Eukaryotic Gene Expression

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

Lecture No. # 3

Gene Regulation in Eukaryotes: Diversity in general Transcription factors

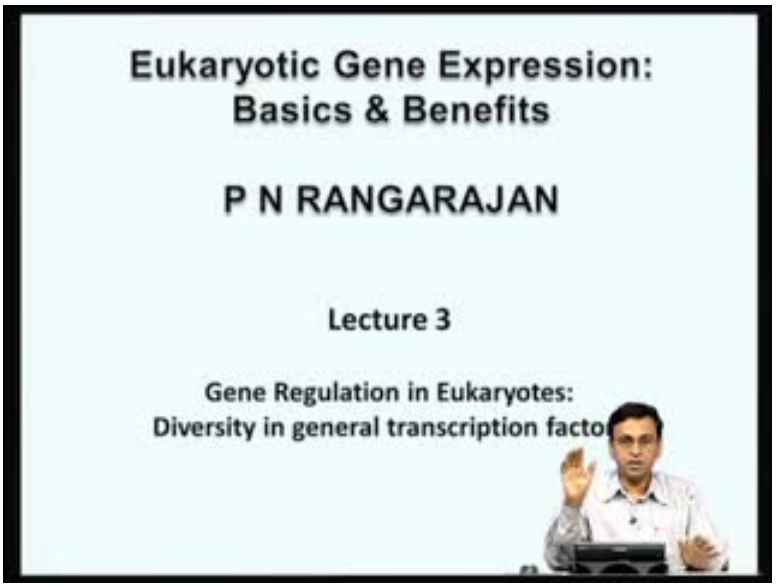
This is the third lecture in this particular course, and in this lecture, we are going to talk about diversity in general transcription factors. Now, so far, we have discussed about the diversity in RNA polymerases in eukaryotes. There are I, II, and III RNA polymerases, and more recently, there is a polymerase IV, which was discovered in plants. Then, we discovered that there are what are called as core promoter elements, to which these general transcription factors and RNA polymerase bind, and that is what contributes to transcription initiation.

And in the last class, I have actually mentioned, the core promoter elements means not just TATA box; there are many variants of these TATA sequences, and these variations in TATA sequences actually contribute to differential gene regulation, and I gave a number of examples of…, like a initiator motif, BRE, DRE, and so on and so forth, how various core promoter sequences are actually contribute to differential gene regulation. And I also mentioned to you that these core promoter sequences are very important not only for transcription initiation, but they also play a very important role in transcription activation by upstream activators.

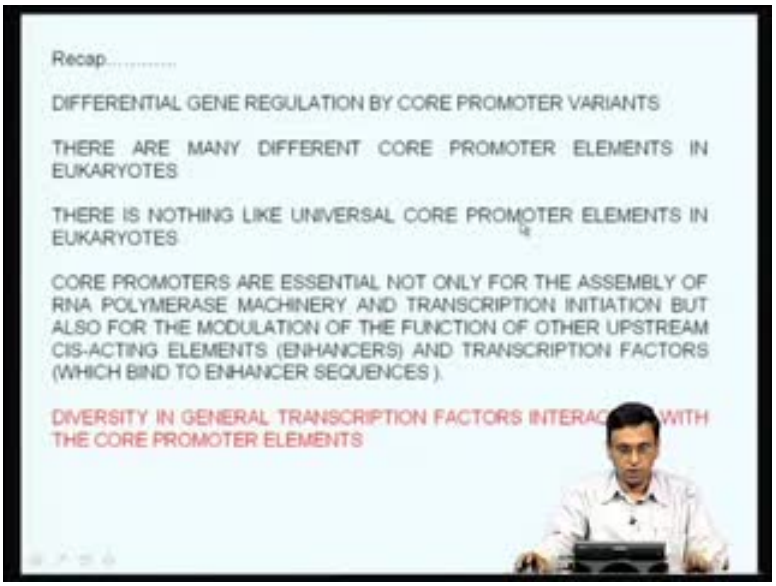
There are many transcription factors which bind to upstream sequences called enhancer sequences, and they function only when a specific core promoter element is present. For example, I gave you examples where, even within the TATA sequence, if you replace the TATA sequence of one promoter with the TATA sequence of another promoter, transcription activation cannot take place by a transcription activator. So, variations within the core promoter elements themselves contribute to diversity in gene regulation. So, the job of core promoter elements is not just to contribute to initiation of transcription, but they also contribute to differential gene regulation. That is what is the

crux of the previous two talks. Today, you are going to focus on the general transcription factors are also diverse.

(Refer Slide Time: 02:00)



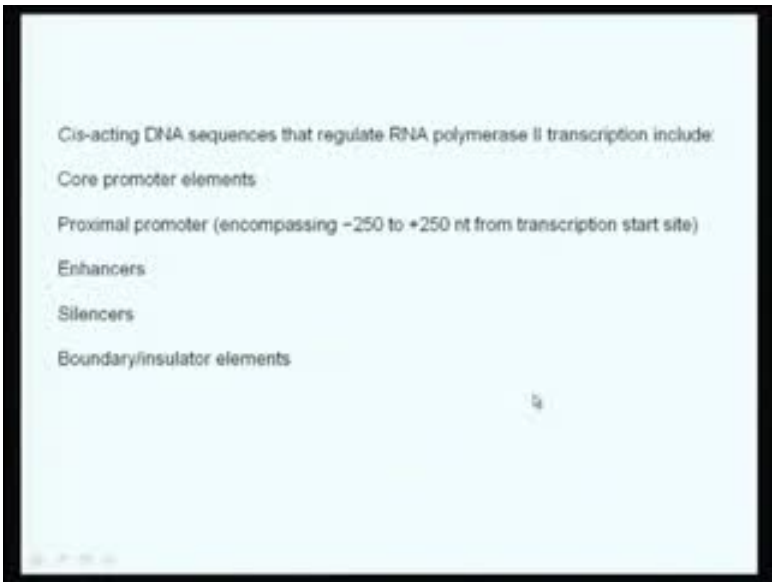
(Refer Slide Time: 02:16)



Now, there are not… there is not just one TFIID, or there is not just one TBP; there are variations of general transcription factors, and this diversity in general transcription factors can also contribute to differential gene regulation in eukaryotes. That is what going to be the crux of the talk. So, just to recapitulate what we discussed, so far, we discussed so far about differential gene regulation by core promoter variants. I discussed

just now, there are many different core promoter elements in eukaryotes; not just a TATA box, there are many variants of these sequences, and there is nothing like a universal core promoter element in eukaryotes, because not a single core promoter element present in all the TATA boxes, and all the eukaryotic promoter, some of them contain initiator sequence, some of them contain TATA motif, some of them contain a BRE, DRE, and so on and so forth. Core promoters are essential not only for the assembly of RNA polymerase machinery and transcription initiation, but they are also essential for the modulation of function of other upstream cis-acting elements known as enhancers, as well as transcription factors which bind to enhancer sequences.

(Refer Slide Time: 03:34)



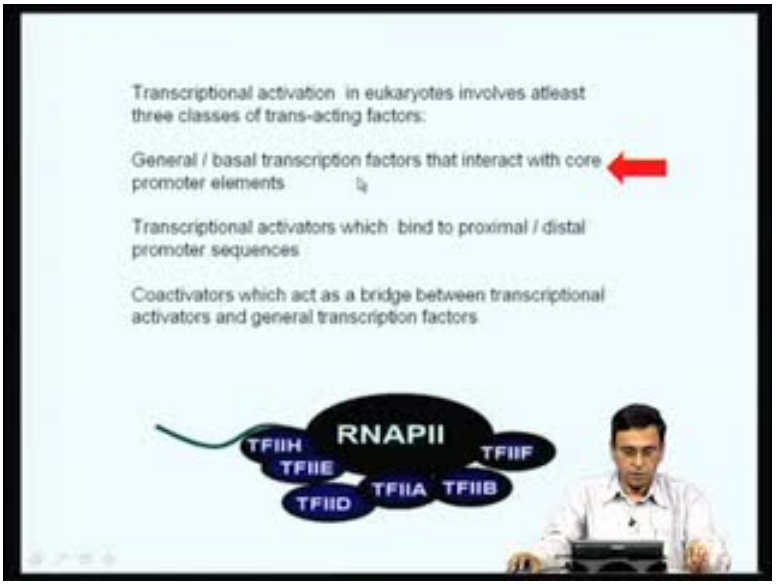
I just discussed with you, there are some enhancer sequences. They function as enhancers only when certain core promoter elements are present downstream. If you replace these core promoter elements with the variants, this upstream enhancer can no longer function. Therefore, there is a link between these different promoter sequences in a eukaryotic promoter, and what now we are going to focus today, is that how diversity in general transcription factors interacting with the core promoter elements can also contribute to differential gene regulation. It is going to the crux of today's talk.

Now, this is again to give you a overall perspective, because I do not want you to be under the impression that we are only talking about a general transcription factors, and there are other regions which are not important. If you now, just like I give an example

of a cricket pitch in a last class, when we talk about cricket match, we have players, we have different players, we have are close in fielders or outdoor fielders, but all the action is focused only on the cricket pitch. So, the promoter is actually analogous to cricket pitch, and you have different batsmen, different bowlers, coming and then doing.

Just in the same way, you have different protein factors, which contribute to differential gene regulation, and we also have what is called the close in field, outfield, and we have a stadium, and so on and so forth. The same way, you have proximal promoter elements, you have distal promoter elements, you have enhancers, silencers, boundary insulator elements; all these ultimately contribute to overall regulation of gene expression, but today, we are only going to focus up on general transcription factors which are binding to the proximal promoter sequences, and how that alone is can bring about several levels of gene regulation. that is going to be

(Refer Slide Time: 04:44)

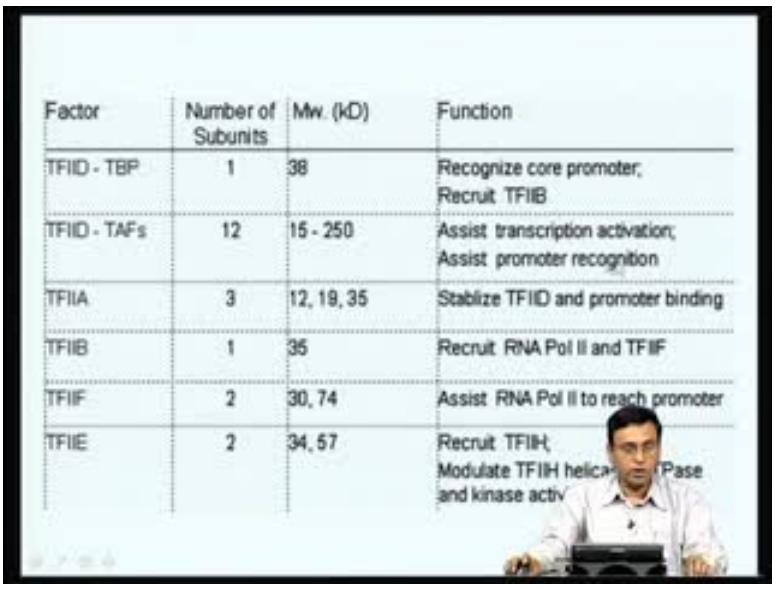


So, in the sub sequent classes, we are going to talk about other promoter elements like enhancers, silencers, boundary elements and so on and so forth. Now, transcription activation when it So far, we discussed only about cis-acting elements. In the last class, we only talked about the cis-acting elements, which are present in the core promoter region. Today, we are going to talk about trans-acting factor– that is, the protein factor which act in trans.

And what kind of transcription factors are actually involved in transcription activation in eukaryotes? Generally, there are about, broadly, three classes of trans-acting factors, which play important role in transcription activation in eukaryotic promoters. The first set is called as the general or basal transcription factors that interact with the core promoter elements, and today's talk is going to focus on this particular set of protein factors.

There also transcription activators which will bind proximal distal promoter sequences; we will come to it at later stage in this course. There are also, very importantly, what are called as coactivators or corepressors, which act as a bridge between transcription activators and general transcription factors. So, just to give a very brief overview, you have a set of sequences called core promoter sequences, you have a set of sequences called proximal and distal promoter sequences, enhancers, silencer, and so on, which are present far upstream, and there are a bunch of protein factors called general transcription factors which bind to the core promoter sequences, and there are a bunch of protein factors called enhancer binding proteins or transcriptional activators, which bind to far upstream sequences.

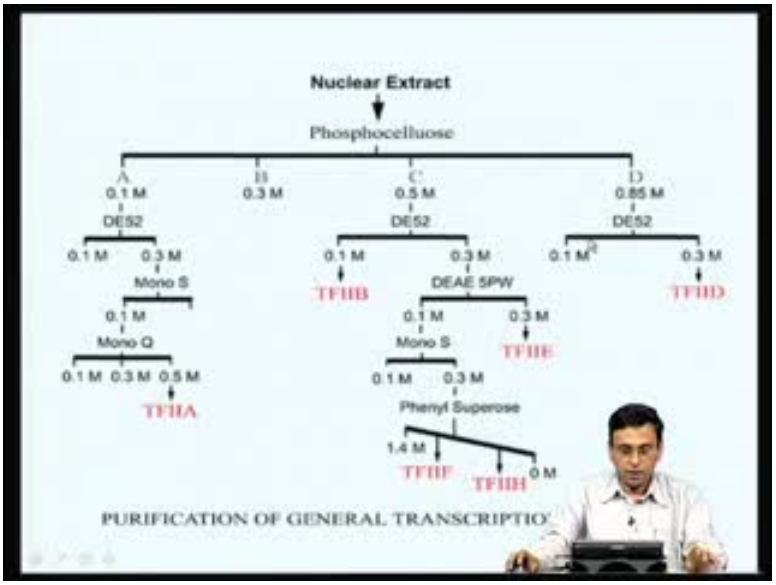
(Refer Slide Time: 06:45)



And more importantly, there are a whole bunch of a protein called coactivators or corepressors, which actually act as a mediator between the upstream activators and the general transcription activators. So, they act as a bridge between these two sets of protein

factors. So, today's talk will focus only about the general transcription factors which bind to the core promoter elements, and how variations of general transcription factors can also contribute to differential gene regulation. Now, so, we mentioned in the first class, very briefly, how RNA polymerase, in conjunction with 6 general transcription factors, namely, TFIIA, B, D, E, H, and F, can actually bring about transcription initiation in eukaryotes, right? Now, I also mentioned, very briefly, in the first class, each one of these general transcription factors has one to many subunits, and they have different molecular weights, and each of them actually have a different function.

(Refer Slide Time: 07:25)

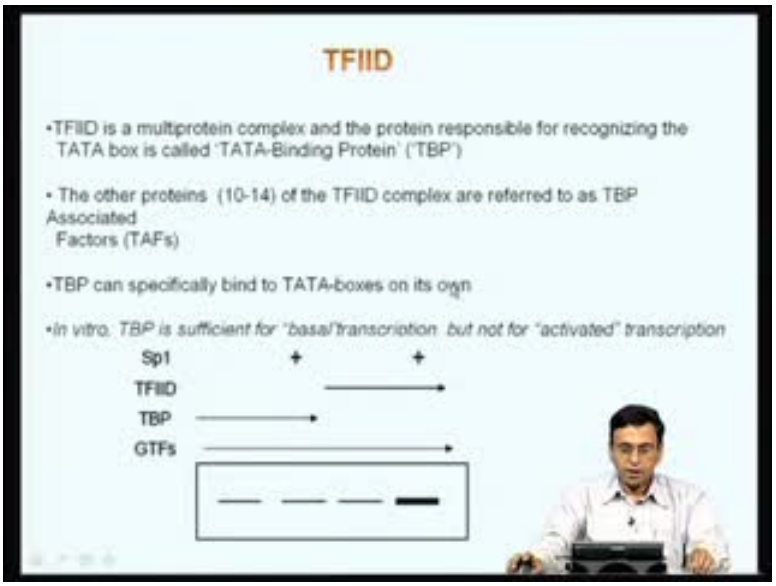


For example, the TBP component of TFIID, actually, is involved in promoter recognition, and its job is to, actually, also recruit the TFIIB, and they revolts. The TFIID, in addition to TBP, also contains what are called TBP associated factors, and they, actually, play a very important role in transcription activation and promoter recognition. You have TFIIA, B, F, and E, and each of them have a specific function during transcription initiation, which we discussed in the first class.

Now, I also mentioned, very clearly, that enormous amount actually on time and effort is, actually, spent in purifying each one of these general transcription factors. For example, people actually have taken nuclear extracts devoid of histones and DNA, and then fractionated them on ion exchange columns, and then eluted proteins by using different salt concentrations, and then re-chromatographed in different kind of ion

exchange columns, and based on all this differential chromatography of various nuclear extractions, proteins, each one of the general transcription factor purified in different set of these column chromatography fractions, and each one of them is purified.

(Refer Slide Time: 08:37)



Then, all these purified transcription factors were actually put back, and then when you put a promoter sequence containing a downstream coding region, and we put nucleotides, we have been able to demonstrate that transcription initiation can actually be mimicked in a test tube, and this is what is called as cell-free transcription, and these cell-free transcription assays, actually, played a very important role in the analysis of promoter sequences, as well as in identifying what are all the various components of the general transcription factors. So, today, let us now start discussing about TFIID; how diversity in TFIID can contribute to gene regulation. That is what we are going to begin with, in this class.

Now, just to recapitulate about TFIID, TFIID is a multi-protein complex and a protein responsible for recognizing the TATA box in the TFIID complex is the TATA binding protein or the TBP, which we all are aware. There are many other proteins, about 10 to 14 different proteins of the TFIID complex, and these are actually referred to as the TBP associated factors or TAFs.

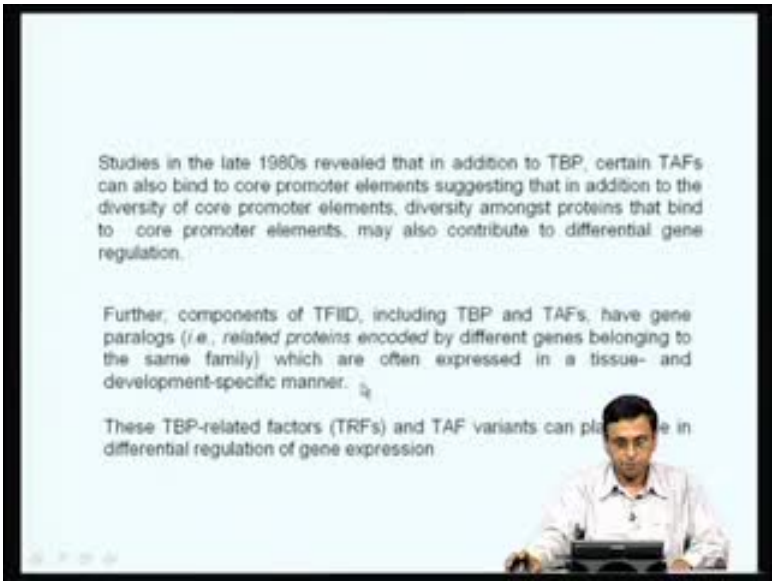
Remember this word TAFs– TBP associated factors. We are going to discuss lot more about these particular TAFs in this class. Now, TBP can specifically bind to TATA box

on its own. So, although TFIID consists of TBP and many other TBP associated factors, if we are just clone the gene coding for TATA binding protein, express the recombinant TFIID, recombinant TATA binding protein, and if you now take a TATA box sequence, the TATA binding recombinant protein can actually bind the TATA box very well. Then you ask the question– why do we require TFIID? What is the function of all these TBP associated factors?

This was actually demonstrated when they actually did cell-free transcription using this recombinant purified TDP alone, as well as TFIID purified from cell extracts, which contains TBP as well as TAFs. They found in a very important difference. Now, TBP, although can support basal transcription, TBP could not support activator-dependent transcription. That means, if you have promoter region, which contains not only the core promoter sequence, but it also has the binding site for a transcription factor.

For example, in Sp1, which is a specific transcription factor which binds to a proximal distal promoter element, when we take a such promoter promoter sequence, which contains a core promoter sequence, as well as the binding sequence of an upstream activator, and if you now put only TBP instead of TFIID, you see, TBP, the level of transcription is more or less same. Addition of TBP does not make any difference, so the Sp1 cannot activate transcription when you have only TBP. Whereas, when you have TFIID instead of TBP, then you see, Sp1 is able to activate transcription. which clearly… These kind of experiments clearly told that the TAFs– the TAFs of TFIID, actually, are essential for mediating the activator function of many upstream activators. TBP cannot substitute the TAFs function, and TBP, in combination with TAFs, are essential for activator-dependent transcription in cell-free transcription assays.

(Refer Slide Time: 11:10)



So, this was the first demonstration that the TAFs of the TFIID complex play a very important role in supporting activator-dependent transcription in eukaryotic cells. So, studies in the late 1980s revealed that, in addition to TBP, certain TAFs also bind to core promoter elements, suggesting that in addition to the diversity in core promoter elements, diversity among proteins binding to core promoter elements may also contribute to differential gene regulation.

So, this is the first point I want emphasize in this particular class. The last class, I told you, diversity in the core promoter elements itself contribute to differential gene regulation, but now I am going to tell you diversity in the TAFs and diversity in TBPs can also contribute to differential gene regulation. This was discovered sometime in the late 1980s. The components of TFIID, including TBP and TAFs, have many paralogs. Now, what are paralogs? Paralogs are nothing but, they are related proteins encoded by different genes belonging to the same family.

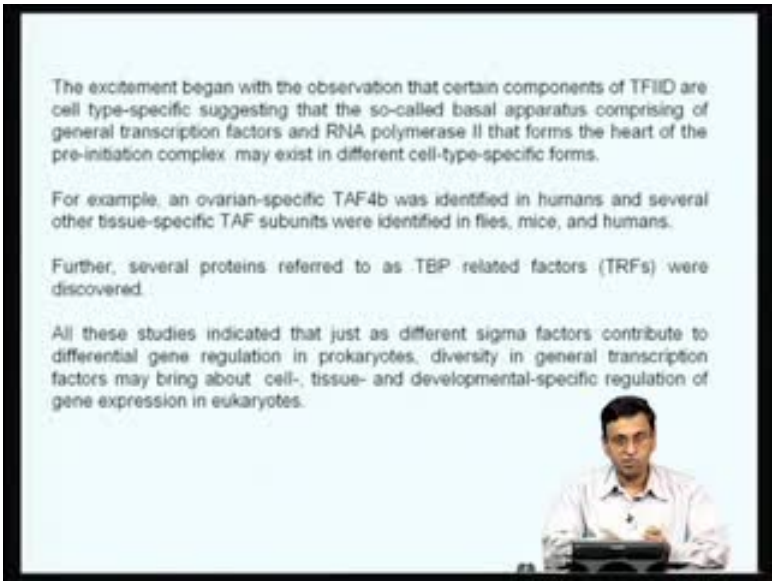
So, when, for example, when I identified the TBP in humans, and if I will have identified the TBP-related protein in Drosophila, that becomes a paralog. So, they all belong to the same TBP gene family, but there is a paralog of TBP which is a TBP-related protein, which I have identified in Drosophila. So, many paralogs of TBP and TAFs were discovered in many eukaryotic species. So, as people start looking at general transcription factors in Drosophila, humans, insect cells, and in the yeast, so on and so

forth, they start identifying that there are many members of these gene families; there are many variants of TBP; there are many variants of TAFs; and these are actually called as paralogs. And what was very interesting, when people started identifying some of these paralogs of the general transcription factors, many of this paralogs or many of these variants of general transcription factors have been found to be expressed in a tissue and development-specific manner. This was the first indication that the general transcription factors themselves are differentially expressed in certain cell types, and therefore, they can and themselves contribute to differential gene regulation.

So, just as I told you in the last class, core promoter element's job is not just to support transcription initiation. Variations in the core promoter sequences can also contribute to differential gene regulation. Today, I am going to tell you, the job of the general transcription factors like TFIID is not just to assist RNA polymerase come and bind to the core promoter region and initiate transcription, but there are variations within these general transcription factors, and these variations also can contribute to differential gene regulation.

For example, there are variants of TBP and these are all refer to as TBP-related factors or TRFs. So, variants of TBP have been identified and there seems to be, for example, certain TBP (( )), TRFs or TBP-related proteins bind to certain, only certain amount of sequences, but not the others. Similarly, there are certain TAFs– they themselves can, actually, bind to certain core promoter sequences in addition to TBP. So, all these variations have, actually, contributed to differential gene regulation.

(Refer Slide Time: 13:57)



So, the excitement in this area of eukaryotic gene expression, if say, a lot of action took place in the early seventies to 2000. That is why, there was, actually, burst of activity in eukaryotic gene regulation, and number of research groups worked on dissected out the various facets of both activator-dependent transcription, and as well as general transcription factors in model systems like Drosophila, yeast, and mammalian cell extracts, and because of these of studies, a new knowledge actually started flowing in the late 1980s upto 2000, and a number of very interesting observations, actually, came out of these research activities.

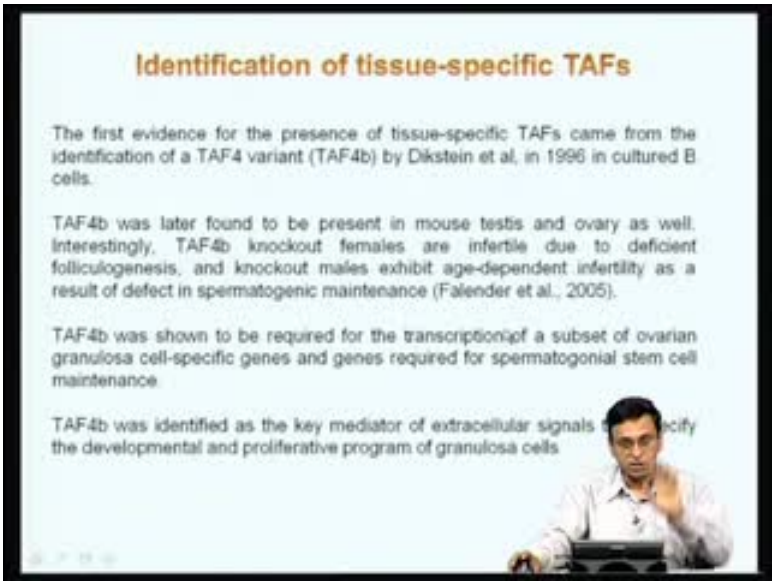
So, the excitement, actually, began with the observation that certain components of TFIID are cell type-specific. Now, so far, you might have been under the impression that general transcription factors, because they are actually present or required for transcription initiation of all protein coding genes, they must be exposed in all cell types. So, whether it make and make TFIID from HeLa cells or Drosophila or yeast cells, they should all be the same, right?

That must be the assumption. But still, people realized that when we make some of the TFIID from certain cell types, they are not the same; they are different, suggesting that there are many variants of this general transcription factors, which are expressed in a cell type in a tissue type-specific manner. So, the so-called basal apparatus, comprising of general transcription factors and RNA polymerase II, that forms the heart of the

preinitiation complex, may actually exist in different cell type-specific forms. So, the first important observation that I want to discuss with you is that there are many general transcription factors which are expressed in a cell and developmental-specific manner. So, they are not the same as we find in the normal cells, for example, in the case of TAFs, once they identified a bunch of TAFs from model organisms like Drosophila and yeast cells and the HeLa cells, and so on and so forth, they identified a new TAF called TF4b, which is actually expressed only in the ovarian cells in humans.

And this kind of a TAF, variant of TAF, TAF4b is not the exact the same of the TAF4, which is expressed other cells, since suggesting that there are, actually, variants of these TAFs which are expressed in other cell types. So, once they found these kind of things, people started looking at TAFs in other cell types and other tissues, and such variants of TAFs are identified in number of other organisms like flies, mice, and humans, and so on so forth. Similarly, the TATA binding protein is not just one protein. People also realized that when you started purifying TBP from other cell types, some other tissues and other organisms, variants of TATA binding proteins called TBP-related factors were identified, and they seem to be having functionally slightly different from the regular TATA binding protein. So, all the studies indicated one thing– just as in the first class, I told you that sigma factors switching in prokaryotes, that means, by using different sigma factors capable of binding to slightly variant minus 10 minus 30 sequences, differential gene regulation is brought about in prokaryotic cells.

(Refer Slide Time: 17:28)



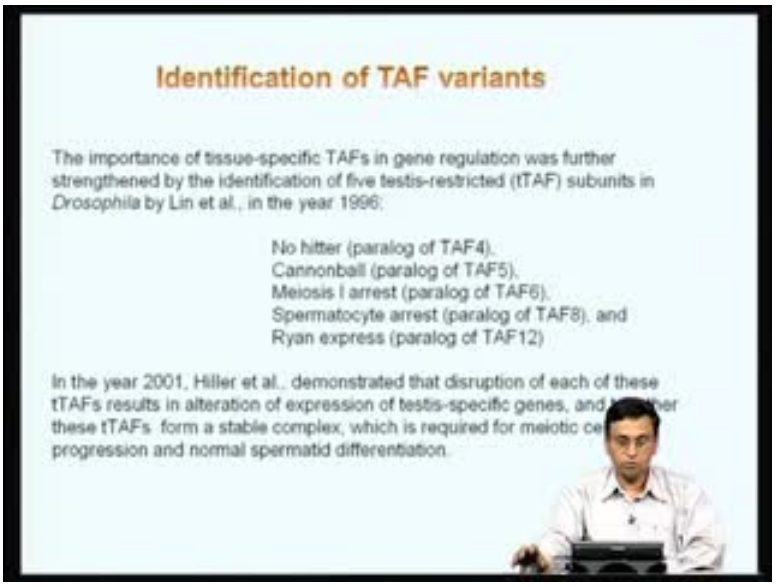
The same way, differential gene regulation eukaryotes can be brought about by diversity in general transcription factors, which are expressed in a cell tissue and development-specific manner in eukaryotes. So, this is the first point I want you to remember, in this class. Just as diversity in sigma factors contribute to differential gene regulation in prokaryotes, diversity in general transcription factors such as TBP and TAFs can contribute to differential gene regulation in eukaryotic cells. That is the first point you have to remember. So, let us now try to focus on what kind of variations in TAFs contributes to differential gene regulation. Remember, what are TAFs? TAFs are nothing but TBP associated protein factors, which are actually in the form of TFIID. So, TAFs, in association with TBP, constitu te the TFIID, and many of these TAFs, the important function of TAFs is that they act as a bridge between upstream activators and the general transcription RNA polymerases.

For many of the transcription factors like Sp1 to activate transcription, you require TAFs. If you provide only TBP, these upstream activator proteins cannot activate transcription, but the moment you provide TFIID, which contains TBP associated factors TAFs, these upstream activators able to activate transcription. So, people thought the only job of TAFs is to assist many of the upstream activators to activate transcription in a better manner, but the first evidence for the presence of a tissue-specific TAFs came from the identification of a TAF4 variant called TAF4b by Dikstein et al., in 1996 in cultured B cells. Now, TAF4b was later found to be present in mouse testis and ovary as well, and when you knockout the TAF4b gene in the females, they became infertile due to diffusion folliculogenesis, and males exhibited age dependent infertility as a result of defect in spermatogenic maintenance. So, this is very interesting; I wanted to pay attention to this.

For the first time, people have realized there are certain TAFs which are expressed in a cell and tissue-specific manner. For example, there is a particular TAF called TAF4b, which is expressed in a ovarian tissues and testis, and if you knockout this particular gene in mice, then certain gonadal functions are affected, very clearly indicating that this particular TAF may be important a set of genes which are involved in spermatogenesis and oogenesis. So, you can see, this was never foreseen, before. People thought all the TAFs are the same and their job is to assist, form a TFIID complex, and assist general transcription activator protein to bind and activate transcription, but once they start

identifying cell and tissue-specific TAFs, and these TAFs seems to be involved in regulation of specific set of genes. So, specific regulation of specific genes seems to be brought about by the expression of specific TAFs in specific tissues. So, TAF4b was shown to be required for transcription of a subset of ovarian granulosa cell-specific genes as well as genes required for spermatogonial stem cell maintenance.

(Refer Slide Time: 20:44)

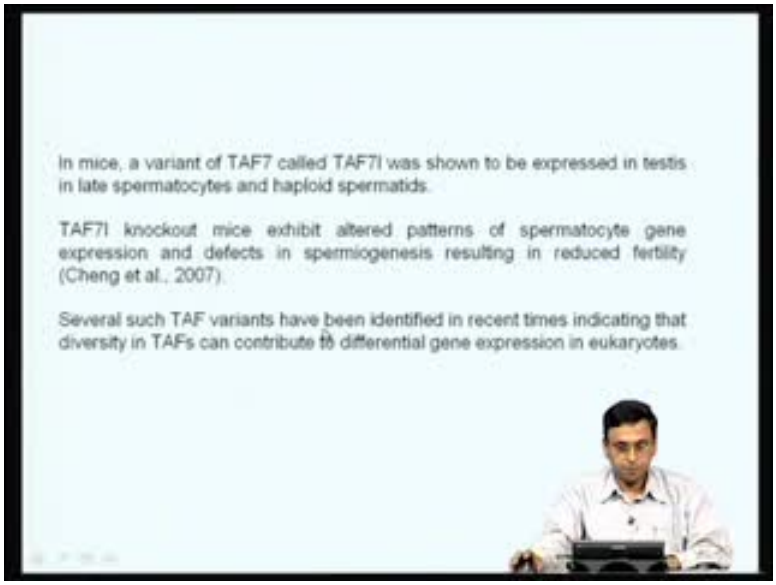


So, you can see, these particular TAFs, which are expressed only in some gonadal tissues, seems to be involved in the regulation of a subset of genes in these gonadal tissues. So, they are not general; they do not have a general function. They are not involved in general activation of, general initiation of transcription of all the gene. They seem to be involved in the activation of a specific subset of genes.

So, TAF4b was identified as a key modulator of extracellular signals that specify developmental and proliferative program of granulosa cells in the ovary. So, for the first time, a tissue-specific TAF was identified. This seems to be having a very specific tissue-specific function, and seems to be involved in the regulation of a subset of genes in these particular tissues. So, such TAFs were now identified in a number of other tissues of a number of other organisms, and the purpose of this talk is not to tell you everything or discuss about all the TAFs, but just to give you one of the examples, just to drive home the point that variations, variants of TAFs can actually contribute to differential gene regulation. That is the crux of today's talk.

So, the importance of the tissue-specific TAFs in gene regulation was further strengthened by the identification of five tissue-restricted TAFs, what are called as tTAFs– T stands for testis– testis-specific TAFs or testis restricted TAF subunits in Drosophila in the year 1996, and these five testis-restricted TAFs are, namely, one of them, for example, it has no hitter, which is actually paralog of TAF4; cannonball, which is actually paralog of TAF 5; meiosis I arrest, which is involved in the paralog of TAF 6; spermatocyte arrest, which is a paralog of TAF8; and is called Ryan express; it is a paralog of TAF 12. So, variants of these TAFs were actually identified and all of them seems to be expressed in testis, indicating that these paralogs of TAFs, actually, have a specific function in the testis. So, in the year 2001, Hiller et al. actually demonstrated that destruction of each of these tTAFs results in alteration of expression of tissue-specific testis-specific genes, and together these TAFs form a stable complex required for meiotic cell cycle progression and normal spermatid differentiation.

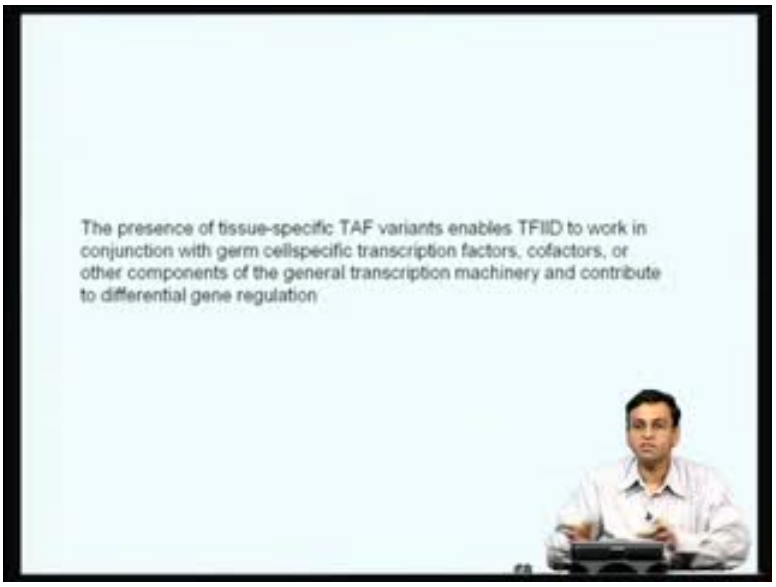
(Refer Slide Time: 22:31)



So, you can already see how the TAFs, which were actually discovered as general transcription factors required for transcription initiation of all the eukaryotic protein coding genes, it soon became clear that there are certain TAFs which are actually expressed only in certain tissues or certain cell types, and they also are involved in the regulation of only certain specific genes involved in a specific physiological process. Similarly, in the case of mice, for example, a variant of TAF7, called TAF7I, was shown to be expressed in testis in late spermatocytes and haploid spermatids. So, TAF7I

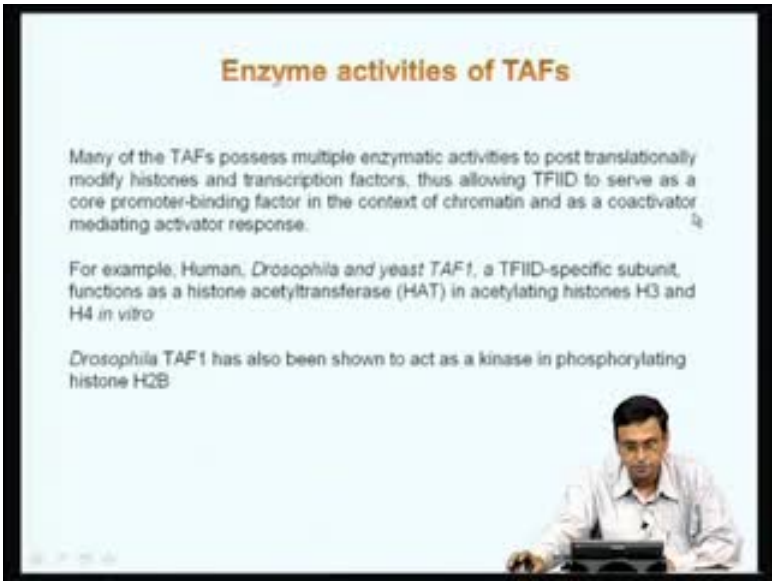
knockout mice exhibit altered patterns of spermatocyte gene expression and defects in spermiogenesis, resulting in reduced fertility. So, you can see, here are certain TAFs which are actually essential for spermatogenesis, and if you have if you block the expression of these particular TAFs it results in infertility or reduced fertility in these mice.

(Refer Slide Time: 23:12)



So, several such TAF variants have been identified in recent times, indicating that diversity in TAFs can actually contribute to differential gene expression in eukaryotes. So, the presence of tissue-specific TAF variants enables the TFIID to work in conjunction with germ cell-specific factors, cofactors, or other components of the general transcription machinery, and contribute to different gene regulation. So, you can see, again, we have to come back to the example of cricket pitch, where, for example, there is a specific cricket pitch which is suitable for a spin bowling. In the same way, there are certain promoters in certain tissues like testis, and if these promoters are actually present in, say, subset of genes, and these promoters are actually bound by specific kind of TAFs, and therefore, these TAFs function in conjunction with specific upstream activators which are actually expressed only in that particular tissue, and in combination, they actually are involved in a regulation of a specific subset of genes. So, they contribute in differential gene regulation.

(Refer Slide Time: 24:00)



The other important observation that came about, in the case of TAFs, is the discovery that TAFs actually contain very important enzymatic activities. In fact, many of the TAFs are actually identified as enzymes. This is another important discovery in the area of eukaryotic gene regulation, which actually changed the entire paradigm of eukaryotic gene regulation. So far, we have discussed TAFs has only proteins, no? Their job is to form multi subunit complexes and assist upstream activators to activate transcription, and help to form a TFIID complex. But the moment they realized that some of the TAFs, actually, have enzyme activities, that actually changed the way the eukaryotic gene expression was actually studied. Let us now understand what exactly we mean by this. Many of the TAFs actually possess multiple enzyme activities to post translationally modify histones and transcription factors, thus allowing TFIID to serve as a core promoter binding factor in the context of chromatin, and as a activator mediating activator response.

Now, so far, we have been discussing as if the DNA is a naked template. In eukaryotes, we all know that the DNA is actually present as chromatin. The histones are actually bound to the DNA template, and histones are organized in a very specific manner, contributing to a chromatin formation and chromosomes. So, people who realized that when you study transcription regulation, you have to explain how these transcription factors are regulating, not just naked DNA templates, but usually, have to explain how,

actually, these transcription factors are regulating chromo gene expression from chromatin templates.

So, the first question that came to your mind is, if the DNA is covered by histone, how is this transcription factor able to bind to DNA and activate transcription? The first thing that has to happen is you have to dislodge the histone. So, this is where when people were actually realized that some of the transcription factors, especially TAFs, they actually can modify histones. For example, the human and Drosophila is TAF one which is a TFIID specific subunit, it actually can acetylate histone.

So, this actually can function as a histone acetyltransferase or HAT, which is capable of acetone acetylating histone H3 and H4. The moment they are discovered, the some of the general transcription factors, such as TAFs, can, actually, add a phosphate or add an acetyl group to the histones, it be can very clear that the mechanism by many of these transcription factors are regulating gene expression of chromatin templates is, actually, by post translationally modifying histones.

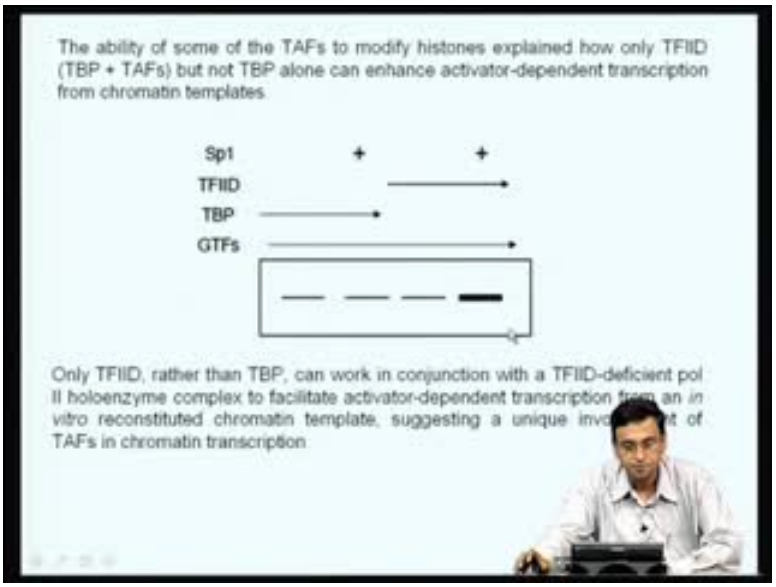
See, very simple. Now, histones are basic proteins, so DNA is negatively charged. So, the basic protein, which has a positive charge can bind very tightly to the DNA, but the moment these TAFs go and add acetyl group to the histones, now the histones are becoming negatively charged, and therefore, this negatively charged histone cannot bind very tightly to negatively charged DNA. Therefore, the histones fall off, and now, these transcription activators can now bind, preinitiation complex can be formed very clearly, and therefore, transcription initiation can takes place.

So, the discovery that TAFs actually contains specific enzyme activities changed the whole field of eukaryotic gene expression. People recognize that the assembly of preinitiation complex and the binding of specific transcription factors is actually facilitated by selective removal of histones by some of these transcription factors, by post translationally modifying histones. Very simply, if you want to shut off the expression of by particular gene, you simply add a phosphate. You simply add a positive charge; you methylate histones. So, DNA histones becomes more positively charged, so they will go and bind DNA much more tightly, whereas if you want to activate transcription, you add a negative charge on histone. Therefore, histone falls off, and now,

the preinitiation complex formed very nicely, and therefore, transcription initiation can takes place.

So, by simply making histones bind or fall off by either adding a positive charge, or by adding a negative charge, or by removing the positive charge, or by removing the negative charge, by using what are called as histone acetyltransferases and histone deacetylases, histone phosporylases, and so on and so forth, differential gene regulation can be brought about.

(Refer Slide Time: 28:30)



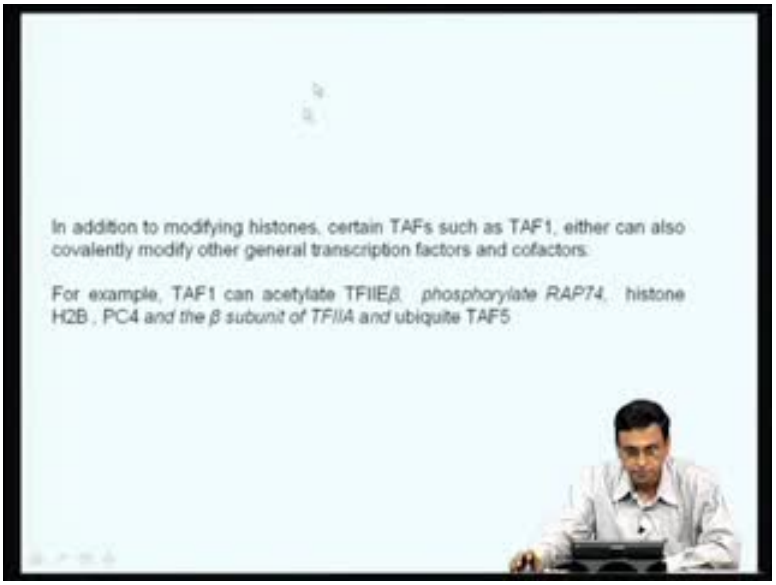
And one important factor, player in this entire area are the TAFs. So, many of the TAFs soon discovered to be enzymes capable of post translationally modifying histones. Similarly, for example, the TAF1 also has been able to act as a kinase by actually phosphorylating histone. So, many of the TAFs can act as a histone acetylases, they, or they can actually act as a kinases by they are acetylating histones or phosphorylating histones. So, this discovery now explained, very clearly, why certain TAFs are actually required for activation by certain transcription factors. Now this slide we are actually discussed in a last line, where I gave an example how TBP– purified TBP or recombinant TBP– can only support basal transcription, but when the TBP cannot support activator-dependent transcription, you can see here, in the first lane, we have only the promoter sequence, so we have only the basal transcription and general transcription factors.

When we add TBP and Sp1, there is no change in the level of transcription here. Similarly, when you add Sp1 to Sp1 cannot actually change the basal transcription, but instead of TBP, if I now add TFIID, and you can see, Sp1, in conjunction with TFIID, can enhance transcription. That means, for Sp1 to activate transcription, you cannot TBP alone cannot do the job. You require TBP plus TAFs, which is, actually, the TFIID.

Now, we explained what TF what the TAFs are actually doing. So, in order for Sp1 to activate the transcription of chromatin templates, some component of the TAFs are actually required, and these TAFs are, probably, are removing histones in the vicinity of the promoter, and therefore, it can facilitate preinitiation complex, and thereby, it facilitates Sp1 to activate transcription from these promoters. So, ability of some of the TAFs to modify histones explain how only TFIID, which consists of TBP and TAFs, but not TBP alone, can enhance activator-dependent transcription from chromatin templates.

This is the very important observation that changed in the entire field of eukaryotic gene regulation. So, only TFIID, rather than TBP, can work in conjunction with TFIID-deficient pol II holoenzyme complex to facilitate activator-dependent transcription from an in vitro reconstituted chromatin template, suggesting that a unique environment, unique involvement of TAFs in the chromatin transcription.

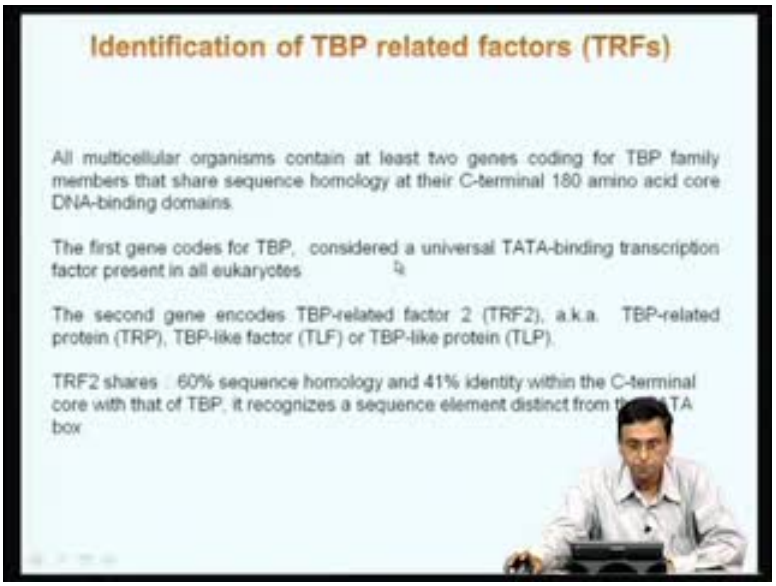
(Refer Slide Time: 30:42)



So, the moment people started studying chromatin templates and asked the question how general transcription factors and activators activate transcription, the observation that

chromatin-dependent transcription, actually, involves modification of histones, and many of the transcription factors, especially general transcription factors, actually, contained histone modifying activities. This, actually, changed the entire field of eukaryotic gene regulation. So, in addition to modifying histones, many of the TAFs, such as TAF1, either can covalently modify other general transcription factors and cofactors as well. I will give you one example. TAF1, for example, can deacetylate some one component of TFIIE. It can also phosphorylate RNA polymerase associated protein called RAP74. It can also modify histone H2B, PC4, and the beta subunit of TFII4, and also what is called as ubiquitin (( )) of TAF5.

(Refer Slide Time: 31:35)

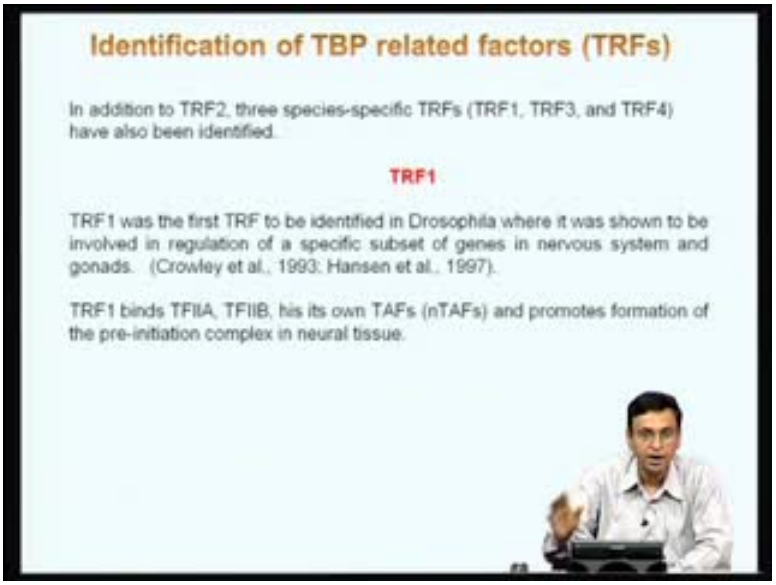


So, what these examples tell you is that the TAFs, or the TBP associate proteins, actually, are enzymes. Many of them– not all– many of them are enzymes capable of covalent modification of number of other proteins, and these covalent modifications can be either acetylation, it can be phosphorylation, ubiquitinylation, so on and so forth; and all these things have a very important role in gene regulation.

So far, I told you TAFs– the TBP associated proteins, actually, can have enzymatic activities and there are actually variants of TAFs. In certain tissue types and some cell types, specific forms of TAFs are expressed, and these specific TAFs can actually bring about tissue-specific and cell or development-specific gene regulation. Now, I am going to tell you, in addition to TAFs, there are variations of TBP itself.

Now, let us look at it. Now, all multicellular organisms contains at least two genes coding for TBP family members that share sequence homology at their C-terminal hundred and eighty amino acid core DNA binding domains. As you know, TBP is a TATA box binding protein. That means, it need to have what is called a DNA binding domain, and this DNA binding domain is actually present in C terminus of the TATA binding protein, through which it actually binds and recognize the TATA box. And usually, the TBP there when you identifying genes coding for TBP, this C-terminal domain is actually consult in many of them the first gene which was identified as a one the coding for TBP, which is actually considered as a universal TATA binding transcription factor, is present in all eukaryotes, and its job is to bind the TATA box. Now, the second gene encoding a TBP related factor called TRF2, which is also known as TBP related protein, TRP; TBP like factor, TLF; or TBP like protein, TLP; these are all various nomenclatures for the same protein.

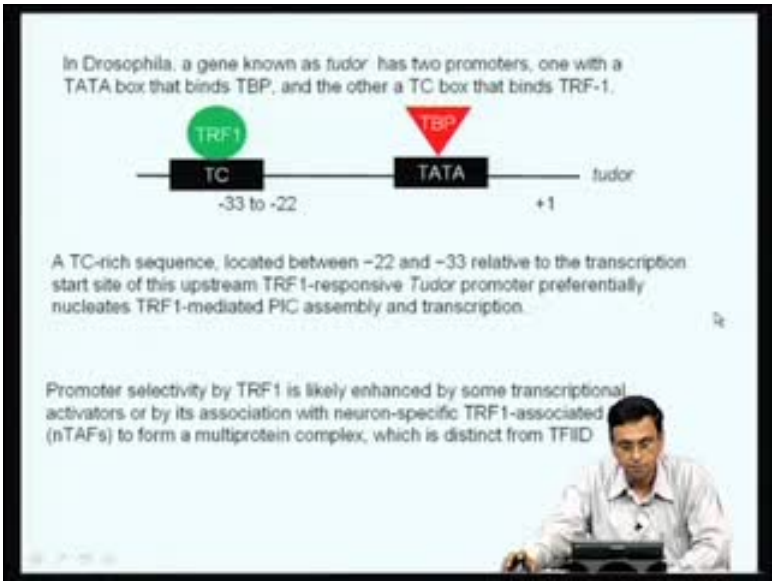
(Refer Slide Time: 33:30)



And this TRF2, actually, shares about 60 percent sequence homology and 41 percent identity within the C-terminal domain of TBP, but very importantly, this TRF2 or the TBP related factor, actually, recognizes a sequence which is quite different from the TATA sequence. So, you can see, a variant of TBP itself can contribute to differential gene regulation, because it recognize the sequence that is quite different from the TATA box.

Although it has sequence identity with TBP functionally, it actually recognize the slightly different sequence than the TATA sequence. So, this identification of TBP related factors again are in a new paradigm in a gene regulation, wherein variants of TBP themselves can activate different genes, depending upon what kind of promoter, core promoters sequence they recognize. So, in addition to this TRF2, three species-specific TRFs were actually identified. These are called TRFl, TRF3, and TRF4, and the TRF1 was actually first identified in Drosophila and was shown to be involved in regulation of a subset of genes involved in nervous system. It is expressed only in nervous system of many of these genes and involved in regulation of specific genes in the nervous system and gonads.

(Refer Slide Time: 34:38)

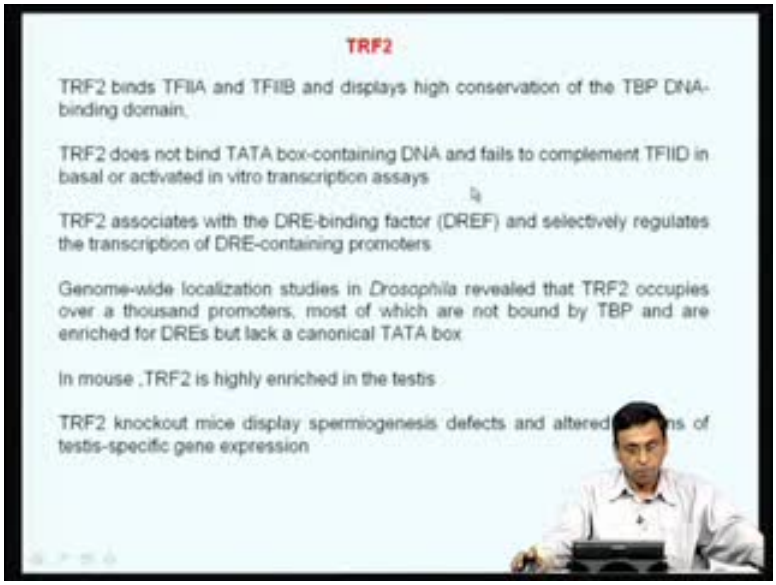


The TRF1 actually binds TFIIA. TFIIB– it has its own set of TAFs which are called as nTAFs, which are specifically expressed in nervous tissue, and actually promotes the formation of the preinitiation complex in nervous tissue. So, we can say, very clearly, there is a specific form of TBP which is called TRF1, which in association with TAFs, which are expressed in specific nervous tissue can regulate a subset of genes in the nervous tissue.

So, tissue-specific gene regulation is brought about by variants of TBP and associated TAFs. For example, in Drosophila, a gene known as tudor has two promoters– one with a TATA box that binds TBP, and another with a box called TC box that actually binds

TRF1. So, TRF1, TBP, both are belong to TATA binding protein family, but the TBP, actually, it is the TATA sequence, whereas the TBP-related factor actually recognize a different sequence called as a TC box, which is present. This is present about minus 33 to minus 22, whereas this is present much closer to transcription start site, and this TC-rich sequence, located between minus 22 to minus 33 relative to the transcription start site, serves as the binding site for the TRF1.

(Refer Slide Time: 35:42)



So, the promoter selectivity by TRF1 is enhanced by some transcription activators, as well as by its association with neuron-specific TRF1 associated factors, and therefore, the same gene, depending upon whether it the TRF1 binds with this sequence or TBP-binding sequences, can be differentially regulated. So, there is variations in the both cis-acting elements as well as the general transcription factor binding to these sequences. Similarly, another TBP related protein called the TRF2 was identified. This TRF2, again, binds TFIIA and TFIIB, and, actually, displays very high conservation in the DNA binding domain. Remember, all these TATA binding protein family of genes– they are highly conserved in the DNA binding domain, but they do recognize variants of this TATA box sequences, and have either specific tissue-specific or cell type-specific recognition, and their interaction is specific subset of TAFs, actually, also, are different.

So, TRF2 does not bind TATA box containing DNA. Just like the TRF1 binds with different sequences called TC box, the TRF2 also does not bind a TATA box containing

DNA, and it fails to compliment TFIID in basal or activated in vitro transcription assays, and remember, this is, again, very important, when you say TBP, when you purify, for example, TFIID from a specific cell type, and then if you replace the TBP present in TFIID with a recombinant TBP it will still act as a TFIID, but instead of TBP, if you add TRF2 to TFIID complex, it no longer substitutes the function of TBP, saying that TRF2 is actually variant of TBP and it has a function that is different from TBP. The reason why it cannot substitute to TDP the TFIID complex is because TRF2 cannot bind to the TATA sequence.

That the job of TFIID is to activate transcription from a TATA binding, TATA box containing promoters. The TRF2 cannot bind the TATA box, and therefore, cannot substitute TFIID, substitute for TBP in a TFIID complex. Instead, TFIID actually associates with the DRE binding factor. It is another upstream activation factor, and selectivity regulates transcription from a DRE containing. Remember my last class, where I told you there are variations co-promoter sequences.

Some of eukaryotic promoters, instead of the TATA box, actually, contains what is called as a DRE motif, and here I am telling you, there are protein factors which specifically bind to the DRE sequence, and a variants of TATA binding protein is actually involved in this transcription activation from these DRE containing promoters. So, this TRF2, actually, recognizes a DRE sequence, instead of a TATA box. In fact, when people carried genome-wide analysis, now, there are now techniques to analyze what kind of protein factors can actually bind to what kind of sequences by techniques called chromatin immunoprecipitation and so on and so forth.

When they carried out analysis to find out what kind of promoter sequences are actually occupied by TRF2 in Drosophila in vivo, when they actually do some of these studies, they found that several thousand promoters are actually occupied by TRF2 in vivo, clearly indicating that the TRF2 has a very important function in vivo.

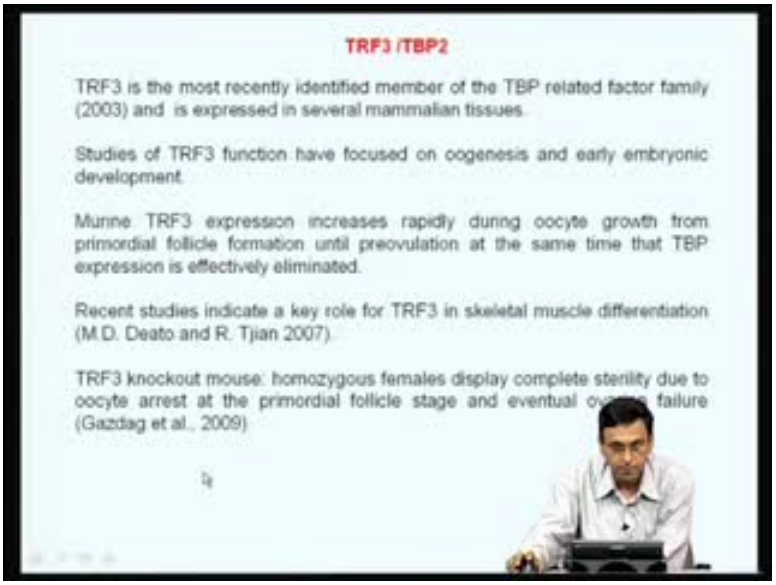
And very interestingly, many of these promoters, which are occupied by TRF2, is not bound by TBP. That means, there are subset of genes in vivo which are occupied by TBP. There are another subset of genes in vivo, which are occupied by TRF2, clearly indicating that TBP and TRF2 may be regulating different sets of genes in Drosophila, and, very interestingly, many of these promoters, which are actually occupied by TRF2,

they are contained DREs, but they lack a canonical TATA box. So, very clearly, that means, the TRF2 is actually involved in the regulation of those eukaryotic promoters which are TATA-less promoters, that which actually contain DRE instead of a TATA box.

So, we can say how variation in TATA binding protein can actually contributes to differential gene regulation. There are genes which contains TATA boxes, and such genes are regulated by TFIID containing the normal TBP; and there are different sets of promoters, which instead of TATA box, contain a DRE and such genes are regulated by this DRE AF and TRF2, which actually occupies such promoters and initiates transcription initiation.

Very interestingly, when you knockout this TRF2, and in such mice they, actually, they actually display defects in spermiogenesis and altered patterns of testis-specific gene expression, clearly indicating that this variant of TBP, namely, TRF2, is actually involved in the regulation of certain subset of genes that are essential in spermatogenesis or spermiogenesis. So, if you knockout this TRF2, the expression of these genes involved in spermiogenesis is blocked, and therefore, you have defects in spermiogenesis. So, you have a tissue-specific function for TRF2.

(Refer Slide Time: 40:00)

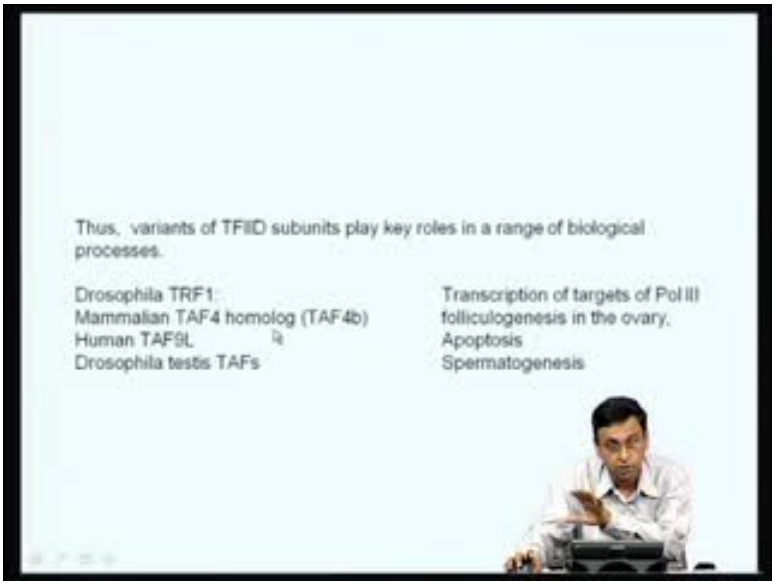


There are other factors called TRF3, or which is also called as TBP2, which are actually discovered. These were discovered as recently as 2003, and the TRF3 was extensively

studied, again, in gonad tissue, as well as early embryonic development, and in the case of mice, the TRF3 expression was showed to be increase very rapidly during oocyte growth from primordial follicle formation until pre-ovulation, and at the same time, TBP expression is effectively eliminated. Very important, you see how differential gene regulation takes place during in the, during in the oocyte growth in the mice. It turns out, during in the primordial oocyte formation, the expression of TRF3 actually goes up, and the same time, the expression of TBP comes down, indicating that the TBP function is taken over by TRF3 at this specific stage of oogenesis, and many recent studies actually indicate that this TRF3, actually, as recently as 2007, Deato and Tjian, actually, have shown, TRF3 has a very important role in skeletal muscle differentiation.

So, I am telling this example to just to emphasize the fact that there are variants of TBP. We seem to be having a very tissue-specific and cell type-specific or development-specific regulation, because their expression pattern is different than the normal TBP, and in the… if we knockout TRF3 mice, the homozygous females display complete sterility due to oocyte arrest at the primordial follicle stage, which tells you clearly that the TRF3 has a very important role for further development from the primordial follicle stage onwards, and if you knockout oocyte development cannot proceed beyond the primordial follicle stage, and therefore, it leads to ovarian failure, clearly indicating that specific forms of TBP, or variants of TBP, have a very important role in oogenesis in mice. So, if you knockoff this TRF3, oocyte differentiation cannot be proceed beyond a certain point.

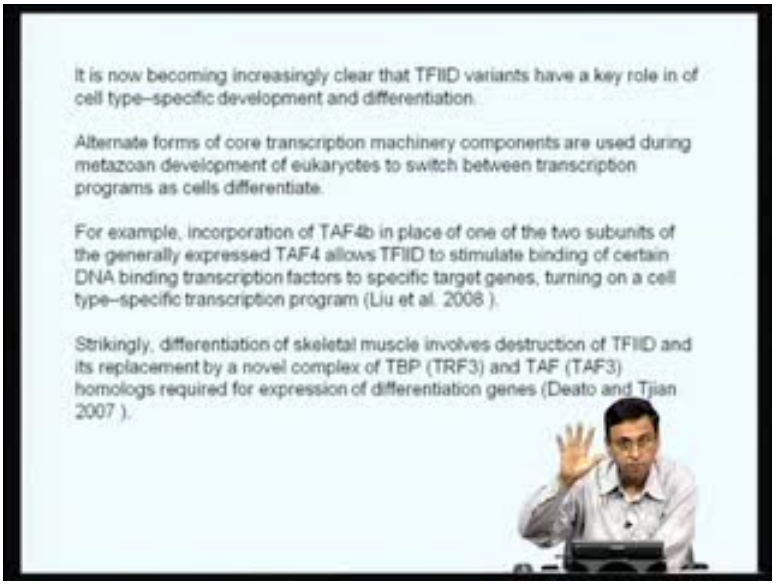
(Refer Slide Time: 42:01)



So, variants of TBP, TFIID subunits, whether it is TAFs or TBP variants, they play a very important role, and I have given you some examples. Here, in case of TRF, and in the case of Drosophila, it has been shown that TRF1 actually essential for transcription of pol III as well certain promoters of pol III, RNA polymerase III promoters are also regulated by TRF1, and in the case of TAF4B, it is essential for folliculogenesis in the ovary.

In the case of mammals, there is a human variant of TAF9 called TAF9L, which is actually involved in apoptosis, and in the case of Drosophila, there are certain variants of TAFs, which are expressed in testis, and these testis specific variants of TAFs are very important in spermatogenesis.

(Refer Slide Time: 42:47)



So, what I what I told you, so far, is that variants of TAFs and tissue-specific or cell type-specific expression of TAFs, as well as TBP variants, can contribute to differential gene regulation.

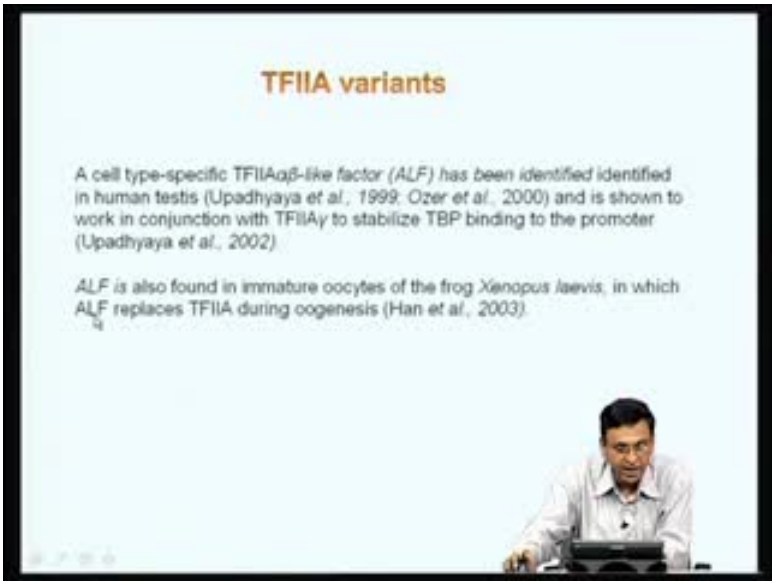
That is the crux of what I told you so far, and it is now becoming very clear that TFIID variants have key role in cell type-specific development as well as differentiation. Alternate forms of core transcription machinery components are used during metazoan development of eukaryotes to switch between transcription programs as cells differentiate. So, you can see, in the case of eukaryotes, by simply expressing different forms of general transcription factors, you can bring about differential gene regulation.

By simply expressing TRF1 or a TRF2, or specific forms of TAFs, certain subset of genes can be activated or inhibited in different tissues and different cell types. So, differential gene expression can be brought about by differential expression of different general transcription factors, especially, TAFs and TBP variants. So, incorporation of TAF4b, in place of or two subunits of TAF4, allows TFIID to stimulate binding of certain DNA binding transcription factors to specific target genes turning on a cell type-specific transcription program.

So, for example, so of this gonadal tissue by simply expressing its TAF4 variant. This TAF4 variant is now able to interact with specific subset of transcriptional factors, and therefore, is able to drive a specific developmental program or a specific cell type-

specific fate in that particular time or space during development. Similarly, differentiation of skeletal muscle involves destruction of TFIID and replacement of a novel complex of TFIIB containing TRF3 and TAF3. This was discovered, recently, in 2007. What was shown, very clearly, in this particular paper by Deato, Tjian, is that during skeletal muscle differentiation, the normal TFIID is destroyed and the normal TFIID function is taken out by a specific TBP called TRF3 and a specific called TAF3, and they all, actually, now go and activate a subset of genes resulting in differential of a skeletal muscle. So, you can see, the variants of TAFs and various TBP, by selectively expressing in a specific time point during development and differentiation, specific genetic programs can be brought about. Specific genes can be activated or repressed.

(Refer Slide Time: 45:03)



Similarly, there are also examples, wherein variants of TFIIA can also contribute to differential gene regulation. Not only TFIID, not only TAFs or TBP variants– even variants of TFIIA can contribute to differential gene regulation. For example, a cell type-specific factor called TFIIA alpha beta like factor called ALF has been identified in human testis as early as the late 2000, and this was shown to work in conjunction with TFIIA gamma to stabilize TBP binding to the promoter. So, not only there are variants of TBP, variants of TAFs. There are also variants of other general transcription factors, which can contribute to differential gene regulation, and this variant called ALF is also found in immature oocytes in frog, Xenopus laevis, in which ALF replaces TFIIA during oogenesis.

(Refer Slide Time: 46:27)

Now, what I told you, so far, is that how specific stages during development, or specific stages during differentiation in specific tissues, especially nervous tissue, gonadal tissues, etcetera, so specific variants of general transcription factors are expressed, and such variants of the general transcription factors, in conjunction with certain upstream activators, can activate a subset of genes contributing a specific differentiation program like muscle differentiation, gonadal differentiation, nervous system, or expression of specific tissue during nervous system, and so on and so forth.

So, I think what I told you, so far, in this particular class, is that just as sigma factors switching in eukaryotes contributes to differential gene regulation by different sigma factors recognizing different set of minus 10 or minus 30 sequences, in the case of prokaryotic promoters, general transcription factors can also be expressed differentially in eukaryotes, and this differential expression of general transcription factors can also contribute to differential gene regulation in eukaryotes.

So, I told you, very briefly, about how tissue-specific or developmental space-specific expression of certain variants of TAFs can contribute to differential gene regulation, and I also told you, how variants of TBP itself, like TRF1, TRF2, TRF3, can also result in activation of a different set of genes at different time points, or different tissue. So, the important point that you have to remember is that TFIID means, it is not just there just to initiate transcription initiation from pol II promoters. There are variants of TFIID; there

are variants of TFIIA; and there are variants of other general transcription factors, and by selectively expressing these variants of TBP or TAFs at different time points during development in different cell type and different tissues, different set of genes can be activated or repressed, contributing to cell type or tissue-specific regulation of gene expression. This is the crux of today's talk.

So, as in the previous class, I have also given the number of original research articles here, which, actually, are the original research articles that are actually demonstrated how expression of different TAFs or different TBPs can actually contribute to differential gene regulation in eukaryotes. And other important aspect that we discussed today, and which is also mentioned in many of these papers, is that many of this TAFs, actually, contain enzyme modifying activities, that are, actually, enzymes. Many of the TAFs can be, actually, the phosporylate histones; they can acetylate histones, and this histone modifying activity of TAFs is playing a very important role in the activation of transcription in the chromatin context.

So, by their ability to modify histones, these TAFs can either make histones bind DNA more tightly or loosen histones from the chromatin, and this actually facilitate upstream activators to go and bind transcription factors, bind those specific elements, and preinitiation complex can be assembled in the vicinity of these promoters, and this removal of histones is actually a very important component of transcription activation of eukaryotes.

So, transcription regulation of eukaryotic promoters, basically, takes place by modifying histones, and the job of many of these TAFs is, actually, act as a bridge between general transcriptional factors and the upstream activator proteins, and the upstream activator proteins, many times, the ability by which they can activate or modulate enhancing transcription, they actually recruit these TAFs, and these TAFs, in time, modify the histones or remove the histones in vicinity of the TATA box or the transcription start site there, so that the general transcription factor can come and bind to the TATA sequences and initiate transcription, and enhance the rate of transcription. So, I think I will stop here, clearly telling you that the differential expression of general transcription factors, or their ability to modulate, or their ability to bind different sets of core promoter elements, can also contribute to differential gene regulation.

So, in the first three class, I have basically covered various aspects of basal transcription of eukaryotic protein coding genes, how RNA polymerase II assembly takes place in protein coding genes, how, in the first class, I told you there is an RNA polymerase II, which, in conjunction with general transcription factors, form the preinitiation complex, and the second class, I told you that there are variations in the core promoter sequences. There are variations of TATA box; there are very other many other core promoters sequences. And variation in this core promoter sequence itself can contribute to differential gene regulation. And the third class, I told you, just like you have differential sigma factors regulating gene differential gene expression in prokaryotes, different expression of… differential expression of general transcription factors like TAFs and TBP-related proteins can actually contribute to differential gene expression, and therefore, one important facet of eukaryotic gene regulation involves a general transcription factors around the core promoter sequences.

In the next two classes, we are going to talk about the role of general transcription factors in the activation of transcription by RNA polymerase I as well as RNA polymerase III, because transcription initiation means not just protein coding genes; we also have to talk about transcription initiation from pol I and pol II, how ribosomal RNA genes are transcribed, how tRNA genes are transcribed. So, we discuss little bit about how tRNA gene preinitiation complex is assembled on the tRNA gene promoters, and how preinitiation complexes are assembled on the rRNA promoters. We are going to talk about general transcription factors are actually involved in this, and with this, we more or less covered the basic aspects of transcription initiation in pro eukaryotes, and once we covered them, we are going to discuss the transcription initiation also plays a very… the general transcription factors are not only involved in the transcription initiation, but they are also involved in transcription elongation, transcription termination, polyadenylation, and so on and so forth.

So, it is not just enough if the train starts off from a station, but it has to monitored throughout, and ultimately, it has to reach the destination. Still, there a number of process are involved. Similarly, the general transcription factors are not only involved in transcription initiation, but they also contribute to transcription elongation, transcription termination, mRNA capping, polyadenylation, and so on and so forth. So, with all maybe in next three classes, we would have covered more or less all aspect of transcription

initiation, and now the RNA how organize actually synthesize from normal promoter. After this, we are going to talk about upstream activation sequences; how and upstream transcription factors how binding of this upstream transcription factors to specific sequences can bring about modulation of gene expression, and how these transcription factors can modulate the activity of general transcriptional factors binding to the core promoter elements, and what is the role of histones in all these things, and we are also talk about epigenetic regulation of gene expression. I think I will stop here.