RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

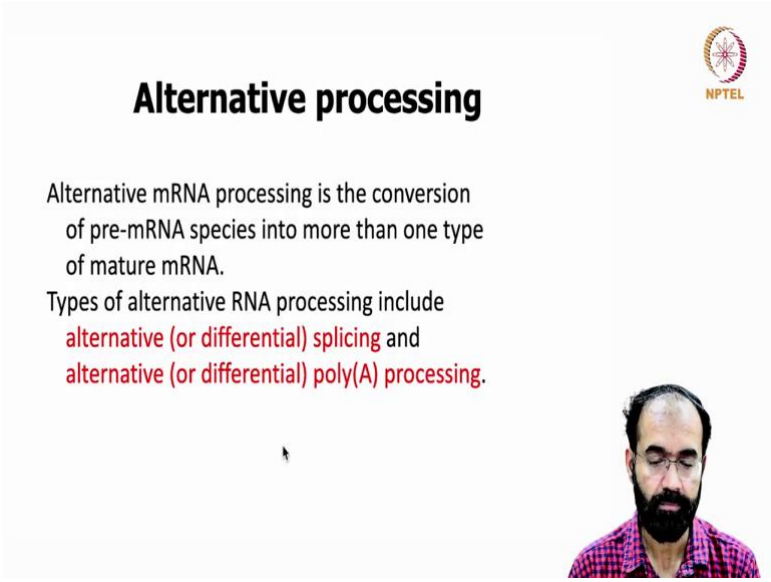
Lecture - 18
Alternative RNA Processing and Editing: Alternative Splicing

(Refer Slide Time: 00:21)

Welcome back to another session of RNA Biology and we were here in last class about alternative mRNA processing and it can do in four different ways and one is the name itself says alternative processing of different promoters, we will see them in detail and utilization of alternative poly adenylation sites. And then alternative splicing; that means, utilization of differential exon intron boundary; that means, it can incorporate few bases from the intron into the exon or it can include some part of the exon as intron.

So, different ways are possible and then comes lastly comes the RNA editing that is you are changing the existing RNA to perform something else. So, all these ones we will see them step by step because it is quite interesting and it is quite exciting you realize that you know how systematic the world has evolved and how systematic high intriguing the evolution has taken place at molecular level.

(Refer Slide Time: 01:36)



What is alternative processing? Alternative mRNA processing is the conversion of pre-mRNA species into more than one type of mature RNA and if it gets converted into more than one type of mature RNA, you end up getting more protein or little bit different protein. When you say different protein that does not mean that it is entirely different say if a protein have got 100 amino acid you make 10 amino acid change then technically it is a different post-protein.

So, I am not saying that in order to have a change you need to have 10 amino acid change. Say one simple example I will tell you all of you would have heard about sickle cell anaemia. Sickle cell anaemia is a condition in which the RBC red blood cells can form the shape of sickle.

Normally we know it is a concave circle like it look like a circle little fat circle which has got central portion little bit pressed that is little bit concave on either side. So, that is the typical shape of the RBC red blood corpuscles which is the haemoglobin bearing molecule which because of which you are able to deliver oxygen to various parts of the body. So, in the sickle cell anaemia the RBC instead of their circle shape, it can become shape of a sickle means, extremely elongate and it can form the shape of a sickle.

Sickle is basically circular sword which you use for you know cutting weeds etcetera. Farmers use it that is the you can look about how a sickle will look like. It is like you know semi lunar shape or something you can say, but it is definitely not circle. It has got

huge implication, it can clog your coronary arteries and you can end up getting you know a heart attack etcetera.

When you have more sickle cell shaped cell. However, a RBC will become sickle cell shaped only when there is a reduced oxygen tension like if you go to high altitude like this was initially discovered during the First World War when in the First World War or Second World War I do not know exactly.

But one of the World War when there were pilots were flying in very high altitude suddenly they were you know they are getting a cardiac arrest and then when they analysed their blood they found out that you know the reason is sickle cell shaped cells in their blood stream which is something more prevalent in Africa and also in India it is prevalent in people who are living in some you know endemic groups especially who are like some types of tribes etcetera it is prevalent and there is also can be done some screening for sickle cell anaemia.

But now how this is possible a single amino acid mutation? That is a place where glutamic acid should be there that place a valine comes and one single base amino acid change can create havoc in the structure of the haemoglobin and that can create the death of the individual if they are not lucky.

However, it comes with a interesting advantage you know the malaria parasite which grows usually in the RBC. So, but if you have a sickle cell RBC the malaria parasite cannot grow inside that it is somehow find it less attractive or do not able to grow. So, what happens in a malaria endemic area those individuals with the sickle cell RBC will be able to make a better living.

So, you can understand a malaria endemic area the sickle cell shape or means, do not think that any given time they will have the sickle cell shape only when there is low oxygen tension and there is low oxygen then only this gets otherwise the haemoglobin or the shape of the RBC look the same.

So, but the malaria parasite is not able to grow. So, there is a trade off you have a sickle cell phenotype, but you get advantage over malaria infection your malaria parasite will not be able to kill you because of the sickle cell RBC you still have some RBC left which is resistant to malarial infection.

So, this is a kind of trade off I gave this example just to you for you to understand that even if one amino acid change is there and we will see some example with RNA editing many a times RNA editing happens at one amino acid or one you know one base level which may bring in another amino acid; because of which you end up getting a advantage and some of this editing are so important that the organism will be able to leave if the editing is stopped organism will die embryonic development will not happen.

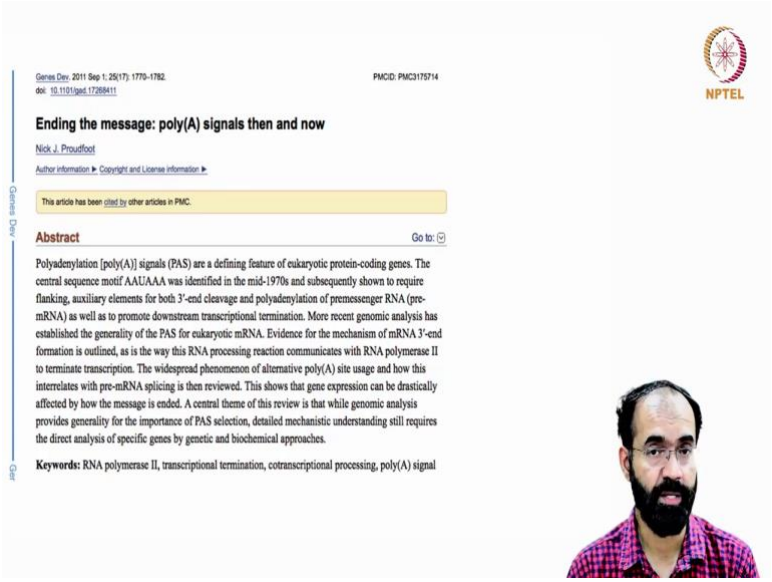
So, it is so, crucial. So, alternative processing means it can cause change in the amino acid sequence eventually the amino acid sequence in the in that when you when the organism is doing the protein translation ok. So, alternative mRNA processing is the conversion of pre mRNA species into more than one type of mature RNA the purpose is quite clear you should be able to produce diverse proteins.

And types of alternative RNA processing include alternative or differential splicing and alternative or differential poly A processing both can give rise to differential outcome; that means, a tissue A and tissue B you have because of this differential processing tissue A can have protein A and tissue B will have protein B.

However, both have come from same gene they are not coming from different gene. So, it is a clever way of enhancing the complexity. One more thing I would like to tell you when human genome is sequenced it is assumed that around 23,000 genes are there, but a complex organism like human to have around one lakh genes are supposedly be present.

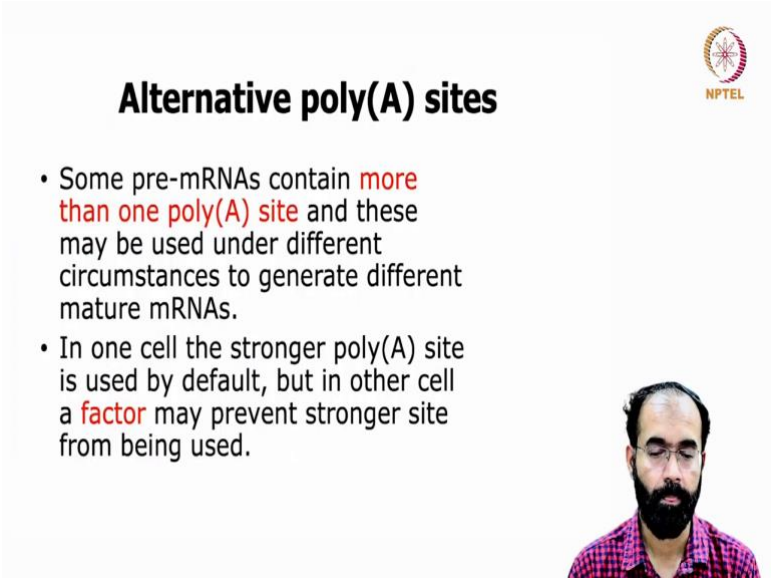
So, 23,000 is only one by fourth of one lakh. So, where from that extra 75,000 genes are coming from? So, it is believed this alternative way of producing different types of protein could be the only way in which you can have variety of proteins and you will be able to perform just like remember about the Swiss Army Knife example which I gave in the previous class. You are able to perform multiple task with one mRNA means multiple proteins can be produced with one mRNA.

(Refer Slide Time: 08:36)



So, those who are interested can read this article that is ending the message poly A signals then and now. So, you will know more in detail how poly A signals are effectively made use of in the diversification of a RNA species from a single gene.

(Refer Slide Time: 08:58)



So, what are alternative poly a sites? So, same pre mRNA contain more than one poly A site. poly A site we know AUAAA or AAUAAA. So, this is the poly A site based on which the poly A polymerase will decide which area onwards they should be start adding the poly A tail. So, it is basically an identification mark also for the endonuclease

around 18 base pairs 16 to 18 base pair downstream of the poly A site there will be a cleavage and the RNA will be released which will be acted upon by the poly A polymerase.

And the poly A site they may be used under different circumstances to generate mature RNA. Now, in a simplistic term we can say a gene have got 3 poly A site. One is say 50 base downstream of the stop codon another may be 100 base pair downstream of the stop codon, another may be 200 base pair downstream of the stop codon.

So, now the question is will it use the 50 base one? Will it use the 100 base one or will it use the 200 base one? The answer is it depends which tissue you are talking about. If the first site is masked somehow like we saw the raw protein binding onto the RNA. Like that many proteins can bind onto this poly A site and mask it.

If it is masked then that will not be recognized by the endonuclease and there will not be any cleavage. So, it will keep it. So, it will keep transcription, the transcription will continue and it will encounter the second poly A site. If that is not masked then that will be recognized and it will be cleaved. What if that also masked? Then it will keep going for the third one.

So, remember if the third one is recognized you have 200 base pair and translated region between the stop codon and the poly A site which is actually worked. Actually it could have had a 50 base pair UTR, but that was skipped and it could have had 100 base pair UTR. That is also skipped. Now, you have got a 200 base pair UTR. Sometimes this extra UTR comes with a burden also.

Sometimes this extra UTR comes with a blessing. Let us see what is the burden, what is the blessing. Burden is if this UTR contains some of the micro RNA recognition element. Like I told you we will discuss more in detail about micro RNA, but remember micro RNA and si RNA normally find their target. They are of around 22 base in length. si RNA or micro RNA they are around 22 base in length. They can if they find a target on some mRNA's untranslated region then it can cause the degradation.

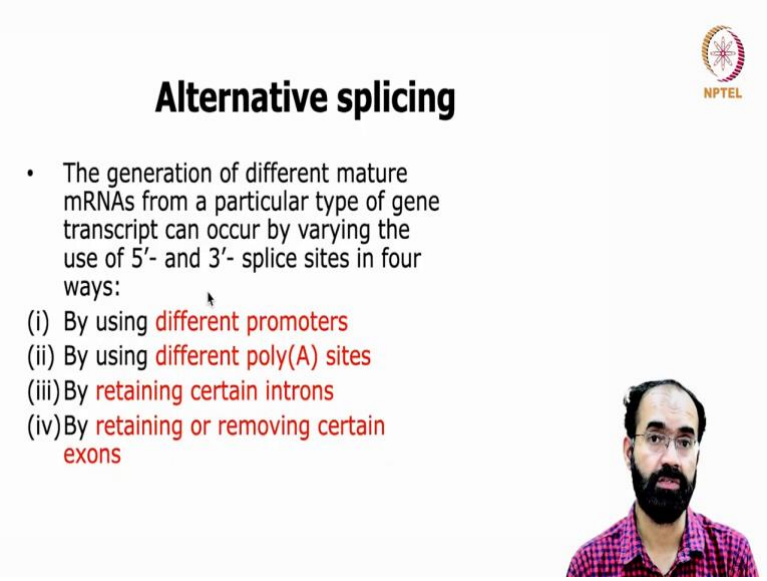
So, basically it is a red flag for the mRNA's survival because extra UTR is there, extra UTR brought in extra nucleotide sequence and extra nucleotide sequence may or may not have target for a si RNA or micro RNA. This is one way. Another way advantage can be

this UTR, extra UTR may provide some special secondary structure which will provide protection from the degrading molecules, degrading enzymes.

So, this will be beneficial. So, we cannot say long UTR is bad or we cannot say long UTR is good. The simplistic way we can say is that it depends which RNA you are talking about, which gene or which tissue you are talking about. So, the system evolves in such a way that with minimum with less you can do more. So, that is the whole agenda or the whole principle lying behind alternative poly A site.

In one cell the stronger poly A site is used by default, but in another cell a factor may prevent the stronger side from being used, like I already told you. One side, two side, three side, four side, there can be multiple sides are possible.

(Refer Slide Time: 13:04)



Now, alternative splicing simply means the generation of different mature mRNAs from a particular type of gene transcript that can occur by varying the use of 5 prime and 3 prime sites in four different ways. So, I told you that you can include an exon into intron or you can include an intron into exon partially to enhance the complexity and it can also skip sometimes one whole exon can be skipped or sometimes one whole intron can be skipped. We will see examples how it is done.

One way is using different promoters. Promoters we know that they are regulatory elements which tells the system. System here what I mean is, the transcriptional

machinery like polymerase and other factors that where onwards the gene is situated from where onwards the transcription should start. So, this is what the promoter or regulatory element. Other than that promoter or regulatory elements means nothing. It is basically a guidance like people say you can see Everest or the Great Wall of China from space.

So, which basically indicates a human construction if it is a Great Wall of China its only an indicator. Like that the promoter elements also an indicator. It can be anything. Who knows maybe another larger huge big structure also will be seen. That is not the only big structure seen from space. There can be other huge structures that also can be seen from space.

So, we should understand this indicators can vary. That means, for simple example for a given gene maybe from the start point transcription start point say some 50 base upstream or maybe 100 base upstream is good enough to act as a promoter. For another gene it may be maybe 10,000 base upstream has got a recognition sequence. It depends on who is the recognition factor or which protein is the recognition factor and what gene you are talking about.

So, we cannot generalize when you talk about individual genes we cannot generalize every gene should have 1 KB promoter or every gene should have 10 KB promoter we cannot say. That is why when you are trying to express a transgenic line when you want to clone the promoter of a gene nobody knows exactly how long the promoter is. Best is more is always better. It is safer.

Like you may if any of you are following this by the Epic of India that is the Ramayana there is a story where the Lord Hanuman brings the whole mountain instead of bringing the medicine for which it has been sent for. So, if you have the whole mountain one can always climb happily and search for the actual medicine which you have instead of searching a medicine and bringing a wrong one.

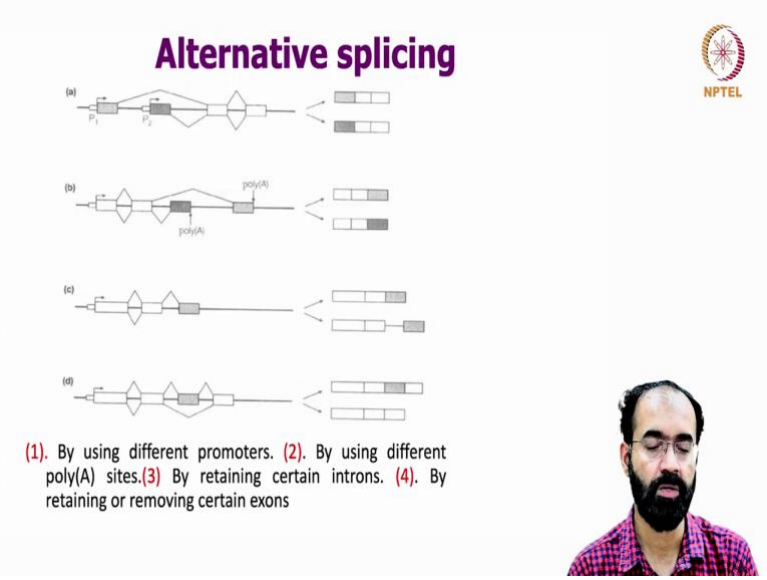
Like that if you have a big clone big chunk of the regulatory element, there is always a chance that you have included all the necessary element. Sometimes it may fall apart you may go into say 20 KB region, 20,000 base pairs. Sometimes you go for 50,000 base pair. For some genes it will go to 1 million. 1 million bases a lot 1 megabits.

So, that can be tricky. But thing is it is rare. It is not that common usually as a thumb rule you can say every gene have got its promoter or regulatory elements within some first 10 KB or so. 10 KB upstream of the transcription start point it has got all the necessary element. However, system make use of different promoter. Say a gene is there.

We are talking about gene A. So, gene A is needed in the skin of an individual. Genes A is needed in the say thymus of an individual and gene A is needed in the brain of an individual Now, how will you regulate these three genes in these three different issues. So, it make use of alternative use of the promoters usually takes place. We will see them more in detail.

Another way is by using different poly A sites which we kind of already discussed. It can use a first poly A site, second poly site or third poly site. And then third by retaining certain introns. That means, it is including certain introns as part of the coding region. And then retaining or removing certain exons. So, these are all the way in which it adds up to the complexity.

(Refer Slide Time: 17:47)



Let us see how is it done. See you can see in the first version that is A that is utilization of differential promoter. See you can see here P1 that is the promoter 1 and then P2 is promoter 2. So, now you can see this arrow indicates the transcription direction, direction of transcription. And second one also is directional transcription.

Now, let us see how is it happening. Say in a given tissue, let us talk about skin. You have promoter 1 and it is transcribing here. And this what you see in the box is ash colour 1 is the exon, first exon. Then it has got second, third, fourth like that multiple exons are there. Say if this promoter is used, it make use of this entire stretch as a single intron.

So, for this promoter, this is first exon and this is the second exon. Not this one. Second exon, this is the third exon. So, when this splicing happens, you end up getting this less coloured first exon and second and third is the same. Whereas, if the second promoter is utilized in another tissue, which was part of the intron, intron of the first promoter and first gene.

Now, the intron of this downstream of this first exon, you have got the promoter 2 and a part of the intron is now acting as the exon and the exon 2 and exon 3 are same as the first one. So, you end up getting this strong stained box and you have got these two exon 2 and 3 the same. So, you end up getting one totally different RNA and totally different another RNA.

Remember the second exon and third exon are same. But first exon is completely different, it can have totally different functionality of the protein. Then you can see another example use of differential poly A site. How is it done? Here only one promoter is there exon 1, exon 2, exon 3. So, what happens? They splice together exon 1, 2 and 3 normally happen because the poly A site is made use of and you end up getting one product. Remember exon 1 and exon 2 are identical exon 3 is unique.

Now, same promoter is driving the transcription, but it did not use this poly A site, it skipped it, somehow missed it, some protein masked it and then the transcription is continuing and it encountered another exon and then downstream of that it has got another poly A site.

Now, the question is first exon, second exon and third becomes fully part of an intron and whereas, this third one is third one of the first gene is now become part of an intron whereas, the fourth one identical fourth one is acting as a third exon for this gene as you can see here.

So, you end up getting two identical first exon and second exon and the third exon is unique. In one gene third exon is actual third exon, in another gene fourth exon is the third exon. Fourth exon is coming in the place of third exon because the actual third exon is included in the intron.

So, in the first case alternative use of the promoter you had diversity in the beginning part of the RNA and which will in turn bring a change in the beginning part that is the N terminal of a protein. Whereas, in the second case you will have variation in the end part of the mRNA, hence you will have variation in the C terminal of part of a protein. C terminus of a protein will be different.

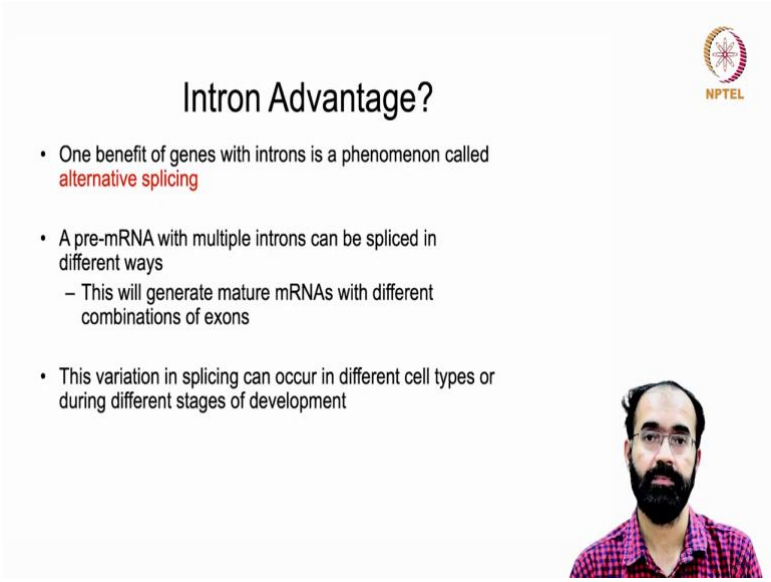
Now, the third example, what is third example? It is making use of differential utilization of you saw here retaining certain introns. So, now you see here promoter is mentioned here you have exon 1, exon 2, exon 3 and it is giving rise to one product here. Now, as a variation what you end up getting you are getting exon 1, exon 2 and it has now retained a portion of the intron and then you have got the exon 3.

So, what happened? You got a longer exon 3 and this will add up portions of an intron. So, you can see here in this case this splice to here, this splice to here and you got this product. In the second case this splice to here and this did not immediately splice, but it included a portion of this intron which was completely removed in the previous case. So, you end up getting a extra sequence from the intron.

And then what was the other one? Retaining or removing certain exons. So, that is what you are seeing in the last example. That is you have the promoter you have the first exon you can see here top first exon, second exon, third exon and fourth exon. Everything happens smooth. Not this ash coloured box because that has got some importance that is skipped in the second case.

First exon, second exon it did not go into the third exon because the third exon is included in the intron. So, you have first intron, second intron, third exon and third intron is now acting as a one giant intron. So, you end up getting first exon, second exon and fourth exon, third exon is skipped. So, this is the way it will bring in extra diversity.

(Refer Slide Time: 23:44)



So, let us quickly think that, what is the benefit of having intron? From now we know that introns are important because introns have got plenty of you know roles to play. You know one more thing I would like to tell you if you are making a transgenic animal say you are making a transgenic mice.

So, every gene have got you know exons and introns. So, normally if you are making a gene from the cDNA, it is already spliced like you make RNA, RNA reverse transcribe it, you make the cDNA, this is a protocol in the molecular biology. So, the advantage of making cloning a gene from the cDNA you are devoid of introns.

So, it is full gene, its ready to go. This if you use for making the transgene then the expressivity of this transgene in the embryos is highly compromised. On the other hand, if you include the introns. So, you cloned a gene with its full introns and you made a transgenic animal, expressivity of that transgene is 100 fold higher 100 fold higher than using just a cDNA alone.

So, what it indicates, introns contain elements that can accelerate the expressivity of a trans gene. So, naturally it will influence the expressivity of that gene when it is intact also. So, one benefit of genes with introns is a phenomenon called alternative splicing. If you do not have introns, you do not have variety. It is just like if you have lots of money you can spend nicely. But all you have is 10 rupee. So, you are limited, maximum you can have is one Tea. You cannot think of you know going for a tour, vacation etcetera.

So, introns give this luxury of having introduced diversity. A pre mRNA with multiple introns can be spliced in different ways. If it has only one intron you are limited with just one way of splicing. Otherwise, you cannot. If you have a ten intron it is a huge variety is possible. This will generate mature mRNAs with different combinations of exon.

As long as you have different introns it gives you opportunity to explore different ways. This variation in splicing can occur in different cell types or during different stages of development. Because in a given cell if a gene is expressed it will not keep you know playing around. It has a it has a routine way. It is just like you have a defined way of making chapati using atta. You will do only that way.

On the other hand, if you are making chapati with besan or puri with besan you will not follow the same procedure. It may have slightly or if you are making paratha. You will make a different procedure. Same logic applies when you have a different cell type then it will employ this alternative splicing approach.

If you have say liver cell one particular liver tissue gene A, it will not do all possible combination of alternative splicing. It will follow stick with one. But if the same gene is expressed in kidney or same gene is expressed in the intestine it may follow another approach. But the organism have got only one gene. So, this logic is what you should be keeping in mind. We will continue more in detail about this in the next class.

Thank you.