So, once this methylation signature is established after fertilization in the zygotic nucleus, let us now try to understand what is the significance of the DNA methylation, what kind of the functions that it regulates?

(Refer Slide Time: 25:22)



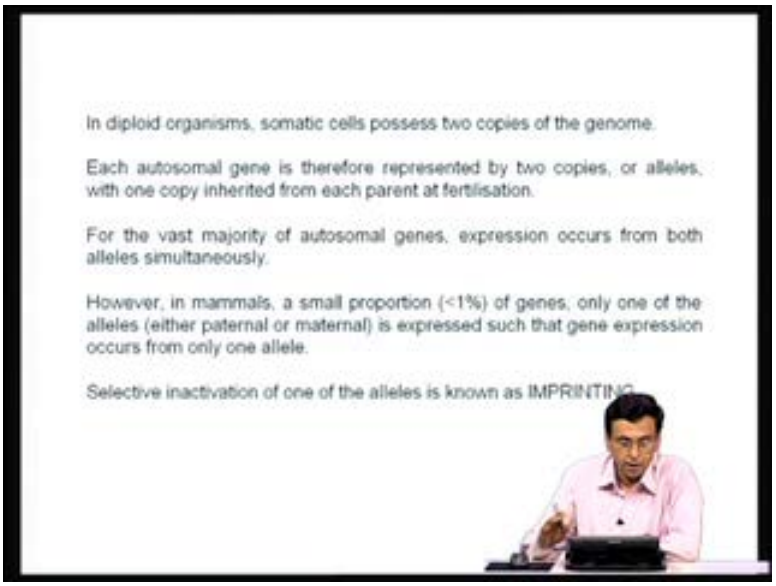
DNA methylation plays a very important role in differential gene expression during embryonic development. I am going discuss today only very 2 examples, one is called genome imprinting and another is called X-chromosome inactivation. Let us take these 2 processes, see how DNA methylation plays a very important role in these 2 processes.

(Refer Slide Time: 25:38)



Now, what is imprinting? Imprinting is nothing but selective inactivation of maternal or paternal allele, such that only the paternal or maternal element allele is expressed. Let me, let me repeat, imprinting is nothing but selective inactivation of the maternal or paternal allele, such that only one of this allele is expressed during the development.

(Refer Slide Time: 26:03)



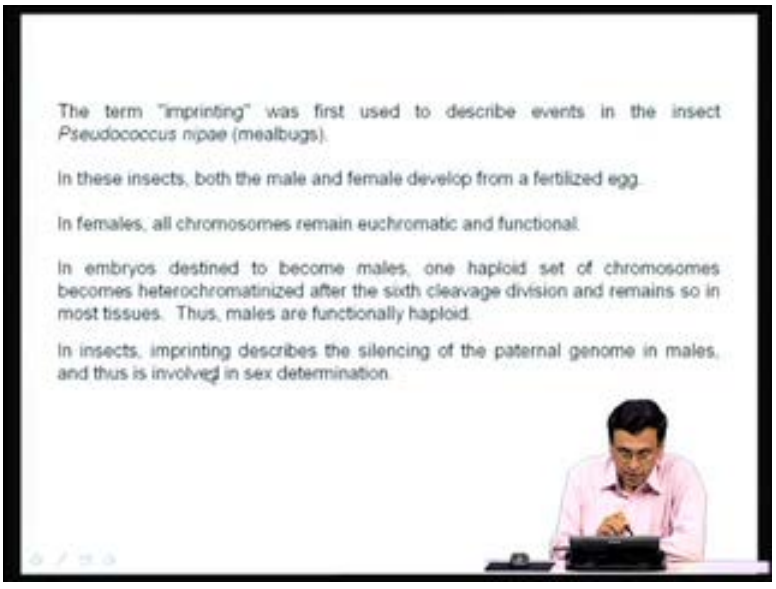
Let us elaborate little bit what exactly is this. Now, in diploid organisms, the somatic cells possess two copies of the genome because one comes from the father, one comes from the mother and each autosomal gene, therefore is represented by two copies, or

alleles, with 1 copy inherited from each parent at fertilization. So, every gene has 2 alleles, 1 comes from the father and 1 comes from mother.

Now, in majority of the genes, let us say for example, in humans we have some anywhere, 30 to 50000 genes, 99 percent of them or even more than that, both the alleles are expressed, that is, the allele inherit from the sperm, as allele inherit from the oocyte, both are expressed.

But, in some cases, a very small proportion, about 1 percent of the total genes, only one of the allele, either paternal or maternal is expressed such that gene expression occurs from only 1 allele. So, you can see, although we have 2 alleles for every gene, one coming from the mother, one coming from the father, all the 99 percent of this genes, both the chromosomes, both the alleles are expressed, about 1 percent of the genes, either a paternal allele is expressed or maternal allele is expressed and this is what is known as imprinting. So, selective inactivation of one of the alleles is known as imprinting.

(Refer Slide Time: 28:14)

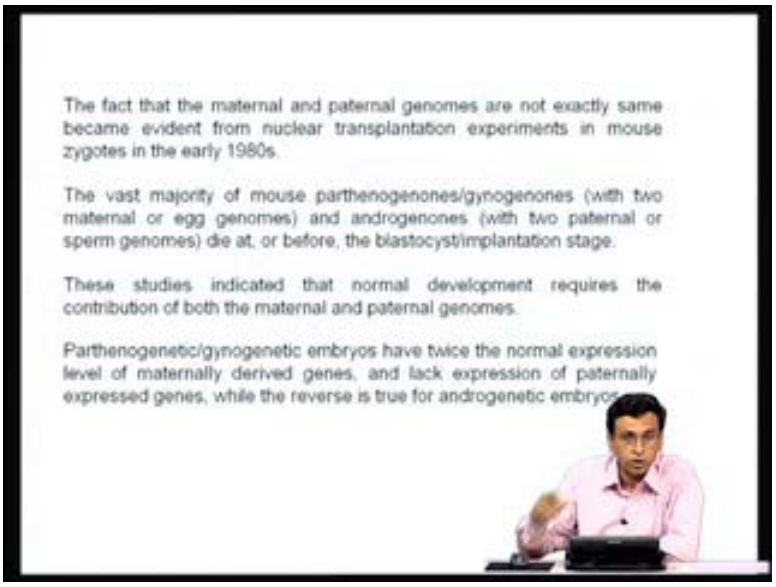


The term imprinting was first used to describe events in the insect Pseudococcus nipae, was popularly known as mealbugs. When these insects, both male and female develop from a fertilized egg; once fertilization is over, both male and female develop from a fertilized egg. However, in females, all chromosomes remain euchromatic and are functional, there is no inactivation. Whereas, in embryos, which are destined to become males, 1 haploid set of chromosomes becomes heterochromatinized after the 6th

cleavage division and remains even in most of the tissues. So, the males' 1 set of chromosomes get hetrochromatinized or gets inactivated, so that there is no expression of the gene from this particular set of chromosomes.

Although, therefore, although the process' two sets of chromosomes, one from the paternal, one from the maternal side, functionally they are haploid because only one set of genes are getting, only one set of alleles are going to be expressed. So, in insects, this kind of imprinting leading to expression of only one set of alleles, describes a silencing of paternal genome in males and thus, therefore, involved in sex determination. So, if that fertilized egg has to be different to males, the one set of alleles get inactivated. Therefore, functional becomes the haploid and such eggs developed into male, whereas those eggs with both alleles are active, they develop into females. So, this kind of an imprinting that takes place in males in the insects, actually determines the sex of the fertilized egg, whether it should develop in male or it develops in female.

(Refer Slide Time: 28:44)



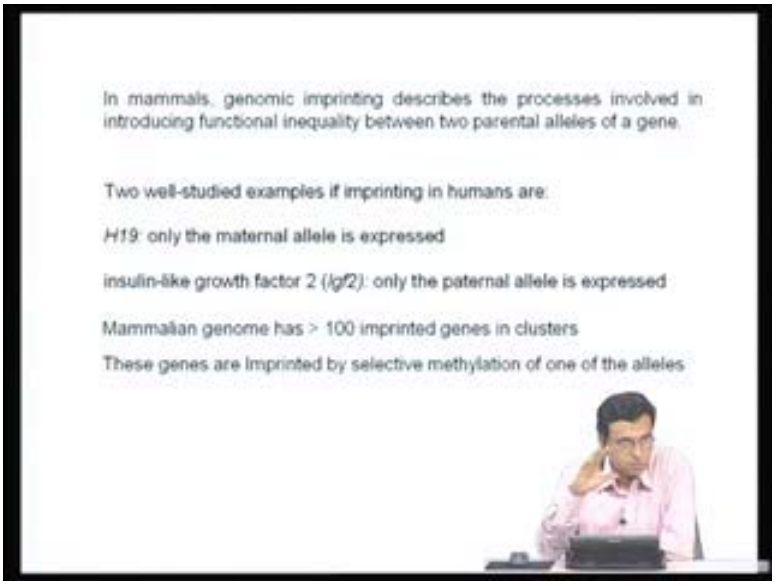
The fact, that both maternal paternal genomes are not exactly same became evident from nuclear transplantation experiments in mouse zygotes in the early 1980s. So, all that I have told you so far clearly indicates, that the, although a number of chromosomes and genes sequence, everything is more or less same in the sperm as well as the oocyte, all the information I have told you very clearly tell, that as far as epigenetic signatures are concerned, it is not the same, they are different. How did they prove that?

In vast majority of the mouse parthenogenones or gynogenones with, that is, these are the ones (( )), which contain with the two maternal or the egg genomes, no sperm genome or those, which contains two sperm genomes, but more oocyte genome, they all die before the blastocyst or implantation stage. So, if the genome from the oocyte genome, from the sperm, is exactly the same, if you replace the sperm nucleus and instead put 1 more oocyte nucleus, which will still develop the same, so these are called as the gynogenones. That is, we have genome from two eggs or you replace, you take out the oocyte nucleus and introduce two sperm genomes, so these are called as the androgenones. But when we do these things in such kind of, fertilized eggs were such kinds of eggs, this is called as parthenogenesis, where to develop at all, so you need to have a maternally inhibited genome and a paternally inhibited genome for proper development, clearly indicating, that they are not exactly the same.

These studies indicated that normal development requires the contribution of both, the maternal and paternal genomes. If you have either of them, development is abnormal, so parthenogentic or gynogenitic embryos have twice the normal expression level of maternally derived genes and lack the expression of paternally expressed genes, while the reverse is true for the androgenetic embryos. So, if you take some of these embryos, which although they die at early stage, but those early stages, if you look for the expression of genes, there are lot of abnormality in the gene expression. Some of the genes are expressed twice and some of the genes, which are normally expressed from the parental allele, they are not expressed at all.

So, in mammals genomic, this is all in the insects. In the insects, there is a complete (( )) of alleles and the early studies have clearly said, that if you now put either 2 oocyte genomes or 2 sperm genomes, it raises abnormal development and abnormal gene expression.

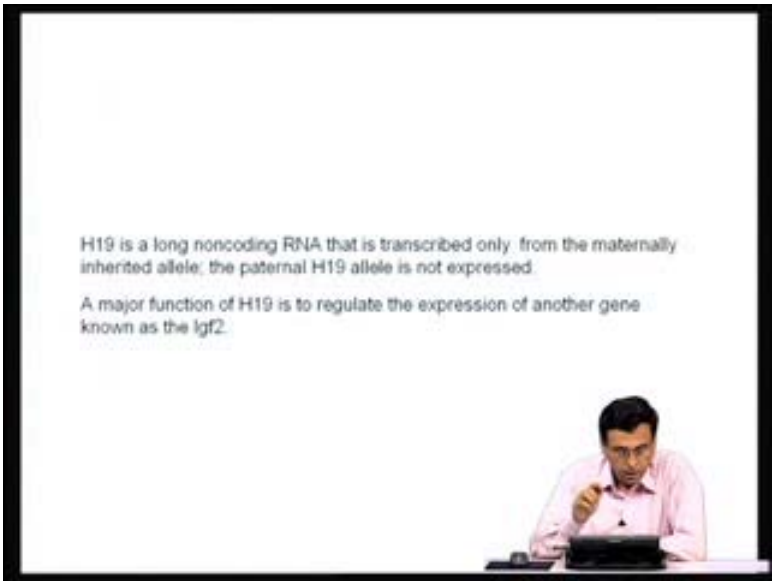
(Refer Slide Time: 30:58)



So, in mammals, genomic imprinting describes the processes involved in introducing functional inequality between two parental alleles of genes. There are some genes, which are not expressed in double dose; they are expressed only in the single dose. We will understand better if we give one or two example of this imprinting and tell you, how exactly is imprinting takes place.

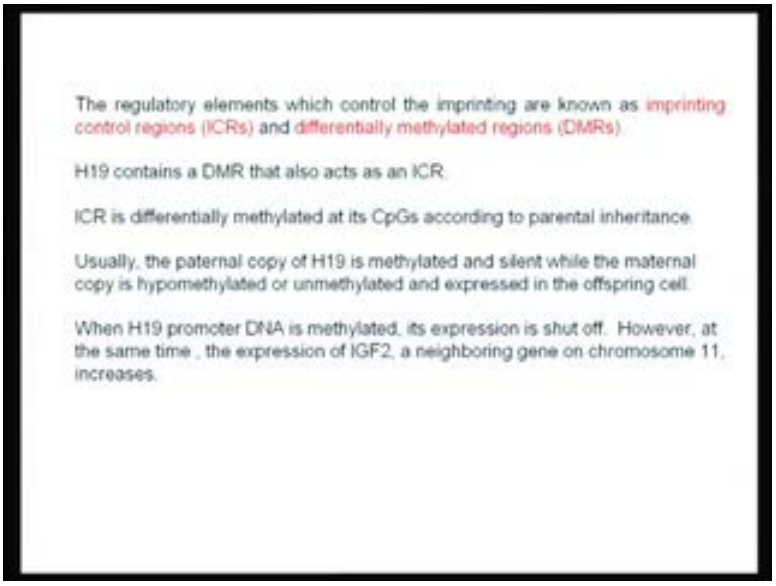
Two of the very well imprinted genomes or imprinted genes in humans are, one is called as the H19, another one is called as insulin-like growth factor two or Igf2. Now, interestingly, although the H19 gene is present in both, paternal chromosomes as well as maternal chromosomes, the H19 is expressed only from the maternal chromosomes, that is, only the maternal allele of the H19 is expressed during embryonic development in the adult. Similarly, the Igf2, only the paternal allele is expressed, but not maternal allele. So, you can see, these are reciprocal. The H19, always the maternal allele is expressed, paternal allele is not expressed; whereas, in the case of Igf2, only the paternal allele, which gets expressed, not the maternal allele. Like this, in the mammalian gene at least hundreds such imprinted genes have been identified so far. In these cases, although the, both the alleles are present, only one of the allele should be expressed at the given time, so turns out these genes are imprinted by selective methylation of one of the alleles. So, this kind of imprinting, the basic mechanism that is responsible for the selective inactivation of one of this allele is nothing but DNA methylation. Let us now try to understand how exactly this takes place.

(Refer Slide Time: 32:43)



Now, H19 goes for a long noncoding RNA that is transcribed only from the maternally inherited allele, in the paternal H19 allele is not expressed. Now, I will not complicate at this time what is called as noncoding RNA, we will discuss this in later lecture. A noncoding RNA is nothing but it is RNA, it is not transferred into protein, that is, the functional molecule is RNA and not the protein. In fact, there are now several noncoding RNAs in addition to ribosomal RNA and transfer RNA, which do not, which are also non-coding RNAs, these are well known noncoding RNAs. Studies in the last 2 decades have led to identification of several such noncoding RNAs; we will discuss some of these noncoding RNAs in the later stages of a lecture series. Now, a major function of this H19 is to regulate the expression of another gene, known as Igf2, let us see how exactly this takes place.

(Refer Slide Time: 33:39)



Now, the regulatory elements, which control imprinting, that is, which are responsible for selective expression of either the maternal and paternal allele are known as imprinting control regions and differentially methylated regions. The cis-acting regulatory elements, which are responsible for this imprinting or expression of one allele or other, are known as imprinting control regions or ICRs or DMRs are differentially methylated regions.

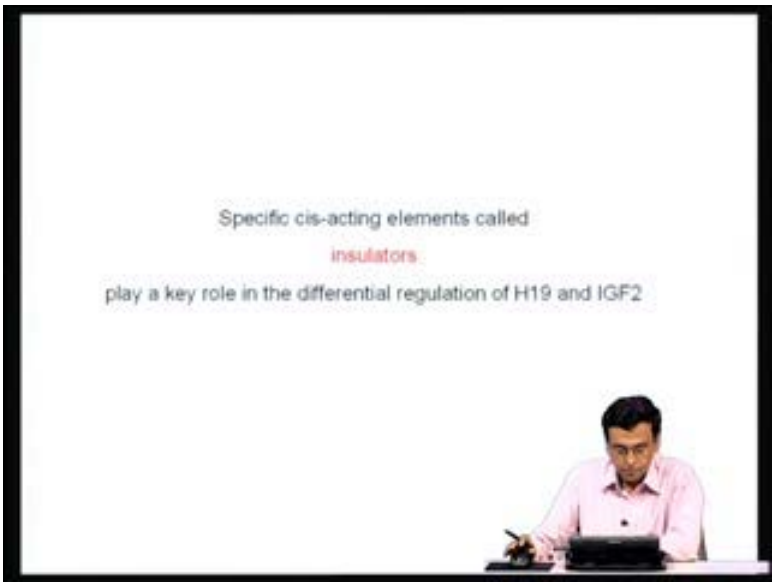
Now, let us take H19 for example, the H19 contains a differential methylated region that also acts as an imprinting control region or ICR. The ICR is differentially methylated at the CpG residues according to parental inheritance. So, there are CpG sequences within this imprinting control region, are the imprinting control regions of the H19 gene and this region of the ICR, the cytosine region undergo differential methylation depending upon where it is a paternal allele or the maternal allele. Usually, the paternal copy of H19 is methylated and silent while as the maternal copy is hypomethylated or unmethylated and therefore, expressed in the offspring. So, we can see, we have two alleles, one contains from the mother, one contains from the father. In the case of H19, the maternal allele is hypomethylated and therefore, it is expressed, where in the paternal allele is methylated and it cannot get expressed.

So, when H19 promoter DNA is methylated, its expression is shutoff as we, as we studied earlier. DNA methylation results in the attraction of the system deacetylasers,

which are deactivated histones and therefore, it can result in condensation of chromatin and therefore, inactivation of gene expression. So, an H19 promoter DNA is methylated, its expression is shutoff. But very interestingly, that H19 promoter at expression is shutoff. The same time, the expression of Igf2, which is a neighboring gene H19, Igf2 are present side by side in the same chromosome.

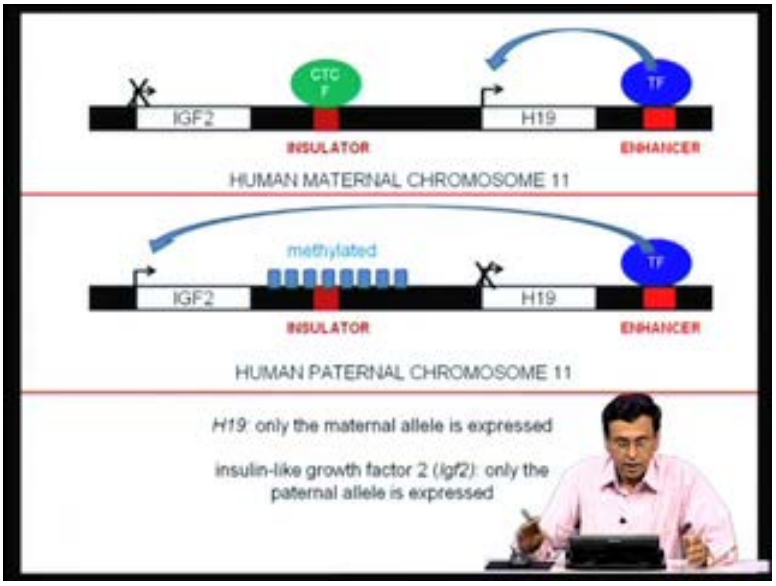
When H19 expression is shutoff, the expression of Igf2 is stimulated or is turned on. Now, let us try to understand how repressing the transcription of one gene results in the activation of expression of another gene; how does it take place?

(Refer Slide Time: 35:48)



Very simple, there are specific cist-acting elements called as insulators, which play a very important role in this differential expression of gene expression. That means, when the expression of H19 is turned off, the expression of the neighboring gene Igf2 is turned on and a certain specific cis-acting elements or promoter elements, called insulators, play a very, very important role in this differential expression of these 2 neighboring genes.

(Refer Slide Time: 36:11)



I have now put a cartoon here to explain, although it is a very complicated phenomenon, I try to explain in a very simple manner. Let us now understand a mechanism by which this actually takes place. Now, this is the human chromosome 11; in the human chromosome 11, both the H19 gene and Igf2 gene are present side by side, they are neighboring genes. Now, the H19 has an enhancer sequence at the 3 prime end of H19 gene, usually (( )) 5 prime end, but in this case enhancer is present in the 3 prime end and enhancers can act in an orientation dependant manner, independent manner.

Now, in the maternal chromosome, where the H19 gene is normally expressed, transcription factors recognizes this enhancers sequence and they bind to it and therefore, this transcription factor now activates transcription from the transcription (( )) and therefore, H19 RNA is made.

But, the, this transcription factor cannot activate the expression of Igf2 because there is a sequence called insulator sequence and to this insulator sequence, a protein called CTCF binds. So, when CTCF binds to this insulator motive, this transcription factor can only activate the expression of H19 and it cannot activate the expression of the Igf2 promoter.

Therefore, in the maternal chromosome only the H19 gene is expressed, but the Igf2 gene is not expressed. Now, let us see what happens, that paternal chromosome? In the paternal chromosome the insulator region in the, in the H19 region, the insulator sequence near the H19 region is methylated and when this gets methylated, now this

CTCF protein can no longer bind to the insulator sequence. And therefore, this insulator can no longer prevent the activation by this transcription factor of the Igf2 promoter.

Now, since this H19 promoter is methylated, H19 gene cannot be transcribed because now, the histones are deacetylated here and therefore, the chromatin is condensed. But however, because the insulator sequence is not bound the CTCF, now the transcription factor, which is binding to the 3 prime end is not H19 enhancer. The activity, its, it can now, there is no barrier here, this insulator barrier is now removed. Therefore, this now goes and activates the expression of Igf2 and that is how, in Igf2 now gets expressed.

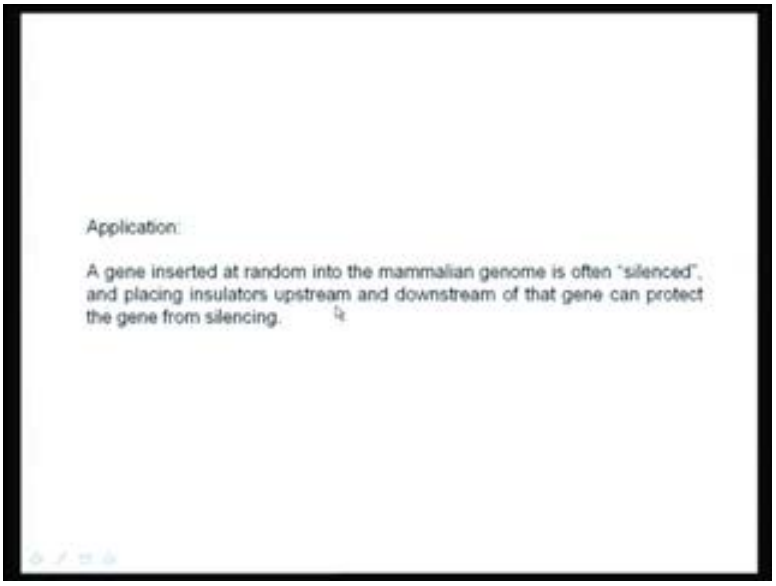
So, you can see, methylation of this insulator sequence removes the barrier and therefore, a transcription factor bound to the H19 enhancer can now activate the transcription of Igf2 gene, whereas when the H19 promoter and insulator sequence is not methylated, a protein called CTCF binds to insulator and therefore, it acts as a barrier and prevents the transcription fact from activating the transcription of the Igf2 sequence and you can see, how nicely these two now reciprocally regulate the expression of each other. That is why, whenever H19 is expressed, Igf2 is not expressed and whenever Igf2 is expressed, H19 is not expressed. So, it is a wonderful example, how epigenetic regulation takes place, regulates the expression of two different genes.

What is the application of this? As I said, this course, whenever possible I always try to give applications to see, how understanding some of this basic mechanism have led to the benefit of mankind.

Now, one of the major problems in the areas, like gene therapy, which again we are going to discuss a few classes later is, that when I want to treat a genetic disorder, when I want to express my gene of interest, when I insert the gene along with the promoter into the cells, the DNA goes inside and integrates into the genome, but expresses only for a short time and then it gets inactivated.

It could be either because the DNA goes in heterochromatin regions or it could be undergo hailstone or histone deacetylation may take place in the promoter region, whatever the reason, the foreign DNA does not get expressed very highly. This is one of the major problems in a number of applications, like gene therapy and so on and so forth.

(Refer Slide Time: 40:12)

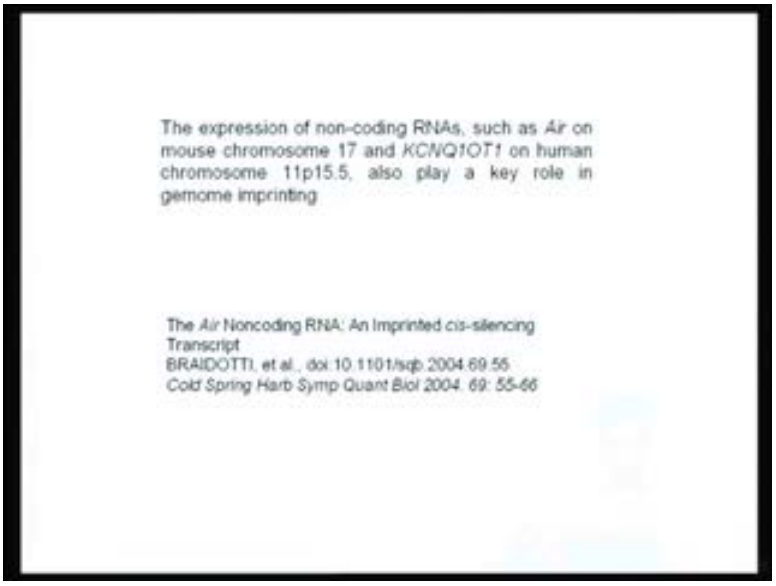


But now, what people have now found out is that if you now take these gene and you, if you now construct the vector in such a way, that you now put insulators, both upstream and downstream of these gene sequence, if you now put that, then the chances of this kind of a silence effect acting on this, your trans gene of interest is very, very low. So, a gene inserted at random into mammalian genome is often silenced. This is one of the major problems of gene therapy; we will discuss this again when we come to gene therapy.

But if you now place this kind of insulator sequence on either side of the gene and then, introduce this gene into the nucleus, this insulator sequence will prevent the silencing effect on to the trans-gene and therefore, your trans-gene can be protected from gene silencing and therefore, your gene can be expressed for a very long time.

This has tremendous implication in the area of gene therapy. We will discuss this in detail when we talk about gene therapy in the later classes.

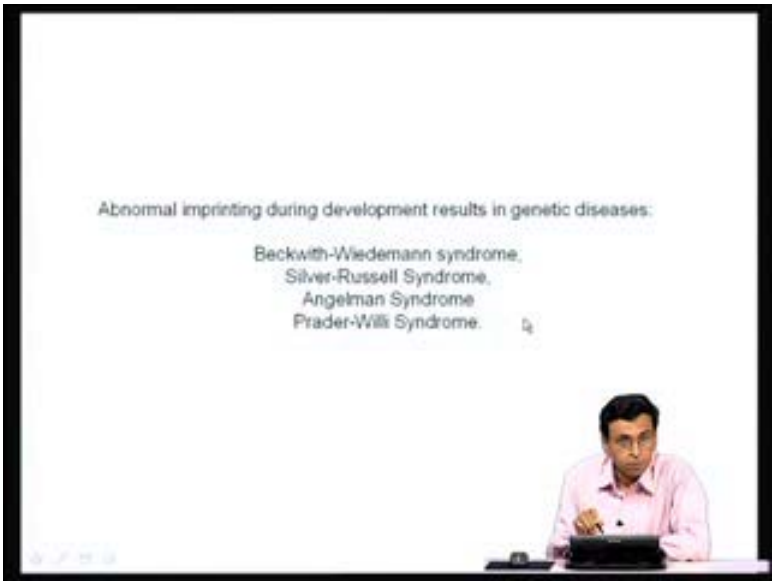
(Refer Slide Time: 40:58)



In addition to this mechanism, what I told you, in the case of H19 and Igf2, there is number of mechanisms that are involved in imprinting. For example, expression of certain non-coding RNAs such as Air on mouse chromosome 17 and KCNQ1OT1 on human chromosome 11p, they all play a very important role in genome imprinting.

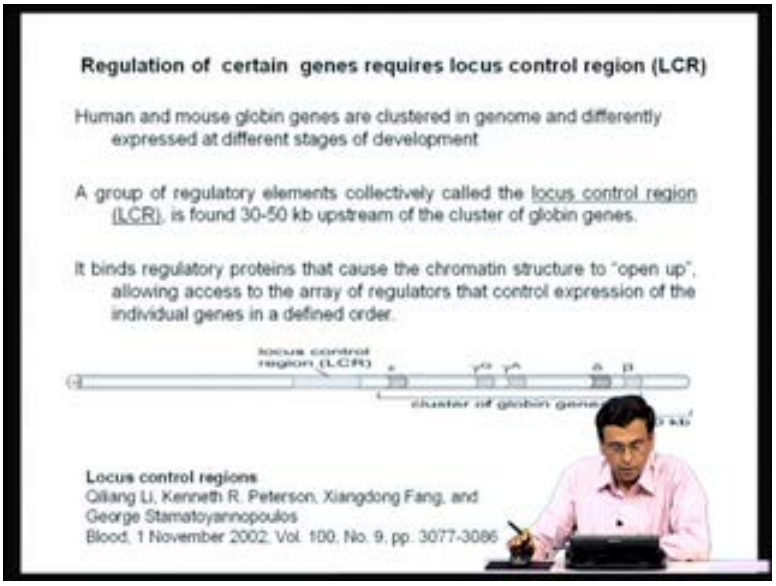
What I am trying to say is, that recent studies indicating, that many non-coding RNAs, that is, these are RNAs, which look like messenger RNAs, but they do not code for any proteins. But they seem to have very important regulatory functions including imprinting and such RNAs into play very, very important role in the chromosome genome imprinting. I will not go in the details, if you are interested, you can go through this article the Air noncoding RNA imprinted cis-acting silencing transcript and then, study how the exactly this noncoding RNAs regulate gene expression.

(Refer Slide Time: 41:50)



Now, this imprinting is very, very important and as I can see here, abnormal imprinting during development results in number of genetic disorders, like Beckwith-Wiedermann syndrome, Silver-Russell syndrome, Angelman syndrome, Prader-Willi syndrome. So, if this imprinting does not takes place properly and this maternal versus paternal gene activation is not done properly, it can result in number of genetical disorders.

(Refer Slide Time: 42:11)



There are many such interesting examples of differential gene regulation involving imprinting and so on and so forth. During embryonic development, another interesting

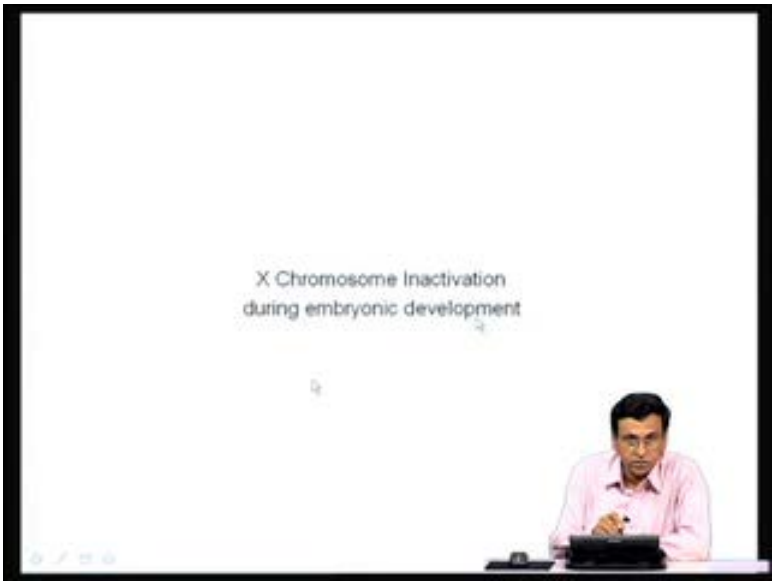
example is, what are called as, there are certain cis-acting elements known as locus control regions. So, regulation of certain genes during development requires, what are called as a locus control region or LCR. Now, the best example to understand this locus control region is the expression of the globin genes. During development, human and mouse globin genes are clustered in the genome and are differentially expressed at different stages of development.

Now, there are, this is the globin gene cluster. What I have shown is clear, that is, what is called epsilon gene, gamma, delta and beta. The beta is finally, expressed in the adult, whereas in the embryonic development, this other variants of this globin gene is expressed only in the adult. The beta globin is expressed, but they all come from the same locus. So, there is a common cis-acting regulatory element called as locus control region, which differentially activates these genes depending upon the various stages of development. So, a group of regulatory elements collectively known as locus control region present 30 to 50 kb upstream of the cluster of genes, plays a very, very important role in this differential activation of these various genes. In the beta globin locus, this locus control region binds to regulatory proteins and causes chromatin structure to open up allowing access to array of regulators, that control the expression of individual genes in defined order.

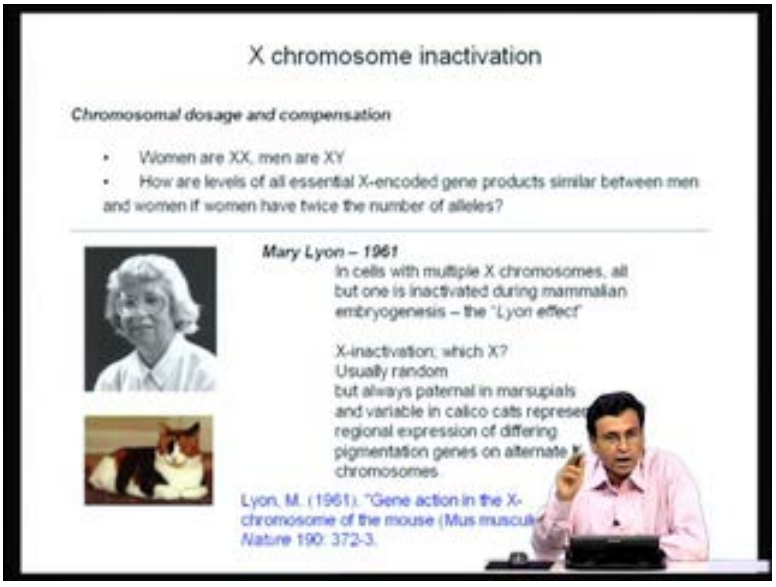
I will not again go into the details; it is a very complicated subject. Again, there is a very interesting article in journal Blood, called, the title Locus control regions and one can always go through this and try to understand, how these locus control regions play very, very important role in differential regulation of gene expression during embryonic development.

I am going to spend may be a few minutes to another important mechanism of epigenetic regulation of gene expression, but again, has a very lot of implications, that is, inactivation of the x chromosome during embryonic development.

(Refer Slide Time: 44:05)



(Refer Slide Time: 44:08)



X chromosome inactivation during embryonic development – now, as you know, the females usually have 2 X chromosomes, whereas the males usually have only 1 X and 1 Y chromosome. That means, genes, which are present on the X chromosome, they are represent at twice in females, but they are represented only once in case of males.

So, if you have X linked genes in the females, the level of expression be 2 times than the same level of gene expression in the males, which is not correct. That means the genes present in the X chromosome, whether it is male or female, should be expressed only at

the same level. To attain this, what happens, that is, what is known as the X-chromosome inactivation. So, although females have 2 X chromosomes, one of the X-chromosomes is completely inactivated, so that the genes, which are expressed in the males and females in the X-chromosome are same.

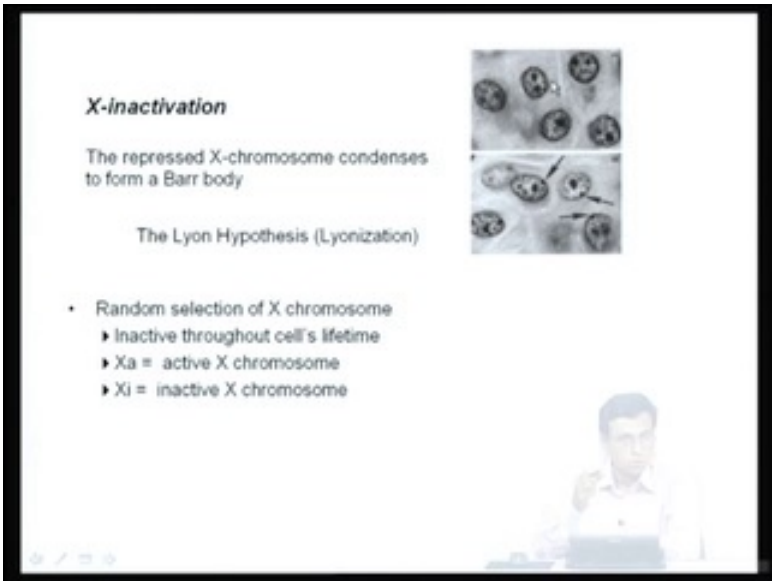
The expression of the X linked genes in the male and females are usually same because one of the X-chromosome females is inactivated, so women are XX and men are XY in their chromosome computation. All levels, all levels, how are the levels of all essential X-encoded gene products are similar between men and women, if women have twice the number of alleles? This was the question Mary Lyon asked this way back in 1971 and what she set to discover is that in cells, which contain multiple X chromosome, all but 1 is inactivated during mammalian embryogenesis and this what she called as a Lyon effect.

If you have, for example 3 X chromosomes, only 1 X chromosome remains inactive, all others get inactivated. In the females we have 2 X chromosomes, only 1 remains active, another remains inactive. Now, question is, which one gets inactivated is the paternal X or maternal X? Again, there is lot of interesting differences, usually with, at the paternal X chromosome should be inactivated or maternal x chromosome inactivated. It is a random phenomenon, in some cells paternal X chromosome gets inactivated, in some other cells maternal X chromosome gets inactivated during embryonic development.

But in certain mammalian species like marsupials, that is, kangaroos for example, it is always the paternal X chromosome, which gets inactivated. Whereas, in (( )) mammals like humans, in some cells paternal X chromosome gets inactivated, in some other cells maternal X chromosome gets inactivated. In cats for example, again it is random and that is why, you have, this variegations in cats are, you always have this pigmentation genes because it, depending upon what kind of, in some cells paternal X or maternal X is going to get inactivated. That is why, have this varying code colors.

This is the first classic paper, which was published by Marry Francis Lyon in 1961, Gene action in the X-chromosome of the mouse in nature, which actually threw light upon this differential X-chromosome inactivation and inactive X-chromosome can actually be stained as a very dark body and is often referred to as the Barr body.

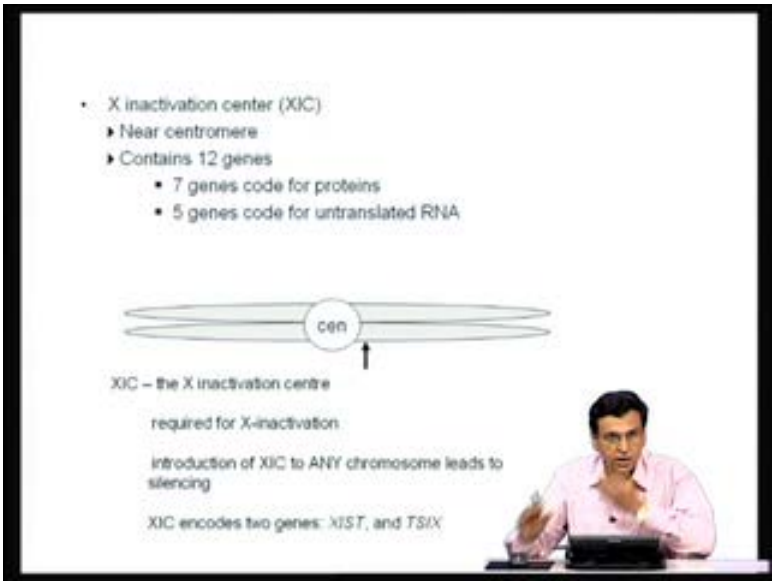
(Refer Slide Time: 47:12)



So, if you (( )) staining using typical DNA stains and in the female cells, if you simply take the buckle smear from a female and buckle smear from a male, in the female buckle smear, you can stain for DNA, you will find this Barr body. Because one of the X-chromosome is, because it is highly heterochromatinized, it can take a (( )) of the stain and you can see it as heterochromatinized as a Barr body.

Now, what is the mechanism by which this X-chromosome inactivation takes place? It turns out, once one of this X-chromosome is inactivated, it remains inactivated for the rest of the life. So, this is a, you know, this is different from the transient activation/inactivation, that we have normally seen for other genes. Many of these inactivations are for the rest of the life; it just does not change.

(Refer Slide Time: 47:43)

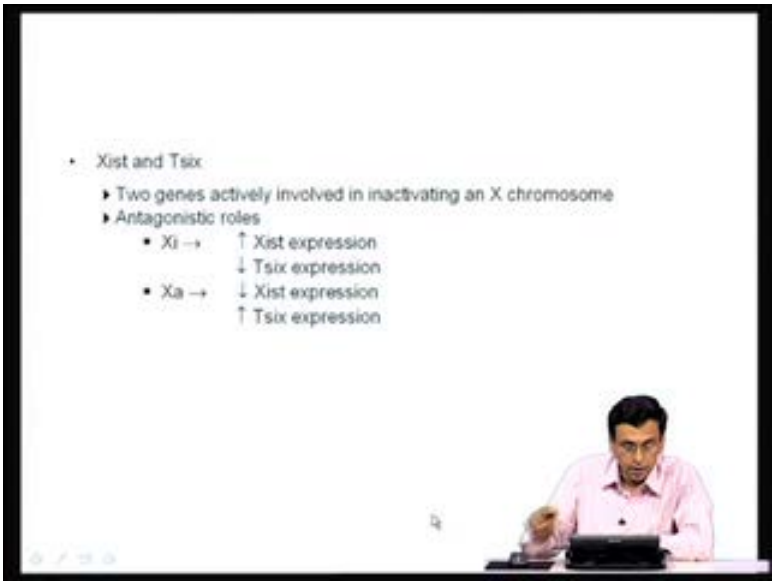


Now, in turns out in the last 2 decades a lot of information has come out on the mechanism by which this X-chromosome inactivation takes place. It turns out, in the X-chromosome, near the centromeric region; there is what is called as X inactivation center or XIC.

Now, this region, which is pointed by arrow here, contains about 7 code genes, which codes for proteins and 5 genes, which codes for RNA, but this RNA does not get translated into proteins. That means, 5 noncoding RNAs and 7 protein coding RNAs arise from this X inactivation center of the X-chromosome.

This X inactivation center is very, very crucial for X-chromosome inactivation. If you remove this X inactivation center, then X-chromosome inactivation does not take place. And if you take this X-inactivation center and put it on another chromosome, that entire chromosome can get inactivated clearly indicating, that the mechanism by which X-chromosome inactivation takes place is this region, is responsible for the X-chromosome inactivation, turns out the 2 important genes, that play very, very important for the X-chromosome inactivation or what is called as XIST, pronounced as XIST and TSIX or Tsix.

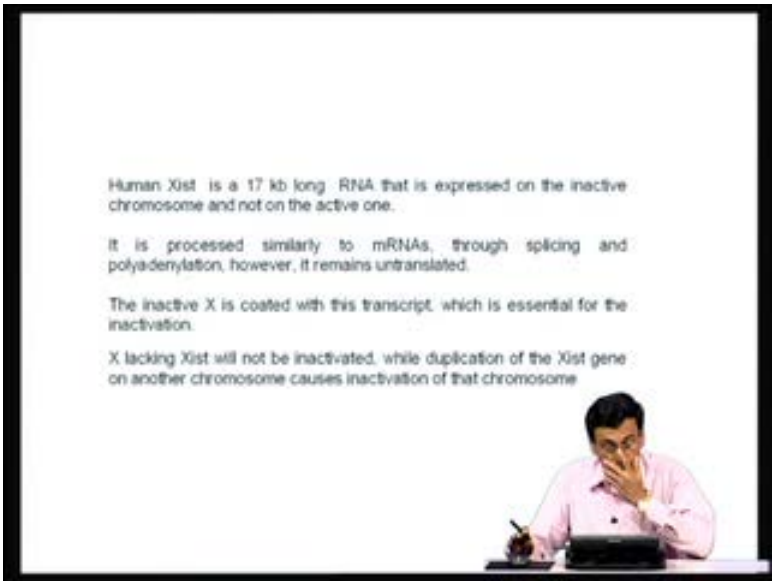
(Refer Slide Time: 49:01)



Now, Xist and Tisx, these are 2 genes, which are actively involved in inactivating an X chromosome. Both of them are noncoding RNAs, they do not code from the proteins, turns out, the X-chromosome, which remains active, the expression of Xist is very high. The X-chromosome, which remains inactive, so I called as X I, the Xist expression is very high, whereas Tsix expression is very, very low.

On the other hand, the X-chromosome, which remains active, Xist expression is very low, Tsix expression is very high. So, the, these two transcripts are, have a reciprocal relationship.

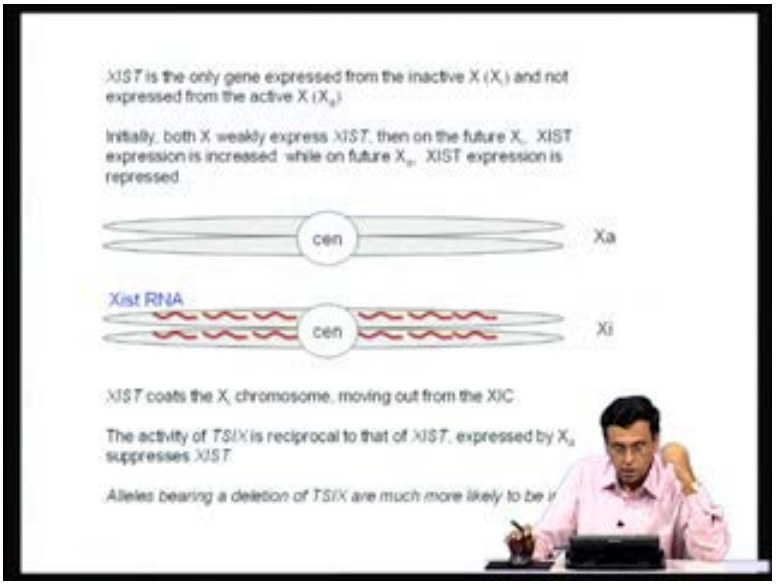
(Refer Slide Time: 49:37)



And what is this Xist? Now, the Xist is nothing but a 17 kilobit long RNA that is expressed on inactive X-chromosome, but not on the active X-chromosome and is processed similar to the messenger-RNAs, but it remains untranslated and this inactive X is coded with transcript, which is essential for the RNA.

So, once this transcript is made from this X-chromosome, this RNA codes the entire chromo, X-chromosome.

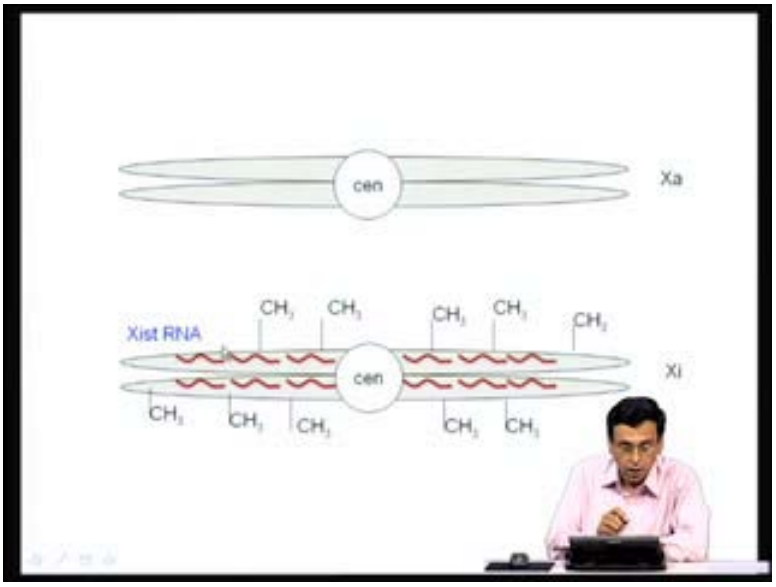
(Refer Slide Time: 50:01)



And therefore, as I show in this cartoon here, once this Xist RNA is transcribed from the inactive X, this Xist RNA codes the entire X-chromosome and once this X-chromosome remains coated in this Xist RNA, it remains inactive, but as the other LL, in which this RNA does not code remains active. So, a noncoding RNA called Xist, which is translated from one of the X-chromosome, when it coats, when it coats the entire X-chromosome that leads to inactivation of the X-chromosome.

Whereas, the activity of the Tsix is reciprocal to that X and is expressed the Tsix, that means, the chromosome in which the Tsix is expressed, Xist is not expressed and therefore, it remains active. Whereas, the chromosome where Xist is expressed, Tsix is not expressed, therefore it remains inactive. So, depending upon which transcript is expressed, the chromosome either will remain active or remain inactive.

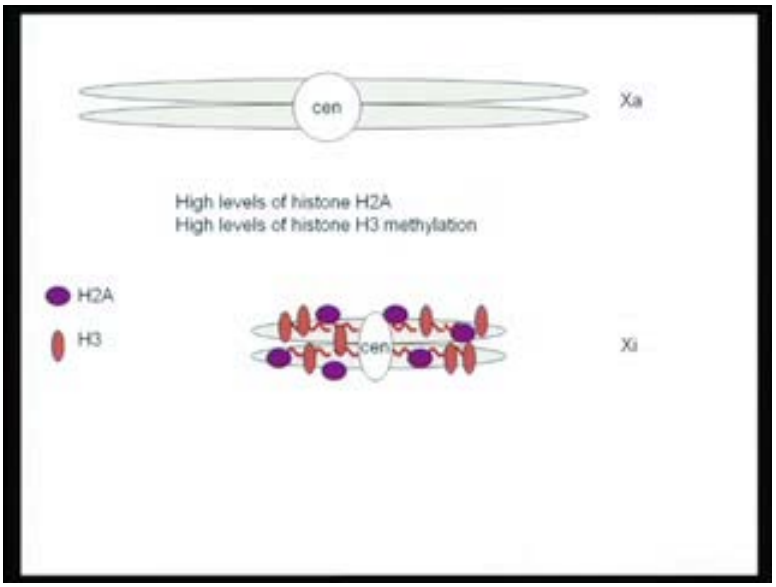
(Refer Slide Time: 50:53)



Now what is the mechanism? The mechanism is again simple. Once the Xist RNA coates this entire X-chromosome, it now attracts DNA methyltransferases and this DNA methyltransferase methylate the cytosine residues all along the chromosome, X-chromosome. And this DNA methylation, now in turn attracts certain chromatin modifying proteins including histone deacetylases and this results in complete heterochromatin (( )) X-chromosome and that is how X-chromosome inactivation takes place.

So, you can see, again DNA methylation plays a very important role in the X-chromosome inactivation.

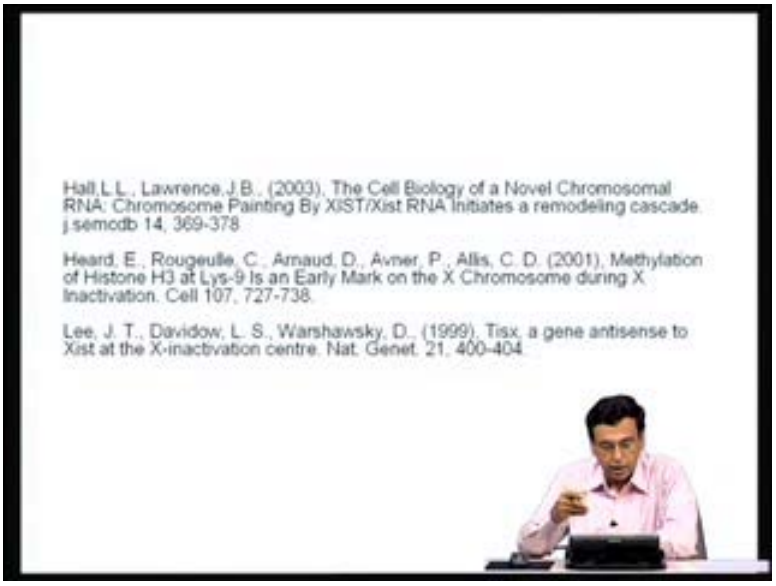
(Refer Slide Time: 51:23)



As I have shown here, once the Xist RNA coats the entire X-chromosome, which is (( )), which has to be inactivated, it attracts number of histones as well as histochromatin modifying proteins and as a result, it controls high condensation. Therefore, becomes heterochromatinized and therefore, becomes inactive and the genes there are not expressed.

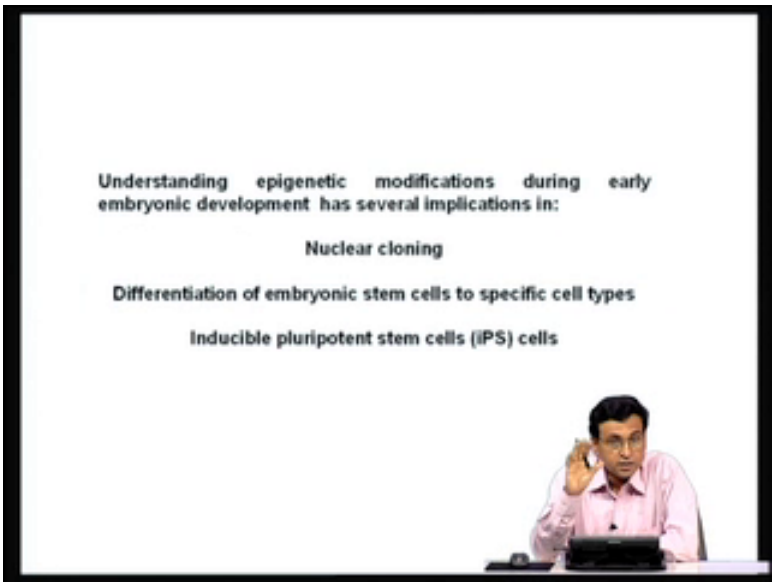
So, DNA methylation plays very, very important role not only in imprinting, but also in the X-chromosome inactivation, as well as gene specific regulation during embryonic development.

(Refer Slide Time: 51:52)



Again, here a number of references, that although I have told you in a very, very simple manner, X-chromosome inactivation is very, very complex subject, that a number of (( )) source, which are very, very important, very interesting and you can read many, some of this papers to understand how this X-chromosome inactivation takes place.

(Refer Slide Time: 52:09)



Now, I will come to the last stage of my lecture here. Understanding is epigenetic modification during early embryonic development has several implications. To understand how this DNA methylation takes place, what, how histone modification takes

place during embryonic development and how differential gene expression and differential chromosome inactivation during embryonic development has number of implications in nuclear cloning, differentiation of embryonic stem cells to specific cell types, as well as inducible pluripotent stem cells.

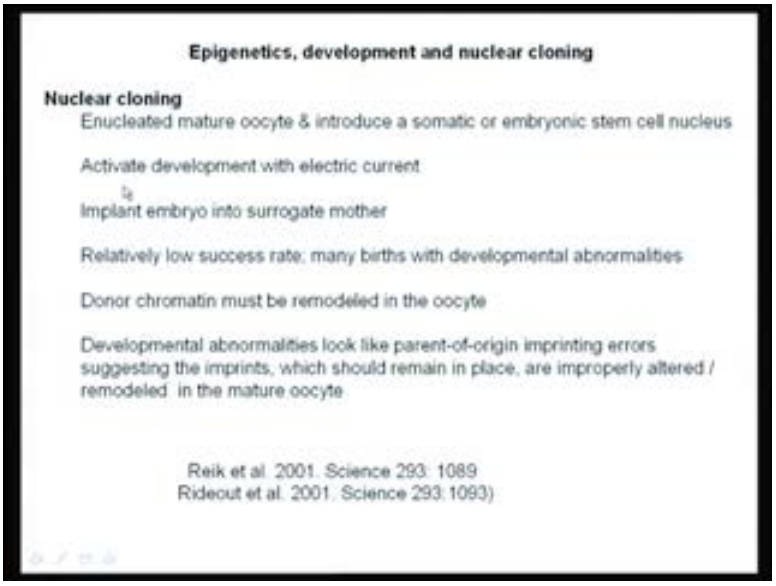
In the next class, we are going to discuss some of these aspects to understand how epigenetic, understanding epigenetic during embryonic development is now paving way for all these studies. Why nuclear cloning has not succeeded to very high levels, why the failure rate is very high in the nuclear cloning?

If you take a nucleus from an adult differentiated cell and put it in an unfertilized egg, enucleated unfertilized egg and then induced to development in embryo, like the famous doll sheep, it, ultimately it is not a very healthy animal, that has problems. That is mainly because the epigenetic program in the adult cell nucleus is not exactly the same as that happens in the fertilized oocyte and sperm nucleus. The diplod nucleus arising out of this sperm and egg nucleus fusion is not exactly the same as the nucleus of the adult differentiated cell and this is the reason, why offspring arising of such kind of a nuclear cloning experiments are neither healthy, nor, nor the success rate is very high.

Similarly, people are now trying to convert stem cells into different cell types. Can I make, can I cure diabetes by developing pancreatic cells from a embryonic stem cell? Can I develop (( )) cells from a embryonic stem cell? Can I develop neuronal cells to treat diseases like Parkinson's or Alzhiemer's from a embryonic stem cells?

All this again, involves differential gene expression. Can you induce expression? Can you reprogram these adult stem cells or the embryonic stem cells to specific cell types, so that it can have number of therapeutic applications? All these, again involve differential gene regulation, epigenetic reprogramming and so on and so forth. So, let us, will now discuss some of these aspects in next class. How understanding the basics science in, are differential gene regulation during embryonic development has now let to better understanding of this kind of a very, very important bio-medical applications.

(Refer Slide Time: 54:38)



Like for example, nuclear cloning, like what I told you just now, enucleated mature oocyte and introduce a somatic or a embryonic stem cell nucleus; activate the development and then implant embryo in a surrogate mother. But the problem in such cases, the birth-rate is very low and many times the offspring, that are born, has several developmental abnormalities. Therefore, the donor chromatin, which is present in the adult nucleus, is not exactly same as that of the octet nucleus, and that is the reason, why it has lot of problems.

(Refer Slide Time: 55:04)

The take home message I want to give is, that there is, what is called as transient or the short term gene silencing, permanent or long term gene silencing. And these are all, again involve lot of epigenetic reprogramming.

Developmental genes that are needed during later stages of development are transiently held in a repressed state during early embryonic development. This is achieved through short-term epigenetic marks such as histone modifications, which can be removed before or within cell division. Like the, in the last class we discussed about homeotic genes now, which have to be expressed at different stages of development. These are all achieved by direct histone modifications; you acetylated or deacetylated histones and depending upon that, you can either activate or repress those genes.

But these are all transient, sometimes you turn on and later, you can shut off and then you can turn on again. So, these kinds of a transient gene expression changes or gene silencing, when you have to temporarily inactivate a gene at a certain stage of development, but can activate later, that usually involves direct histone modifications, but many cases when you want to permanently silence a particular gene, but often involves DNA methylation.

Certain genes of genome, that are marked with epigenetic information, that is stably maintained heritable after many cell divisions, such as imprinted genes, transposons, inactive X chromosome. They require long term silencing and that has to be sustained throughout the development and the life span of an organism. This is generally achieved by DNA methylation, an epigenetic mark that refers to the addition of methyl group to the 5th carbon base of cytosine. Because DNA methylation provides heritable, long-term silencing that is crucial for an organism. So, DNA methylation leaves to long term DNA silencing, whereas direct histone modifications often is involved in transient on short term DNA silencing.

(Refer Slide Time: 56:47)

So, this one article that I recommend you to read, Active DNA demethylation many roads lead to Rome appears in the Nature Reviews Molecular Biology, which very nicely discusses some of this early demethylation, methylation, to changes that takes place during early embryonic development. I will stop here.