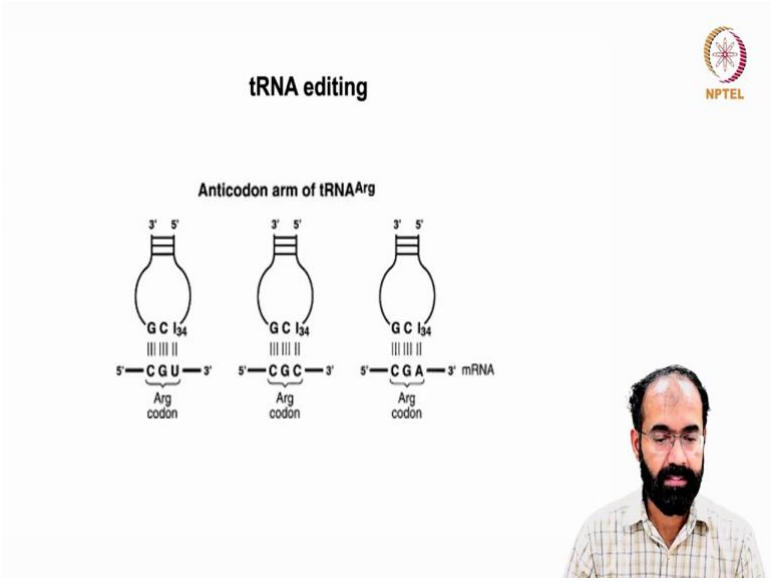
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

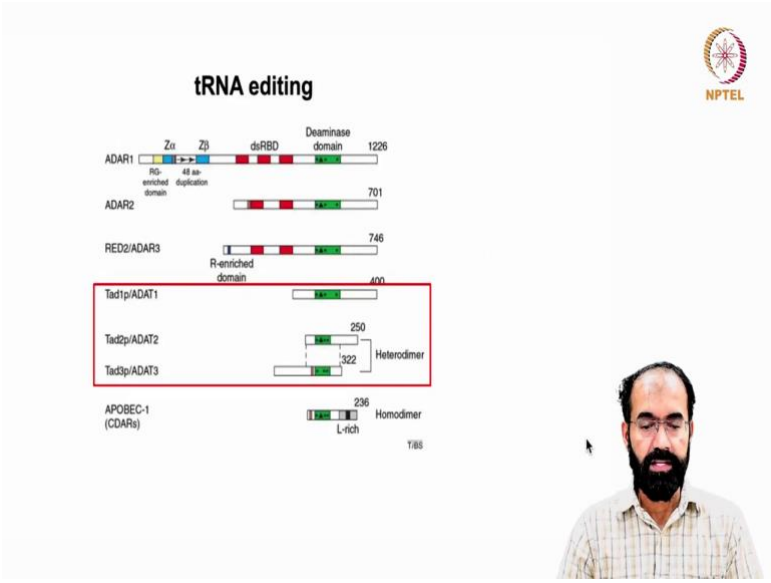
Lecture - 23
Alternative RNA Processing and Editing: Relevance in Immunology

(Refer Slide Time: 00:21)



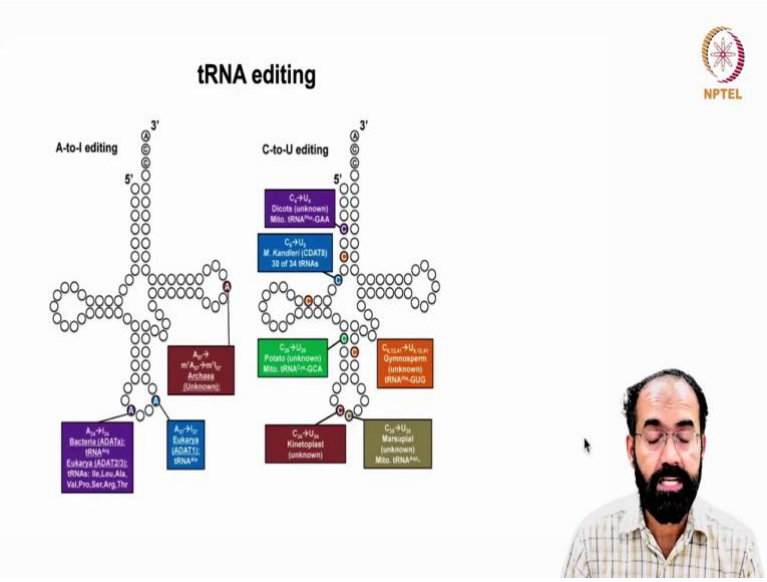
Hello everyone, welcome back to another session of RNA Biology.

(Refer Slide Time: 00:34)



And we were here in the previous class that we were discussing different proteins that are capable of editing the RNA at specific levels. And we know this RNA editing enzymes have a specific loci that is for recognizing the RNA domain and also they have specific domain for causing the deaminase function. And we are now going to see more in detail about the tRNA editing which we kind of looked in the previous class.

(Refer Slide Time: 01:15)

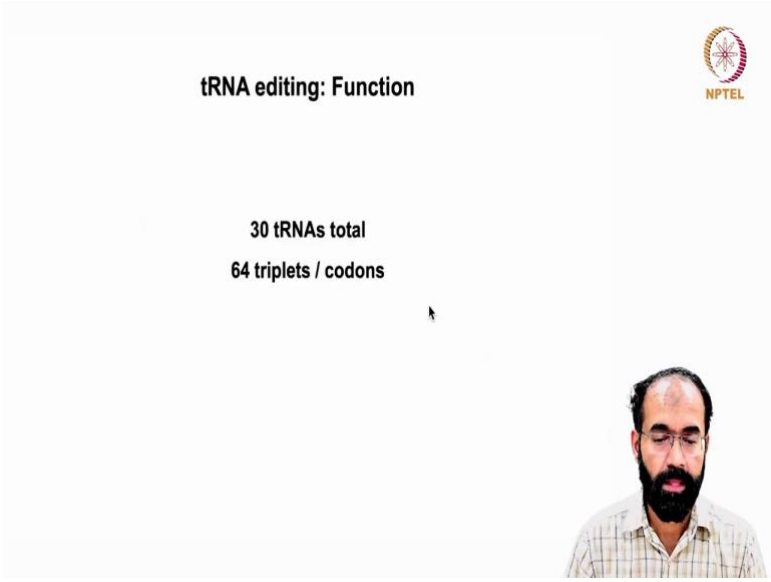


That is specific location in the tRNA need to be edited for purpose of enzyme identification and also for the stability of the tRNA molecule. And we also saw this that the changes in the anticodon part, anticodon part of the tRNA as you can see here allows that with minimum tRNA genes you can cater maximum number of; maximum number of the codons. So, this is an advantage.

So, advantage is quite useful or quite lifesaver when it comes to; when it comes to several prokaryotic or several lower order organisms who do not have the luxury of a wide variety of tRNAs. So, this GCI conversion, GCI conversion is happily be able to cater the arginine codon, because I is neutral, I does not have any biasness towards any of the bases. Otherwise, you must have to have a CGU pairing with GCA only.

So, this problem can be solved and subsequently the other codons also they were able to cater effectively.

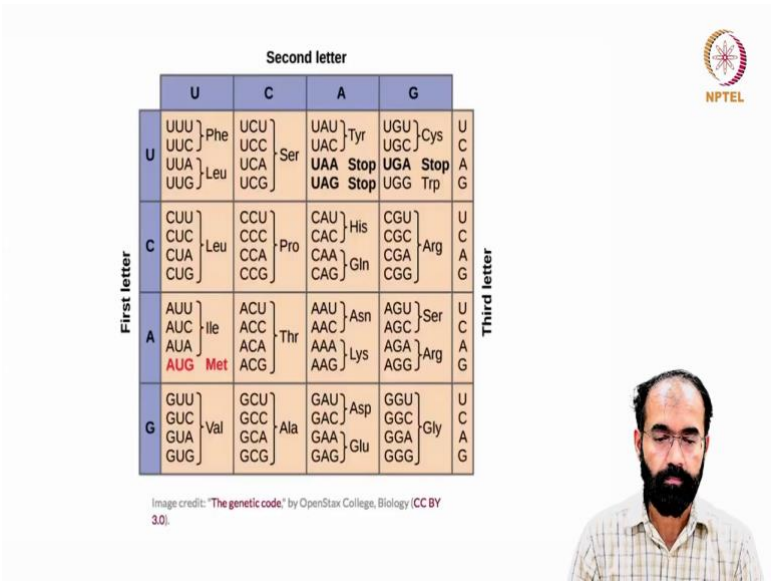
(Refer Slide Time: 02:41)



Now, the outcome is that there are 30 tRNAs are there in total and 64 triplet codons are there, 61 brings in amino acid and 3 of them brings no codons or no amino acids and we call them as stop codons or nonsense codon, which do not bring in any amino acids. So, only 30 tRNAs are there, of course they are sufficient because we have only 20 amino acids.

However, the diversity is useful because you have 61 codons that has to be catered by 30 tRNAs. So, this editing capability allow, this variation to be brought in effectively.

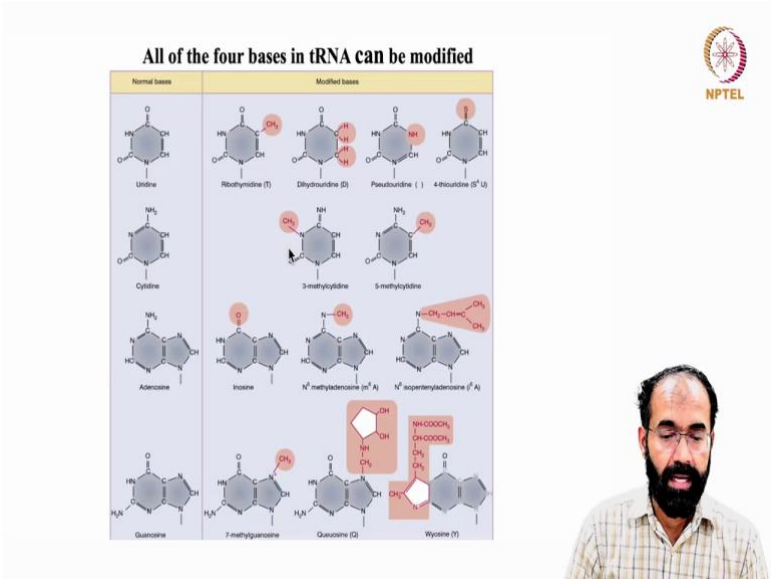
(Refer Slide Time: 03:23)



So, you can see the codon preference table, this is the first base of the codons like U this is written here and this is the second letter U, here it is U and this is the third letter. So, U, U and U it is written U and then comes second U, U, C, U first letter, U second letter and this is C U, U, C and first letter, second letter then A, UUA and UUG.

So, like that it keep varying and accordingly you will have different phenylalanine, leucine like that it keep varying. So, this discovery was done by one of the scientists of Indian origin and his name is Har Gobind Khorana. So, he also received Nobel Prize for the discovery of the genetic code.

(Refer Slide Time: 04:26)

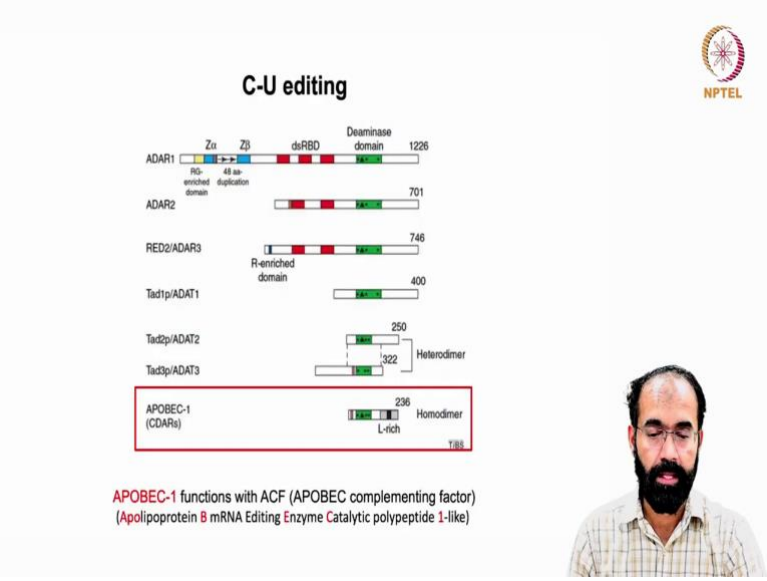


And so, all bases are vulnerable to changes and this slide you may have seen it in the one of the earlier classes also, where we are discussing about the RNA modifications or the base modifications. Uridine, cytidine, adenosine and guanosine all these bases can change into ribothymidine, dihydrouridine, pseudo uridine, 4-thiouridine, etcetera. Like that cytidine also can be changed into 3-methylcytidine, 5-methylcytidine, etcetera.

So, adenosine also can change into inosine, N^6-methyl adenosine and N^6-isopentenyladenosine, etcetera. And guanosine can change into 7-methylguanosine, which we know we all familiar with because it is used for 5 prime cap and then also Queuosine and also wyosine. So, these modifications are in different tRNAs and they can influence the functionality of the or the stability of the RNA molecule.

So, such modifications are not done random, they are done at specific bases and they are done with a purpose. So, we should understand the changes, each and every change that is occurring in individual bases are having one or the other purpose.

(Refer Slide Time: 06:05)

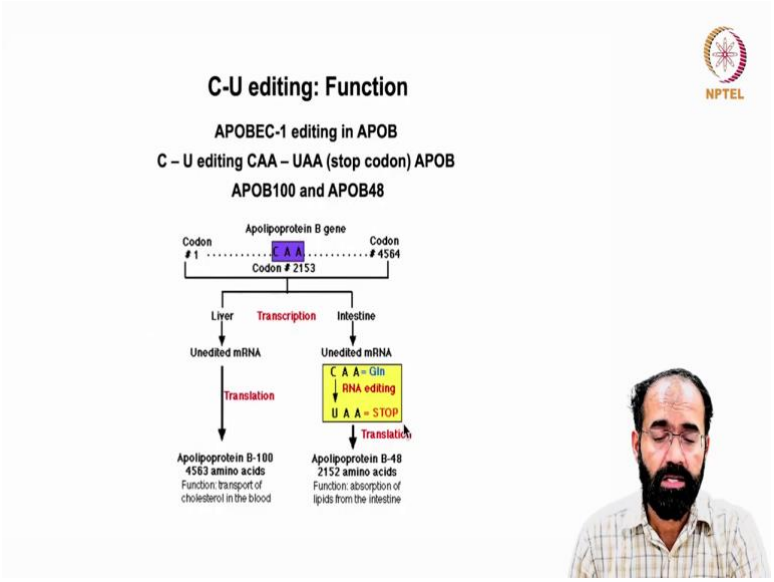


And let us see some example of C-U editing that is cytidine or cytosine to uridine. C-U editing and we know plenty of enzymes we have seen so far, it is the same slide repeating and now we are going to deal with a RNA editing enzyme called APOBEC. APOBEC is a well studied molecule, it has got huge implications is important in RNA editing and it also helps in a lot of other unrelated functions also in the sense it allows it is made use of by various viruses for its own benefit, etcetera.

So, we will see them with some example in detail. So, what is APOBEC? Apolipoprotein B mRNA Editing Enzyme Catalytic Polypeptide 1-like, that is the full form of APOBEC. And some portions have been taken that is highlighted in red color APOBEC and 1-like is given this number. So, this functions with ACF, what is ACF? APOBEC complementing factor.

So, APOBEC is an RNA editing like earlier also I have told you that some molecule, some RNA editing molecule will not have RNA recognition domain, they have the deaminase domain. So, they will collaborate, they will always get along with another protein, which has the RNA recognition capacity APOBEC is 1 of them. So, you see the example ACF here.

(Refer Slide Time: 07:45)



How does it function? So, what is the significance of C-U editing? APOBEC-1 editing in APOB, let us see how is it happening? C-U editing is one what is happening and where which coordinate is editing? CAA is now changing into UAA. And we know we have seen another example earlier where the liver and intestine was functioning a variation in the apolipoprotein B and here you are seeing another example is CAA become UAA.

UAA is a stop codon, you know there are three stop codons are there, UAA is one of them. And what happen? This APOB can be of two different molecular weight, because of introduction of the stop codon in certain cell type. So, APOB 100 and APOB 48 let us see this is the structure of the apolipoprotein B.

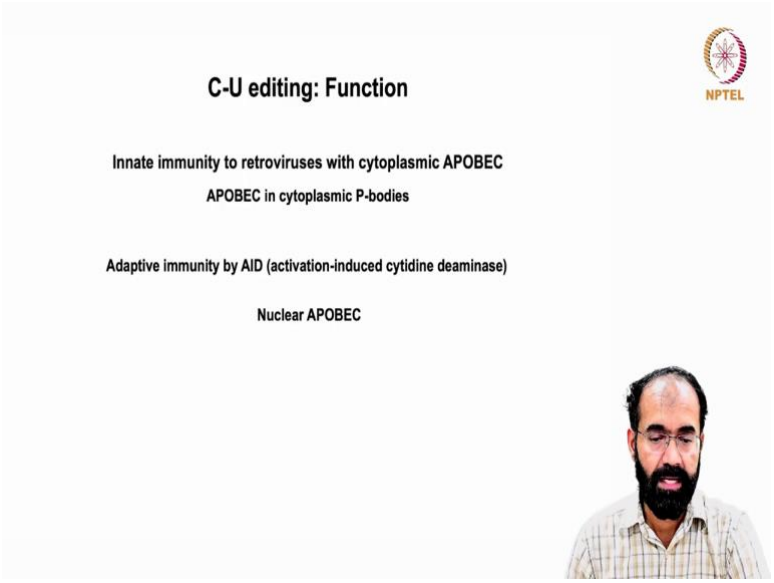
So, codon 1 to codon 4564 and you have a stop codon introduced here 2153. So, now usually in liver and intestine this is how they work and in liver it is unedited and you end up getting a full-fledged protein apolipoprotein B and which has got 4563 amino acid. And its functions in the transport of cholesterol in the blood that is very important function in the liver.

Whereas in intestine you have an unedited RNA usually produced which undergoes editing with the help of APOBEC and this will introduce a stop codon and this stop codon will stop the translation. So, the mRNA will terminate its translation prematurely and this apolipoprotein B is apolipoprotein B-48 which has got around 2152 amino acids.

And its function is totally different, function in the absorption of lipids from the intestine which is quite handy because intestine is a place where the nutrition food, what you take inside should be absorbed effectively. So, in both the tissues liver and intestine the derivatives one is a full length another is a truncated version both have got an effective function.

So, now you see the example RNA editing is not only able to bring in 2 different proteins, of course they have common 5 prime regions or the N-terminus end of the protein they have common. However, what is important is with one gene it is able to perform the task of 2 genes or 2 proteins in 2 different tissues. So, it is very advantageous for the organism.

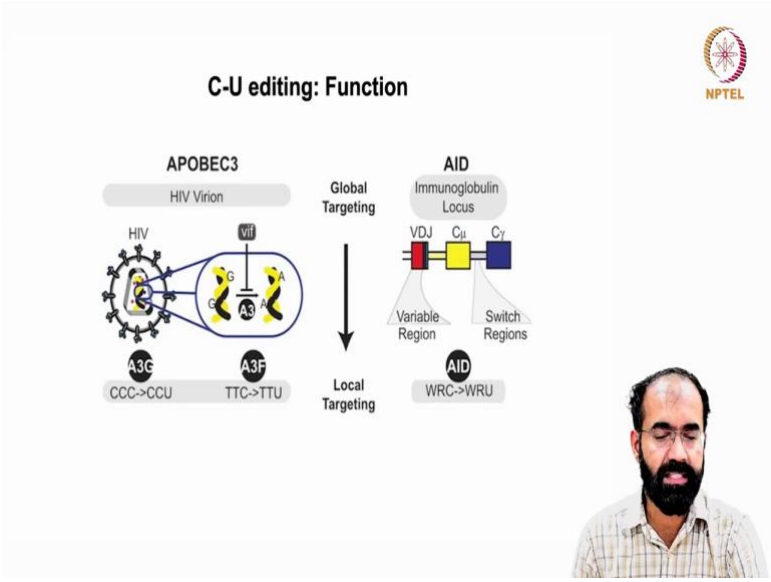
(Refer Slide Time: 10:39)



Now, let us see another angle of functioning of the C-U editing. So, the innate immunity to retroviruses with cytoplasmic APOBEC. APOBEC can be nuclear as well as cytoplasmic. So, RNA can be edited in both the places depending upon which RNA you are talking about. And APOBEC is associated with cytoplasmic P-bodies and the adaptive immunity by AID what is AID activation-induced cytidine deaminase.

So, cytidine Deaminase function, because cytidine if it is deaminated it will become uridine This function is adaptive immunity related. So, this adaptive immunity by AID is something which is very important for us to go further and of course, there is a nuclear APOBEC also we are dealing with cytoplasmic APOBEC.

(Refer Slide Time: 11:40)



Let us see what is their relevance? So, APOBEC 3 and AID they interact together because we have already seen here that it functions along with other factors other proteins and activate adaptive immunity is caused by activation induced cytidine deaminase that protein. So, APOBEC and AID have got a good interaction or good interrelationship.

So, if one example such as HIV virus which caused the AIDS and this HIV once it entered the cell and we call it as HIV virion because the HIV virion is the one which is capable of coming out of the cell as a fully infective particle. So, it can say in a dormant state also. So, what you can see this is a typical HIV particle, where it has got its own genome and also it has got an envelope etcetera.

And then virus also have specific proteins which it produce because which is not available from the host. A lot of virus functions are depend on the host protein, host enzyme, host machinery. But host will not supply everything what a virus needs. So, virus will carry only those things like you can see a simple example.

If you are going to write an exam, they will provide the chair or a table or a bench to sit and write they will also provide you paper to write the exam. They will not provide you dress they will not provide you pen they will not provide you if you want a pencil or eraser, no exam hall will produce you have to carry. But paper you do not have to carry,

paper they will give because your paper is not trustable because someone can you know manipulate it, but pen you do not you cannot manipulate a pen.

So, this is what you should understand. So, virus if it need sometimes if you are calling a skilled worker a carpenter to your home, you need to provide him a place and tool only. Carpenter will bring the tools which is required for his work. You do not have to supply tools to him because without those tools carpenter cannot work. So, same logic applies when virus comes virus makes use of lot of things from the host including some proteins such as APOBEC.

But very unique proteins which no host can provide virus brings in, just like a carpenter is bringing a tool or you carry pen for writing exam like that. So, if is a protein which is viral a specific and what happens is that if there is a specific loci is there, it can like; it can convert into A3, A3 protein can convert into A3G that is it can get converted into A3F and alanine become glycine or alanine become phenylalanine.

So, such kind of editing is possible with the help of when there is APOBEC protein is present in the cytoplasm. And when you are talking about AID immunoglobulin locus because we know we all know that when there is an infection it always brings in a changes at the hemoglobin normally, when you have an infection your blood serum contains plenty of antibodies which can target, but their number is less.

Say for example, if your blood serum has 10,000 different types of antibodies are there, which can target pretty much any known virus in your body, but their number is low. So, as soon as you get an infection one of them, these innate antibodies will go and recognize, but they cannot conquer they cannot completely eliminate the infection

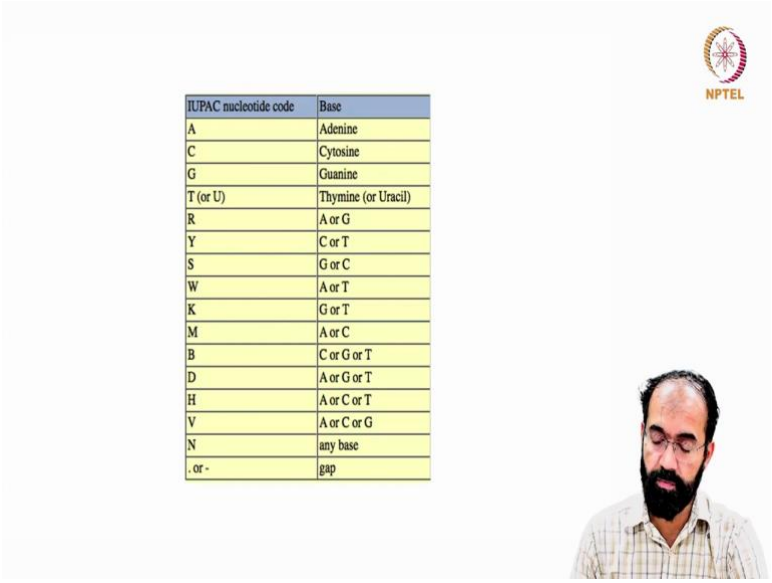
But this small percentage which could detect the host will be amplified. And some normally in the immune cell the VDJ Variable Domain Mutations are very much possible the variable region. So, that there is a concrete backbone of an antibody and the variable loci will which will be the pairing portion which can bring in changes.

And if these changes are able to target a virus, then that particular change bearing antibody will be expanded. It is just like a you are running a shop and you do not know what people need it. Once you produce a product and there is demand you kept 10 products, but 3 products have got heavy demand, then what will you do? You will start

producing these 3 products more and you will reduce the production or stop the production of the 7 other items.

Like that if there is a antibody which detected the host then the organism will expand it. So, that is why antibodies usually have the variable region. And this is a notation WR and C. C we all know it is cytidine and W and R are code for nucleic, R stands for purine and W stands for A or T or A or U in that case. So, in this conversion WRC to WRU is very important for the functioning of the AID and this is basically the switch regions of the antibody.

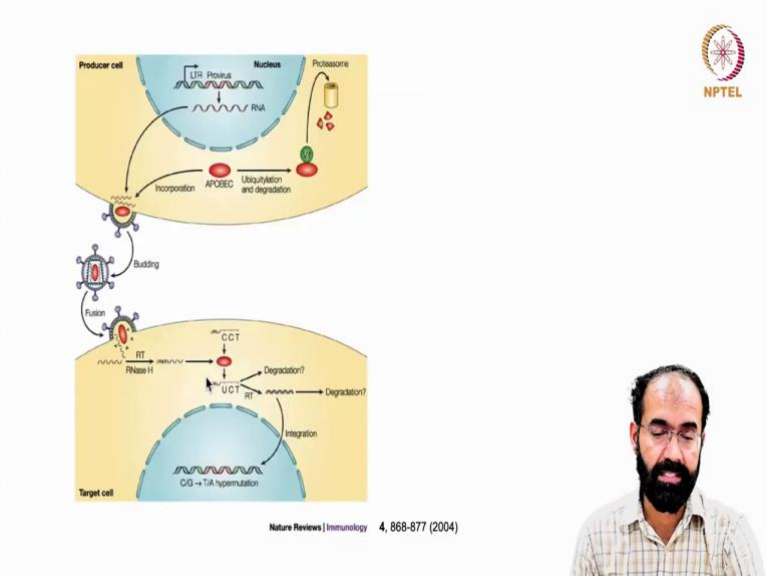
(Refer Slide Time: 17:31)



IUPAC nucleotide code	Base
A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	any base
. or -	gap

For your convenience we will show that this is the IUPAC nucleotide code A, C, G and T or U and that is the I mean the nitrogen base, that is the nucleotide and R stands for A or G which is nothing but purine and y stands for pyrimidine that is C or T so on and so forth you can see it and that is the code which is used in the previous slide.

(Refer Slide Time: 17:53)



Now, coming back to the virus what happens is in a producer cell. Usually the virus gets integrated, the viral because HIV is a RNA virus and it converts itself into DNA and integrates into the genome of the host and that is why we call it as retrovirus. And it is staying in the genome we call it as a provirus. So, provirus basically means it is a virus, but it is not a virus. It genetically it is a virus and that once gets converted into RNA that RNA can get converted into protein.

And I think HIV genome is around 30,000 bases, which is much higher than any other virus usually virus will have around 10 to 12,000 bases only, but HIV genome is little bigger. So, what happens? When this viral RNA is produced, it can happily produce the viral protein, viral code proteins etcetera and it can happily get itself converted into a functional virus come out of the cell and infect other tissue like you can keep spreading like virus once it is entered into the genome it can transcribe and come out and make fresh virus anytime as the virus wish.

If your immune system is strong and you eliminated all the virus then HIV will keep quiet it will stay dormant in your genome. So, whenever your immunity is little low it will start producing from the genome. So, it has the talent. So, what we should understand, the APOBEC is from your body ok. And APOBEC is a very interesting situation because normally if APOBEC is produced, it has to do its job that is causing the cytidine deamination.

Usually, it undergoes degradation if Vif protein is present, Vif protein is from the virus it is sent for degradation. Otherwise, if it is not degraded, what it will do? It will incorporate into the viral particles and it can come out and get into a host cell and it can carry virus along with it

And it can sometimes you know if there is a host derived RNase, RNase H. RNase H is an RNase enzyme that can specifically degrade the RNA part of a DNA-RNA hybrid. If RNase H etcetera comes into action it will be degraded, but what the system will do is that it will make use of this APOBEC and it can decide whether it should be sent for degradation or whether it should reverse transcribe and go into the host or integrate into the host.

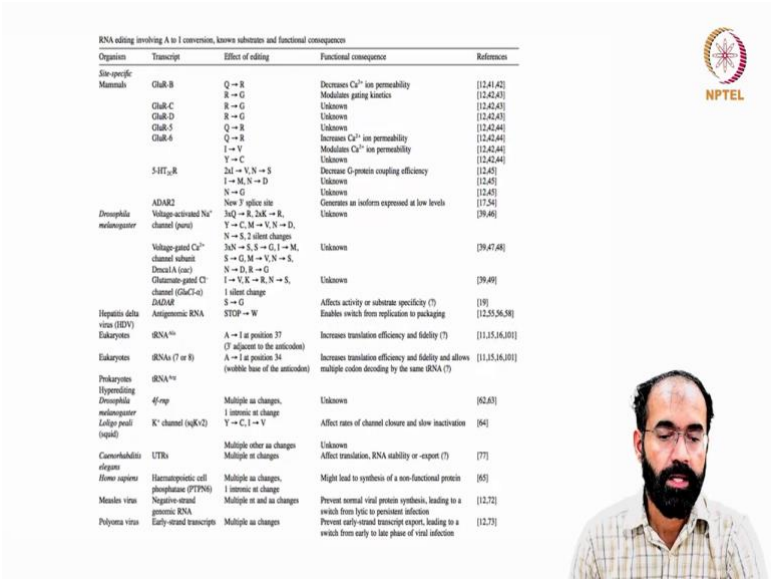
So, if CCT or CCU what you are talking, if it is there CCT if it gets converted into UCT because of the APOBEC because we know APOBEC is a cytidine deaminase. It gets converted into UCT it can have completely different meaning for getting into if it got converted into UCT it is capable of going into the genome. But if it did not a part of it can also undergo the degradation and even this capable of integration also can mark into degradation.

So, what we should understand that this is a strategy it can easily make use of APOBEC either mark the APOBEC for degradation, while in presence of Vif protein. And if Vif protein is not there the virus is not producing it then it can be APOBEC can carry along with the virus and go into the host cell even if host cell is not providing APOBEC it can happily make use of it.

(Refer Slide Time: 21:50)

And those who are interested can read this retroviral restriction by APOBEC proteins, it is a very interesting article because this will have more in depth knowledge and subject to understand because if I start explaining it will take much longer time to understand the entire dynamics of APOBEC protein. And also understand APOBEC proteins are important in the cellular homeostasis of various other tissue types other than this immune related function because they are important even in the regeneration biology also.

(Refer Slide Time: 22:30)



So, this is a table it is a very elaborate table that RNA editing involving A to I conversion and the non-substrates and the functional consequences. Like you can see in mammals, drosophila, hepatitis virus etcetera etcetera. And the transcript, transcript basically we are meaning the gene that is GluR-B, GluR-C like that. And what is the effect of editing? These are all single letter code for amino acid Q to R means glutamine to arginine, R to G means arginine to glycine like that it goes because 20 amino acids also have got a single letter code.

And you also see, what is the functional consequences because decreases the calcium ion permeability in one case, like that you can see the whole list because showing this list is not for memorizing purpose, showing this list is for you to have an idea, how many genes in how many organisms it is even in Caenorhabditis elegans and measles virus, a polyoma virus etcetera also make use of this RNA editing.

So, RNA editing is not a chance even RNA editing is a planned, purposeful event because of which various viruses and various other eukaryotic genes are able to make a proper living and a proper way of functioning in the host. So, this is what you should understand in detail about the RNA editing.

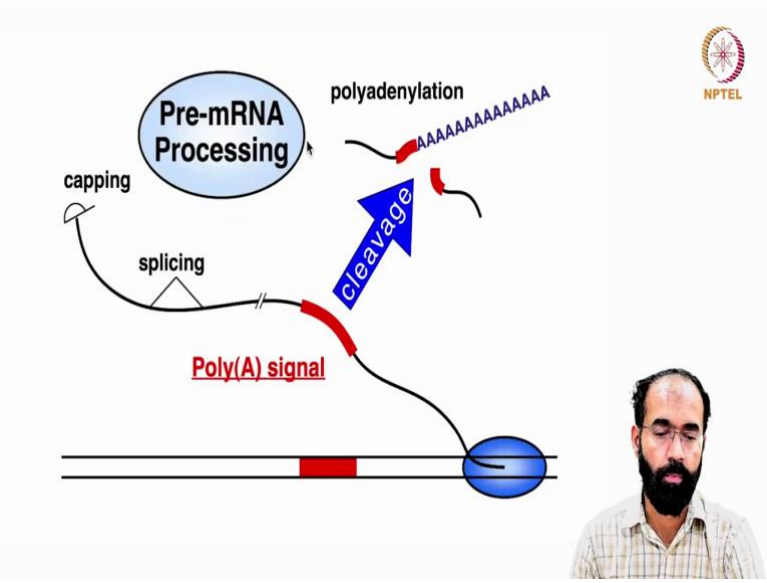
(Refer Slide Time: 24:00)

Now, we can understand more in detail about RNA splicing, RNA export and RNA decay because this is also part of this section. And in this section, we may get into some topics which we may have already covered little bit, but understand this touching upon is

necessary. So, that you can; first of all, you can recap little bit and then it also can bring in the context relevance or the importance of the topic that is being introduced in this section.

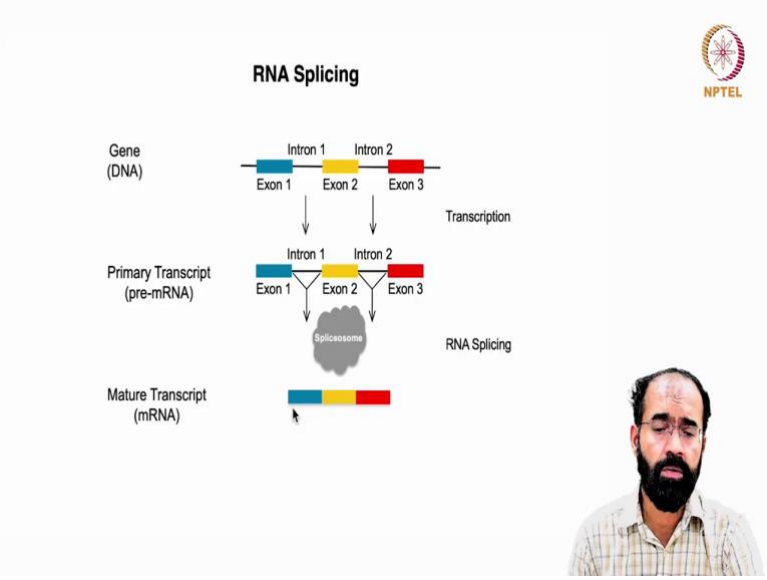
So, do not think that some topics are being repeated because they are intentionally put in, but I have tried the best to be possible to avoid the feeling of repetition because they are basically for memorizing the previous topic.

(Refer Slide Time: 24:47)



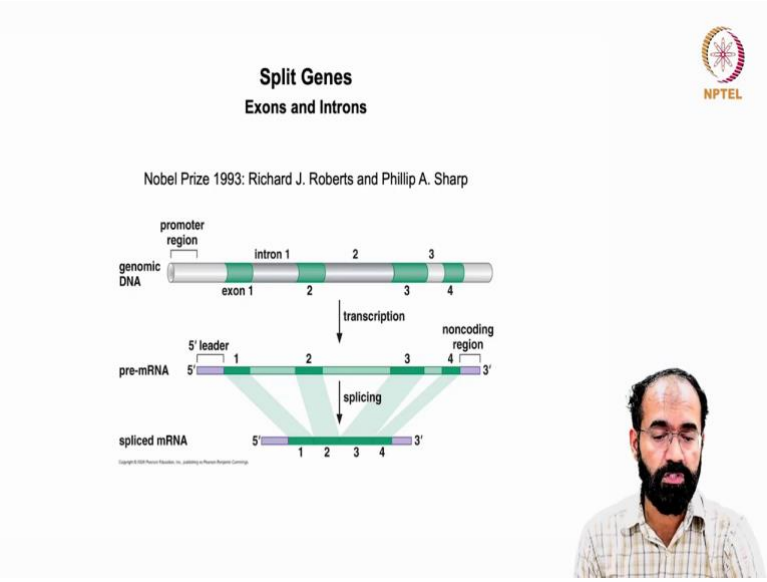
So, we know the pre mRNA processing have got capping, splicing, poly A signal, cleavage and polyadenylation etcetera. This happens from the DNA.

(Refer Slide Time: 25:01)



Now, when it comes to RNA splicing, we know it comes from the DNA and it has got exon 1, exon 2, exon 3 and also have got 2 introns and the spliceosomal machinery works on the primary transcript and or the pre mRNA and causes the production of the mature transcript. We know this is basically the fundamental principle.

(Refer Slide Time: 25:28)



And we also know that this has fetched the Nobel Prize for the discovery of introns by Richard Roberts and Phillip Sharp and that is why we call them as split genes and exons

and introns are very important not as an intruders or introns are not really intruders because they are very important for the normal functioning.

And we also need to understand that introns and exons hold some information about; because once the exons are fused you are technically having no idea, whether base number this to this say base number 1 to 200 was exon 1, 201 to 712 was exon 2 we have no idea, absolutely no idea that is what we assume. But the truth is no there are signatures on the RNA.

And there are even proteins that bind on to individual exons and give clue to the system that this is the boundary of an exon and you may wonder why that is needed, that is definitely needed because that is part of marking whether or not a given RNA should be sent for degradation or not.

So, we need to understand more in detail about that and we will see them one by one because each exon or each intron and the exon intron boundary even in the processed RNA or the post processed RNA or the spliced RNA this boundary have got relevance. So, who is providing this relevance etcetera we will see them one by one in the next class.

Thank you.