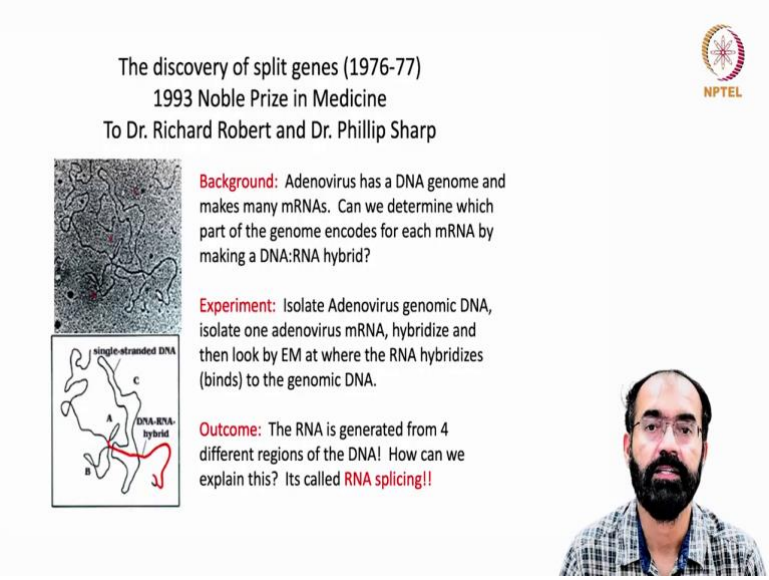
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

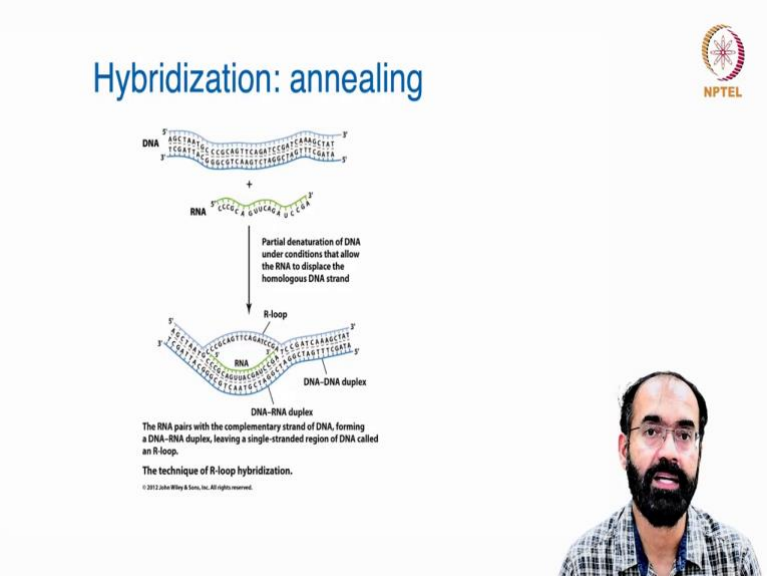
Lecture - 21
Alternative RNA Processing and Editing: RNA Editing in Detail

(Refer Slide Time: 00:21)



Hello everyone, welcome back to another session of RNA biology and we were studying the discovery of introns by Richard Robert and Phillip Sharp in as old as 1976-77 in the year those year and they ended up getting Nobel Prize in 1993. And let us see how this R-loop hybridization takes place.

(Refer Slide Time: 00:52)

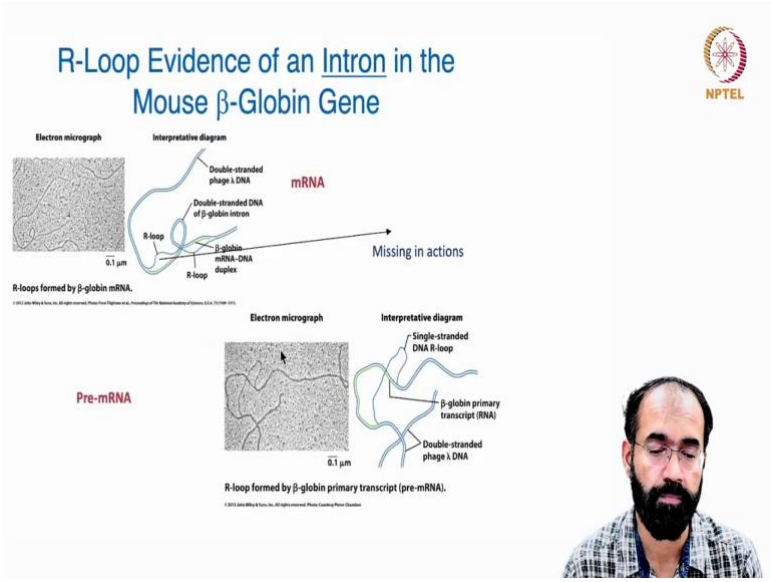


So, hybridization is nothing, but an annealing and DNA usually stays double stranded and the RNA is coming from or produced from this DNA strand. So, the assumption is it is an exact replica of the DNA, lengthwise as well as sequence is only difference is uracil comes in the place of thymine. Is it always the case? And the answer is no.

In some cases you will have variation so, partial denaturation of the DNA under conditions that allow the RNA to displace the homologous DNA strand. So, what it happens like you have an RNA and it finds a place where it can bind and the unwanted region will loop out. This is normal thing because once you denature the DNA, if RNA is bound there DNA count bind and it is kind of identical. The normally RNA DNA duplex when it is formed the DNA strand equivalent span will be moved out.

It is almost like a bracket. If this RNA is 100 base long and 100 base DNA has looped out. So, that is a situation normally should exist or that is a situation where introns are not there that should exist. But is it so always? The answer is no.

(Refer Slide Time: 02:19)

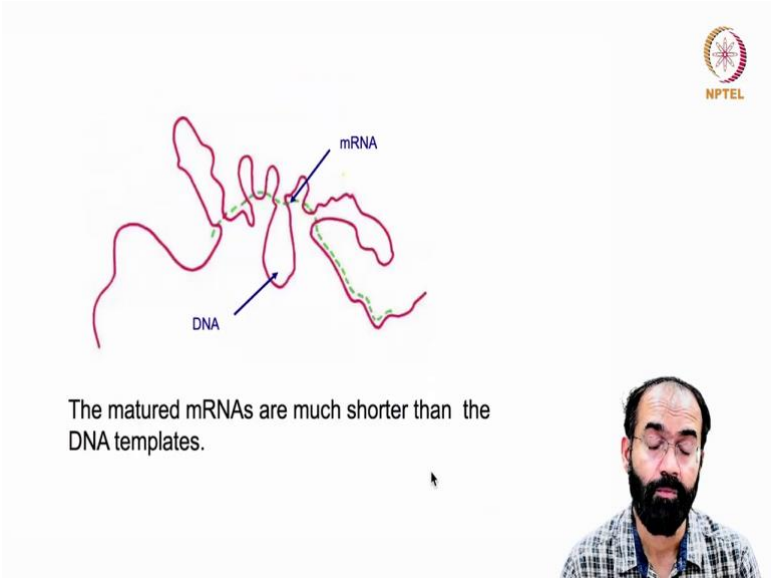


R-Loop hybridization evidence for an intron in the mouse beta globin gene, how does it work? We can see it. This is the real electron micrograph picture what you are seeing in the left hand top side and then you have a representative diagram here. You have a double stranded DNA and also you have a R-loop that has occurred because you have hybridized with certain probe or certain RNA molecule.

And of course, radio labelling etcetera are there because you need to have recognizability onto this probe what you are putting in. So, this is a electron micrograph and this is another electron micrograph. So, if you are using pre-mRNA as the probe which is containing the intron you will have identical size of the loop that is whenever the transcription happens transcription is to the pre-mRNA.

Pre mRNA has got introns and exons intact. So, the hybridization will be equivalent. So, you will have one strand of the DNA that is projecting out and now RNA has hybridized on to that.

(Refer Slide Time: 03:45)



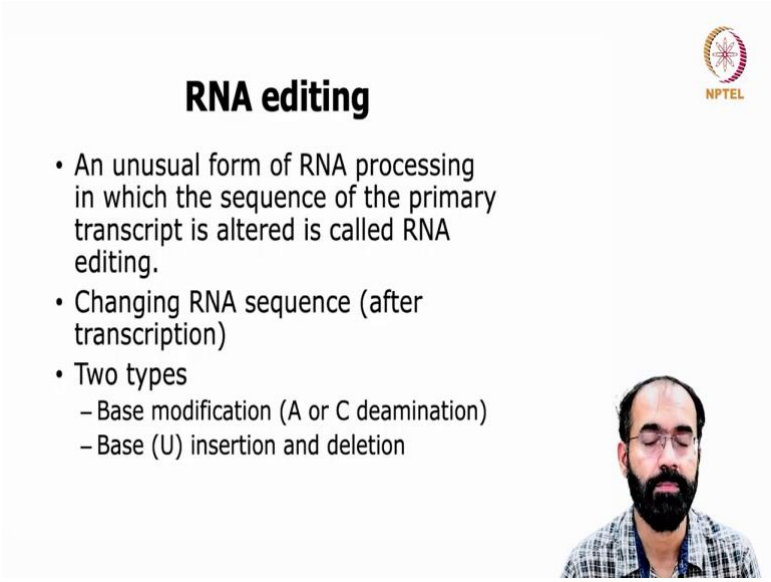
But what happens if you are using matured mRNA that is not pre-mRNA it is a spliced mRNA where introns are removed are much shorter than the DNA template. So, what happens if you are using a mature mRNA you will see this green colour is the mature mRNA it is pairing and then a huge chunk of DNA which is pink in colour has moved out and then then there is a pairing with the DNA then there is a chunk of DNA coming out and then there is a big chunk not pairing.

So, what you can see here and there little bit little bit smaller and smaller fragments of the RNA is pairing with the DNA. RNA is pairing throughout, but lot and lot of fragments of the DNA that is projecting out which is much larger than the RNA size itself.

Usually the RNA size and the displaced DNA strand should match out the loop should match out, but here matching does not happen it is disproportionate for a RNA of smaller in size the displacement is huge size DNA fragments which is nothing, but the intron. So, this is how the discovery of introns are proven in a very elegant, but very simple experiment called R-loop hybridization.

So, the concept and the principle should be very clear to you because the technique that has used that has been used to identify R-loop or presence of intron is very powerful and it has been extensively used for variety of other experiments to discover unknown introns and to discover alternative splicing etcetera etcetera.

(Refer Slide Time: 05:31)



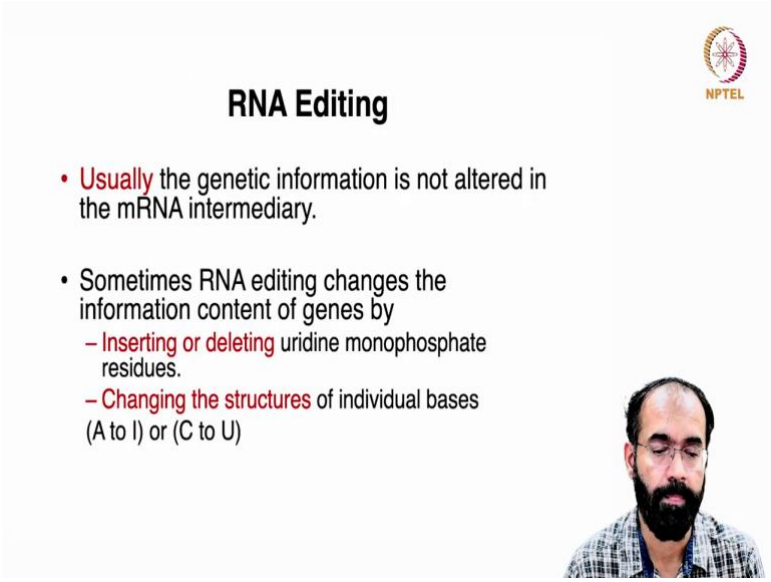
So, and also remember that was a time when the genome sequencing wasn't there. Now, it is rather easy you take the RNA, you sequence the RNA and then also compare it with the genome, we have the genome of pretty much a large number of model organisms, plenty of animals genome size is identified. So, that is not a news anymore, but that time it was a big news and it was not that easy to comprehend.

So, now coming back to RNA editing it is an unusual form of RNA processing in which the sequence of the primary transcript is altered and is called RNA editing. So, whenever there is a change occurs at specific base we call it as RNA editing and RNA editing can cause a do or die situation like in some case the animal will not survive if the RNA is not edited it is that important, we will see with some examples.

So, RNA editing is basically in the simplistic term with the changing the RNA sequence after the transcription or even after the RNA is matured, but it is not sent for translation it is now waiting for the editing. There are two types of editing are there one is base modification A or C the deamination and another is U insertion or deletion.

So, A is a Adenine, C is for cytosine and they can undergo deamination even causing another base or changing into another base and then comes the uridine the uracil element or uracil nucleotide can be inserted or it can be removed. So, these are all the simplistic form two types of RNA editing.

(Refer Slide Time: 07:34)

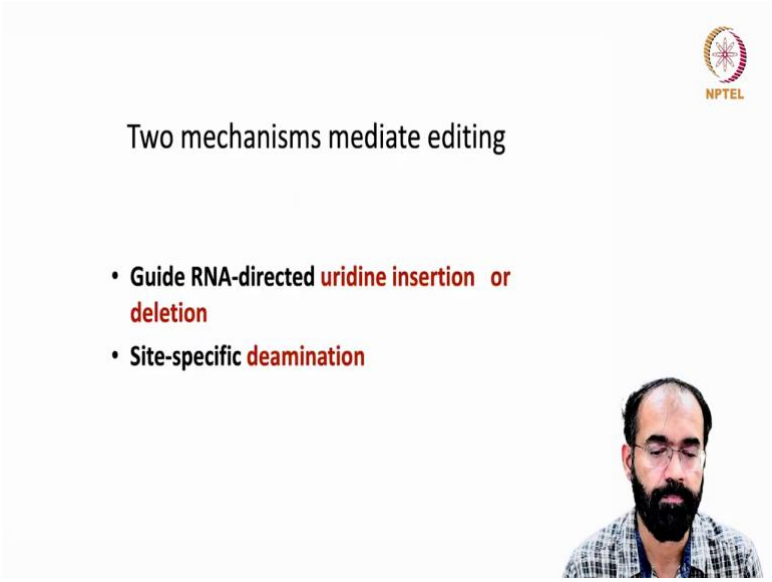


Let us see more in detail usually the genetic information is not altered in the mRNA intermediary. So, we should understand that the DNA or the encoded gene is unaffected encoded gene is intact and it is holding perfect and it is not altered in any way, but it is the post transcribed RNA that is undergoing the changes.

Sometimes RNA editing changes the information content of the genes by inserting or deleting uridine monophosphate residues other way is changing the structures of individual bases like we saw the deamination even A to I or C to U. So, A can be converted into inosine I stands for inosine which is a neutral base it does not have affinity like if there is an inosine it can pair with A, G, C or U.

So, it has a neutral affinity and C to U. So, there was a cytosine present that is now converted into uracil. So, this is another change.

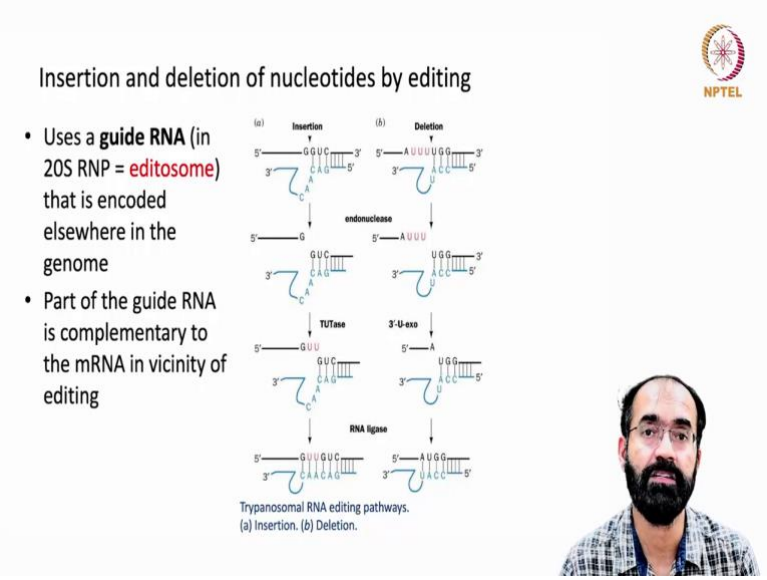
(Refer Slide Time: 08:53)



So, two mechanisms mediate the editing that is one is guide RNA-directed uridine insertion or deletion one way of handling. So, editing like we saw existing number of bases are unaffected. Now, we are talking about a number of bases in an RNA can be affected. So, guide RNA are group of RNA that direct certain protein or enzymes to specific loci either in the genome or into another RNA molecule. And this uridine bases can be selectively added or removed.

And then comes the site-specific deamination, not that every A is deaminated, not that every C is deaminated, not like that some specific loci specific location because of the secondary structure of that RNA will be detected and it will undergo change.

(Refer Slide Time: 09:52)



.

So, insertion and deletion of nucleotides by editing how is it happening and we saw uridine can participate in the insertion or deletion, how is it happening, let us have a quick look. And the guide RNA in the 20S RNP, RNP stands for you know that ribonucleoprotein and 20S stands for a svedberg unit of RNA which is 20S in size and we call it as a editosome.

Editosome means a guide RNA that is capable of influencing uridine changes that is encoded elsewhere in the genome. So, it is produced from somewhere and it is acting on some other RNA. So, it is part of the guide RNA that is complementary to the mRNA in question, the mRNA that is to be edited in the vicinity of editing. So, how does it happen this pictorial manner you can easily know the first section A is basically dealing with insertion second portion is dealing with deletion let us see how is it done.

So, you have an mRNA which is present which has got GGUC sequence ok and this GGUC eventually become GUUGUC. So, you have insertion of two bases. So, what it does. So, this editosome RNP comes in and pairs at specific location and not that it is pairing only in three bases of the real major region and then it is pairing with another four more bases downstream which is immaterial and it stays paired.

Then this editosome can create a nick between these 2 G's it creates a nick and it separates this strand. And remember that strand is not lost it is there in the vicinity. Now, this once this endonuclease action is done what happened, there is a TUTase enzyme terminal uridine transferase that is the name of the enzyme.

So, endonuclease comes and cuts. So, why endonuclease come because this editosome once it is paired with this mRNA can be attractive or can be recognized as a substrate by the endonuclease that is why it is cutting here. Now, this cut portion will become a substrate or this cut nucleic acid becomes a substrate for TUTase terminal UTP transferase which adds 2 U's only just to 2 U not many U, G and G was there, it was cut in between and in between you have 2 U's.

So, it became G G became G U U G and, but that is quite interesting because this unpaired portion is now finding it attractive to pair, earlier it was not pairing because you added 2 U's you end up getting because otherwise only up to here this G it was pairing, here it was not pairing, but addition of 2 U made to pair with these 2 A's and you end up getting G C U A U A G C U A U G pairing and remember G C U A U G pairing was already there, but this extra 2 U happened and this will bring in an RNA ligase.

So, this piece now comes in place because otherwise this G U it is a separate strand and that comes in place and fuses it. So, there will be ligation between this G U U and this U and this G there will be ligation; that means, here right here there will be a ligation by RNA ligase, this you are talking about insertion.

Now, let us see how deletion is possible, can deletion be possible similarly? Same way, but RNP editosome comes in here it catch hold of a specific region in an mRNA and you have a sequence A U U U and U G G. We are talking about this 3 red colour U's will be lost in the end.

So, A U U U U G G became A U G G; that means, it lost these 3 U's how is it possible of course, this pairing brings in a endo nuclease it creates a cut between the last, but U and the previous U it creates a cut endonuclease action and then what happens it removes and this is done by an enzyme that is 3 prime U-exo.

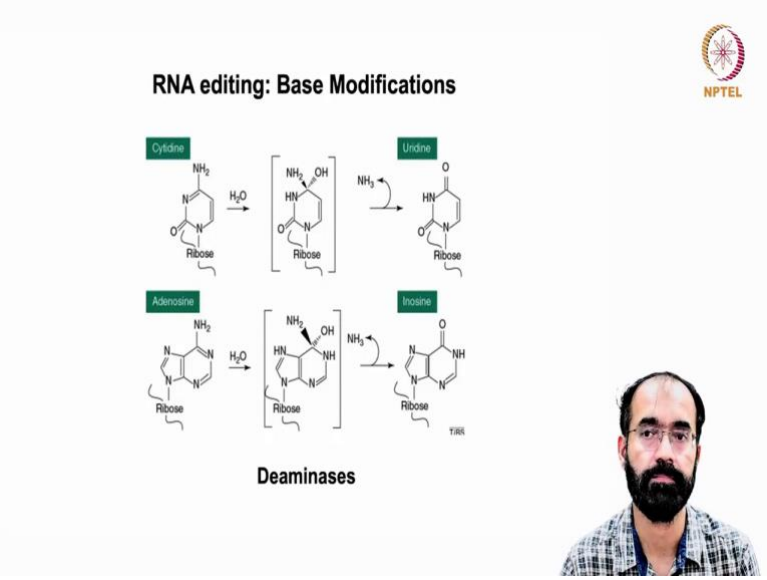
So, this is an exonuclease this is a RNA degrading enzyme, but it does not degrade any RNA it degrades only this conformation RNA and it targets only the Us and it removes

this 3 U and you leave this released a fragment with an A standing here and now this A and this U G G that was there already there in the previous pairing and this comes in contact and it creates a continuous pairing A U G G earlier it was U G G only this A was not participating these 3 U's were not participating in pairing of course, there is downstream pairing here, but that is not participating in any of this modification reaction.

So, you have this removal of this 3 U it brought in a continuous pairing A U G G paired with this RNP molecule 20 S RNP molecule and which brings in a RNA ligase. So, that it will create a phosphodiester bond formation between this A and the U. So, that it becomes a edited RNA. So, earlier this RNA it had a G G you see now it end up becoming G U U G UC, here it was A U U U U G G it became A U G G.

So, in this case you added 2 U's in this case you removed 3 U's, remember these are not the same region these are not the same RNA and this is not even the same RNP, but this is the way you can create changes, this is the way you can bring in edited versions. So, this is the mechanism of addition or deletion of U's.

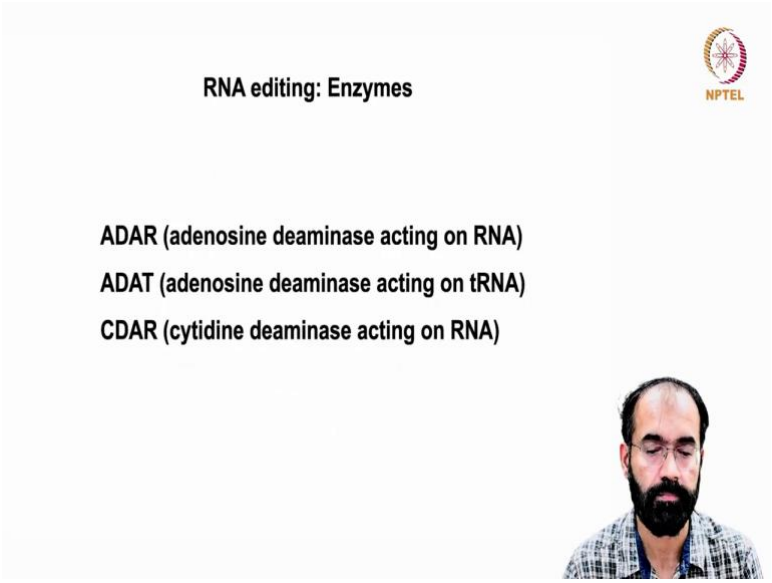
(Refer Slide Time: 16:28)



So, RNA editing can happen through base modifications also. So, cytidine or cytosine residue can undergo some process of deamination to give rise to uridine and adenosine can undergo a process called deamination and to become inosine. So, these are done by deaminases.

So, again I am specifying that not that any random cytidine deaminase look certain specific secondary structure onto the template mRNA and modify only those cytidine modify only those adenosine which are attractive substrate for this enzyme. So, these deaminases are looking for specific secondary structure.

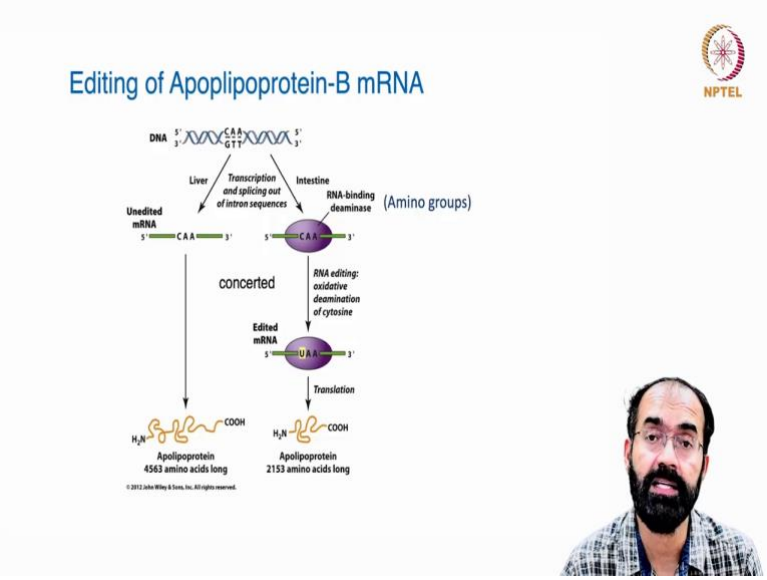
(Refer Slide Time: 17:26)



Now, let us see RNA editing enzymes, what are the RNA editing enzymes? We can have a acronym, ADAR, what is ADAR? Adenosine deaminase acting on RNA. What is ADAT? Adenosine deaminase acting on tRNA. CDAR cytidine deaminase acting on RNA. So, these are some short form.

So, what it indicates ADAR can act on RNA whereas, ADAT can act on tRNA not acting on other RNA and CDAR can act on any RNA. So, they are also while using these enzymes they are also it can bring in specificity.

(Refer Slide Time: 18:18)



Let us see some real time example of this RNA editing, what is the significance, what is the importance. This example is editing of apolipoprotein B mRNA and we can see two different tissues one is liver and another is intestine. So, the gene in question is apolipoprotein B mRNA. So, in liver it is unedited, as you can see here the RNA and this is the portion which you are discussing CAA which is paired with GTT in the case of DNA.

So, in the RNA you are supposed to have CAA and it give rise to a full-fledged protein in the liver which is containing 4,563 amino acids long this is the apolipoprotein B which functions in the liver whereas, in the intestine the cases are different and it is recognized by RNA binding deaminase it can bring in RNA editing and what it does, it causes the C to deaminate.

So, the CAA, C of the CAA is deaminated and it become UAA. UAA we know you may remember it is one of the stop codons 3 stop codons are there UAA UAG UGA. So, this was a codon CAA now it became UAA one base per chain. So, the codon became stop codon and UAA presence will terminate the apolipoprotein.

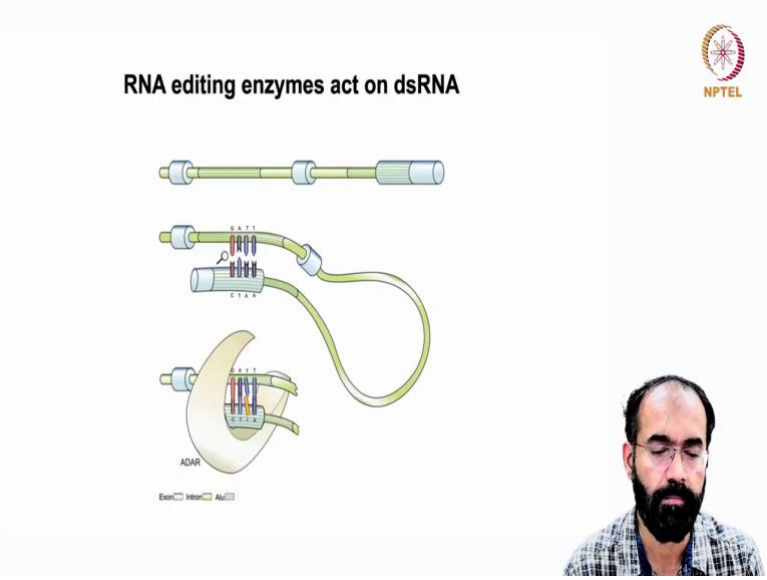
So, actually it should be 4,563 amino acid long if CAA codon was intact, but because of this RNA editing C became U you have an apolipoprotein which is 2,153 amino acids long, which can function in the intestine. Do not think this is by accident because both

tissues need apolipoprotein one say liver needed a larger version whereas, in intestine needed a shorter version because they are totally different tissues.

So, RNA editing becomes handy otherwise system should have had a two different genes and which is extra load onto the genome and you have to have two transcription, machinery, etcetera. So, here effectively having one RNA editing enzyme can bring in the change through same mRNA, but as long as this enzyme is available in the intestinal cell it will never produce this 4,563 amino acid long apolipoprotein B.

It will always produce 2,153 amino acid long apolipoprotein B which is ok and sufficient and necessary for intestine point of view.

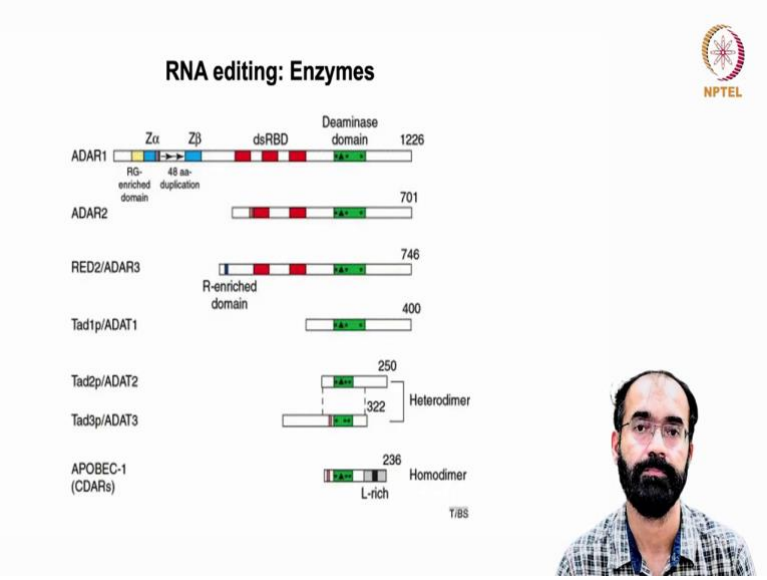
(Refer Slide Time: 21:16)



So, RNA editing enzyme can act on double stranded RNA. So, double stranded RNA basically means secondary structure. Do not think that it is extensively double stranded like DNA double helix it does not happen like that. So, RNA enzymes always look for or this modifying enzyme editing enzymes look for specific double stranded RNA part, as you can see here and that will make it attractive candidate for this an example a cartoon of ADAR and what you see in blue colour boxes is the exon and the introns are green in colour and other small box is basically Alu repeat which is a feature of the genome.

But this pairing region double stranded region makes ADAR a proper enzyme that can recognize this substrate that is double stranded RNA.

(Refer Slide Time: 22:14)



RNA editing enzymes are quite interesting, RNA editing enzymes are having several unique common features and they have their own unique features and they also have several common features and what are they this is ADAR 1, ADAR 2 and RED 2, ADAR 3 and name of these genes are not important.

Let us see with some example, some enzymes have got very large molecular weight in size and they also have specific region double stranded RNA binding domain and this is the Deaminase domain. Deaminase the actual catalytic role played by this enzyme. So, this box is nothing, but a cartoonish way of explaining the structure of this protein enzyme these are protein enzymes ok RNA modifying or RNA editing enzymes and this is a large protein it has got 1,226 amino acid this is got only 701 like so on and so forth.

You have all of them invariably have got a RNA binding domain and Deaminase domain is common and some of them do not have the RNA binding domain, but they will take the help of other collaborating protein which will have the RNA binding domain. Not that every protein need to have in its own way, like you buy rice from the market you need not be a farmer or you buy milk from the shop you are not a farmer right. So, you pay money and you get it.

Proteins also do that they do not pay money, but they will find a suitable partner suitable interacting molecule which have an RNA binding capability and the Deaminase action of

these enzymes become handy for them to interact. So, this is how they control or regulate the RNA binding capacity.

We should understand these RNA modifying enzymes they have evolved over the years and they randomly could have modified these bases and domains and they could have acquired a novel function they could have acquired just like you know how humans learn cooking, you know that the idea is during wildfire it was very easy for humans to get half cooked or burnt animal flesh and they found it rather than eating raw flesh the burnt flesh is more tasty and then they learnt this fire can be tamed and they can they learnt the art of cooking the old human.

Same way many enzymes modify and some of them do not do anything, some of them can be deleterious even in our body many times many modifications are not having a specific function yet and sometimes some proteins are having very high temperature functioning capacity; that means, normally our body temperature is 37 say a plant grows in normal room temperature.

But why a protein should have stability at 90 degree when the organism never sees 90 degree, you will simply die one classic example is you may have heard about the enzyme papain is an enzyme present in papaya plant and raw papaya has lot of papain and people use this for cooking meat they cut the papaya and put along with the meat.

Because the papain enzyme help in making the meat very soft cooking becomes easy you do not need a pressure cooker to cook the meat you cut the papaya pieces because the papain's enzymatic activity is close to around 100 degree, but which papaya plant is growing at 100 degree Celsius.

Papaya plant will die because this is a feature which is there which is yet to be tested, we are smart and we are using it like that many proteins and many enzymes have got a function which may not be put for testing like that many proteins will edit some RNA. If it is not causing harm to the animal it will still be there it is waiting for the time to acquire a novel function. So, this is the concept you should have.

So, we will continue to understand about the RNA editing with some specific example where we have some do or die situation in the coming class.

Thank you.