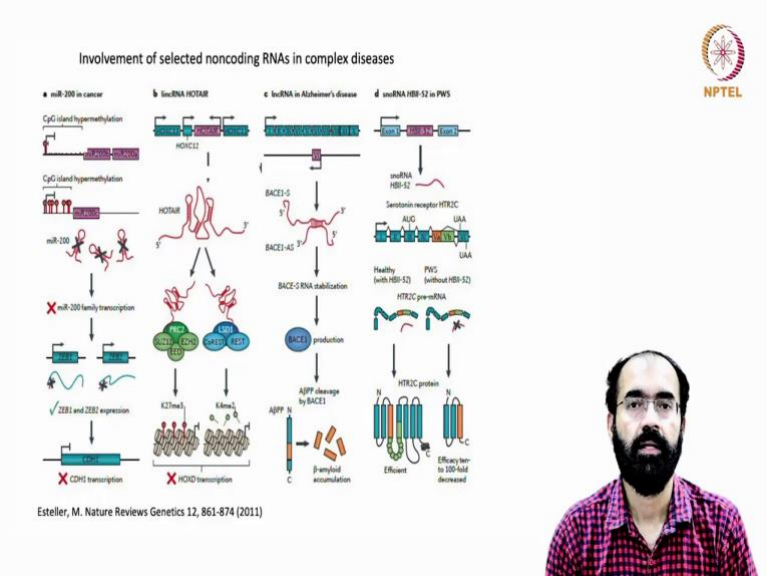
RNA Biology
Prof. Rajesh Ramachandran
Department of Biological Sciences
Indian Institute of Science Education and Research, Mohali

Lecture - 50
Dosage Compensation Xist and ncRNA in Imprinting: lncRNA-Induced Cancer

(Refer Slide Time: 00:21)



Hello everyone. Welcome back to another session of RNA Biology. So, we were here in the previous class, so we will revisit this once again. One we saw the role of miR-200 in causing cancer and long non-coding RNA hot air in causing a HoxD gene repression and lncRNA in Alzheimer's disease.

And we were here in the last class that is snoRNA HB2 S2 in Prader-Willi Syndrome which is caused mainly by a deletion of chromosome in chromosome 15 in and it manifests differential disease in males and females. So, how does it work? It is influencing the complete expression of a serotonin receptor gene HTR2C. So, how does it do? In healthy individuals, the non-coding RNA HB2 S2 is perfectly present. It is a snoRNA, small nucleolar RNA, which has the power of influencing its splicing.

So, in healthy individuals this RNA, non-coding snoRNA is present perfectly and you end up getting a complete receptor of serotonin. Serotonin receptor is complete. Whereas, in PWS that is syndrome affected people that is not having this snoRNA HB2 52.

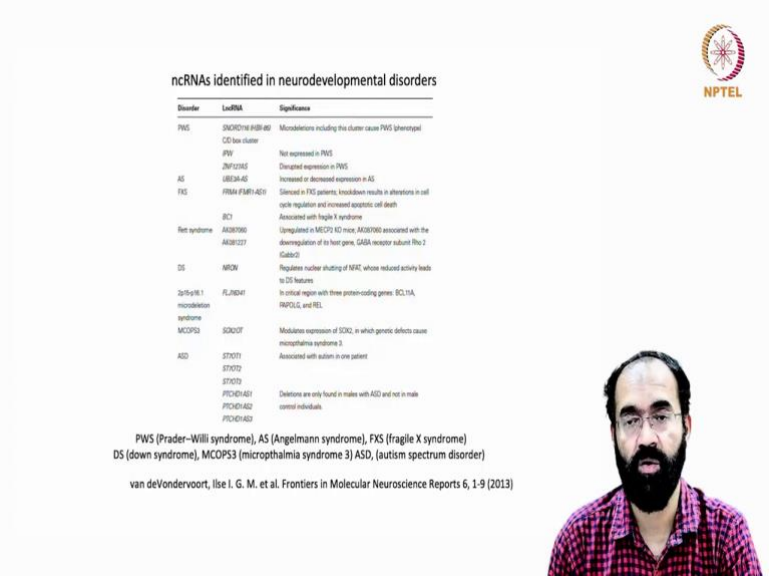
So, this will eventually lead to the incomplete processing; that means, it has got two stop codons. One is stop codon is in the splicing happens such a way that it can happen in the 6th exon or it can influence the beginning the stop codon can come in the beginning of the 6th exon.

So, this is mainly happening because of the alternative splicing that is created by created by the presence or absence of this non-coding RNA. So, when you have this non-coding RNA present, the splicing happens perfectly. When this snoRNA is missing the splicing can happen apparently, because of which the splicing defect can lead to an incomplete RNA, that is RNA, that is devoid of the entire exon will take place.

And the efficiency of this incomplete receptor is around 100 fold less than the wild type, which can leads to the syndrome phenotype. So, there are several serotonin receptor , this is one of the subtypes of the serotonin receptor. Do not think that this is the one and only serotonin receptor. If that is the case the organism will not survive.

This is one of the members of the serotonin receptor family which is not working at all, that situation is not there. But it is decreased functioning or reduced level of functioning that leads to a complex phenotype such as PWS. So, what we should understand that non-coding RNAs can influence various physiological conditions in the organism that can be cancer or a homeotic transformation or some diseases such as Alzheimer's and also influence the syndromes.

(Refer Slide Time: 04:05)

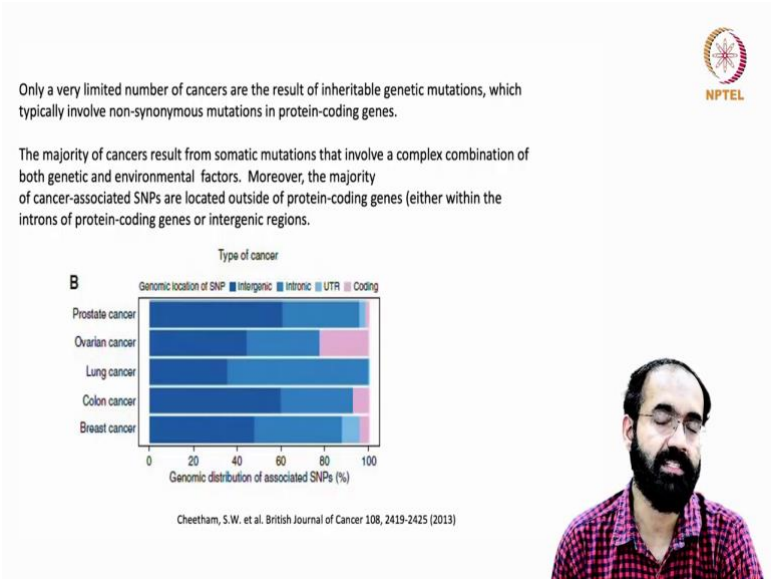


So, now let us see the non-coding RNA that are identified in several neuronal disorders, so or neurodevelopmental disorders. So, a bunch of them are expressed in the deceased conditions and the disorder is PWS and the LncRNA that are affected and the significance.

We will not go into the details through because several disorders are listed here red syndrome and so, going through individual syndrome is not important, like down syndrome we have referred to that it is a caused by chromosome 21 trisomy, like Prader-Willi syndrome, Angelmann syndrome and fragile X syndrome and down syndrome and microphthalmia syndrome and ASD autism spectrum disorder, all of you would have heard about autism. So, these are all to a great extent connected with the non-coding RNA.

And so, in this first column you see the name of the disorder, and second column you see the names of the LncRNA, and in the third column you will see that what are the significance of what is the role played by the micro RNA or the non-coding RNA in the creation of this actual phenotype.

(Refer Slide Time: 05:31)



So, only a very limited number of cancers are the result of inheritable genetic mutations which typically involve non-synonymous mutations in protein coding genes. So, we should understand the mutation in a gene and leading to causing a cancer is very limited. It is actually the tweaking of various gene functions.

So, majority of cancers result from somatic mutations that involve complex combination of both genetic and environmental factors. Moreover, the majority of cancer associated SNP, SNP means change in the single nucleotide polymorphism change in the DNA sequence. So, associated SNPs are located outside the protein coding genes.

So, outside the protein coding genes we should understand that it is not going to affect the protein sequence or amino acid sequence in a given protein. But outside a gene's coding capability, if there is an SNP how come it can lead to cancer. So, this is something to do with the non-coding RNA that is being produced.

So, these SNPs are either within the introns of protein coding genes or intergenic regions; that means, the gap between two genes. So, we should understand majority of these changes that is leading to cancer, cannot simply because of a altered protein structure because of a mutation in the codon.

We are not saying that absence of a amino acid or a change of an amino acid in a protein is not causing going to cause cancer. However, if you look at the number of cancers, many a times the proteins are unaffected still the SNPs and those SNPs that are found in the introns or intergenic regions are causing the cancer.

So, it tells us that the regulation of a gene's function or the even epigenetic or non-coding RNA mediated can get affected that can leads to the cancer. Or in other words, minor tweaking of gene expression is good enough rather than a complete change or a complete havoc in the gene's expression that is leading to cancer.

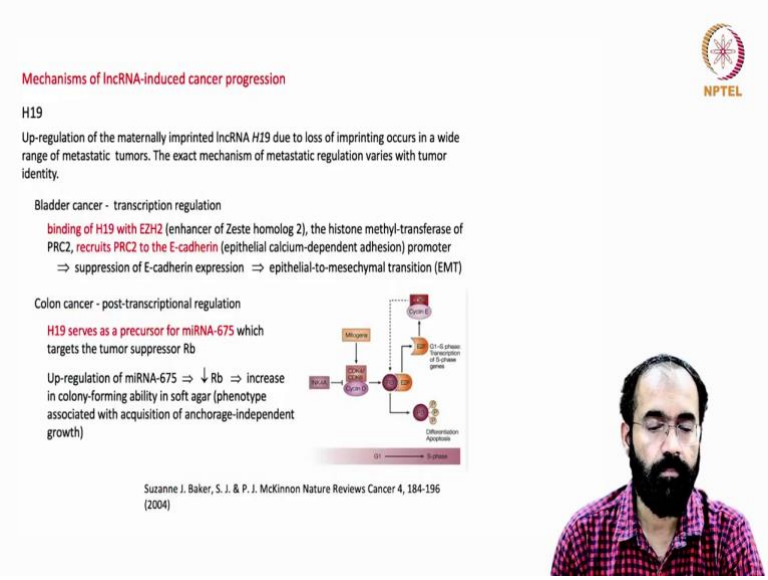
You can see a bunch of cancers in this picture, prostate cancer, ovarian cancer, lung cancer, colon cancer and breast cancer. If you see the SNP, SNP means change in the nucleotide. It can be mutation or it can be just an alteration. Let it be anything, but a change in the a variation. So, you can see the genomic location of the SNP.

The blue colour one, deep blue colour one is intragenic, the moderate blue colour one is intronic and light blue colour is the UTR and the pink colour one is in the coding region. You can see the pink colour is negligible, literally negligible. But in all the cancer, the SNPs are either in the intragenic or intronic or in the untranslated region.

The rate or the importance of coding region having a mutation is almost nil or almost negligible. So, this is something which we should keep in mind. The importance of non-coding RNA or the SNPs that is and they are occurring, they are not inherited from parents.

That simply happened because of your you know living habit or maybe because of the environmental exposure to a mutant etcetera can cause change in the SNP in a given tissue. Your other tissue may be fine which is not cancerous. A given tissue a change that is occurring can lead to cancer.

(Refer Slide Time: 09:18)



Now, let us see the mechanisms of long non-coding RNA induced cancer progression. So, one example we can study is the H19. It is a well-studied example. The upregulation of maternally imprinted long non-coding RNA H19 due to loss of imprinting occurs in a wide range of metastatic tumour's because it is a maternally imprinted means mother's copy, maternally imprinted LncRNA, and loss of imprinting means it is now expressed because imprinting is lost.

So, the exact mechanism of metastatic regulation varies with the nature of the tumours which tissue tumour you are talking about. Let us think about bladder cancer. So, what is the transcriptional regulation that is occurring with regards to H19? So, binding of H19 with EZH2, EZH2 is a catalytic domain of PRC2 complex. Its full name is enhancer of Zeste homologue 2.

So, the histone methyl transferase role played by EZH2 gives PRC2 its identity as a polycomb repressor complex protein. It recruits, so this H19 along with EZH2 recruits the entire PRC2 complex to E-cadherin. E-cadherin is the epithelial calcium dependent adhesion molecule.

So, E-cadherin helps in anchoring the cells onto a specific loci or to the base cell membrane etcetera. So, this H19 brings in PRC2 complex to the promoter region of E-cadherin. And this will lead to the suppression of E-cadherin expression and epithelial to mesenchymal transition takes place. That is also called as EMT. Epithelial to mesenchymal transition is one of the requirement for the metastatisize metastasis of the cancer.

Metastasis is the secondary tumour. Say a tumour originated in the liver. It will find its secondary colonies in the brain, heart or any other or liver or kidney or any other organ and that will inhibit the normal functioning of that organ and this can be very problematic. Especially, when the cancer is metastasized it can cause lots of problem in the normal functioning or normal functioning of that particular tissue and the treatment becomes meagre.

So, in colon cancer what happens? It is a post-transcriptional regulation. What you saw here is a transcriptional regulation with in the bladder cancer H19 on E-cadherin transcriptional regulation. Now, in colon cancer it is a post-transcriptional regulation. H19 non-coding RNA serves as a precursor for miRNA-675 which targets the tumour suppressor retinoblastoma. Retinoblastoma is a tumour suppressor protein. So, the upregulation of micro RNA 675 because H19 is a precursor for that.

What happens? You have a decreased retinoblastoma protein available because which is a tumour suppressor. So, this can cause an increase in the colony forming ability of these cells in soft agar assay where you are checking the colony forming ability of this cell. And this is a phenotype associated with acquisition of anchorage independent growth. Cancer cells hallmark feature is anchorage independence means they can grow without attaching onto the base cell lamina.

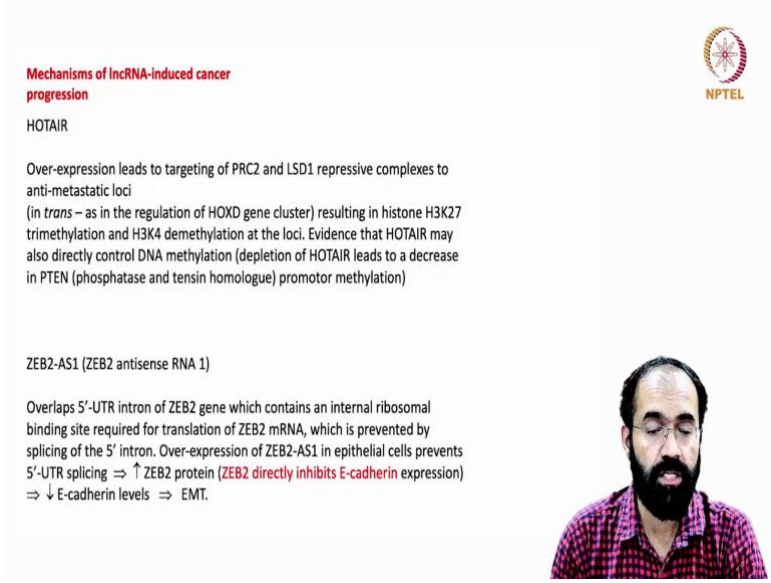
Normally, a healthy cell needs somewhere to attach. Even if you grow in a Petri dish it will grow say uniform layer in the bottom of the Petri dish, but it will not climb on each other. Everywhere it need to have a specific number of contact, means anchorage

dependent way of growing can be lost. So, that quality is lost because the absence of 675 which normally can influence the retinoblastoma which is a tumour suppressor.

So, let us revisit this again. The upregulation of miR 200 miR 675 which is originated from H19 can cause a decrease in the retinoblastoma which is a tumour suppressor protein and increase in the colony forming ability; that means, anchorage dependency can be lost because of reduced retinoblastoma level.

So, this is another way of non-transcriptional regulation. It is post-transcriptional regulation of a non-coding RNA H19 on causing the cancerous property.

(Refer Slide Time: 14:08)



So, now let us see another exam example of mechanism of lncRNA induced cancer progression in the case of hot air. We have seen this hot air because it is normally helpful in recruiting the PRC2 and LSD1 to specific loci.

So, over expression leads to targeting of PRC2 and LSD1 repressive complex to anti metastatic loci, and which in trans, it is also important in regulate in repressing the expression of HOXD. So, the absence of hot air causes the de-repression of HOXD, presence of hot air causes repression of HOXD gene cluster.

And this results, when this PRC2 and LSD1 are recruited, it will result in H3K27 tri-methylation and H3K4 demethylation at the given loci where this HOTAIR is interacting because HOTAIR physically brings in the PRC2 and LSD1. So, evidence that the

HOTAIR may also directly control DNA methylation that is the depletion of HOTAIR has been shown to decrease in another tumour suppressor called PTEN which is phosphatase and tensin homologue and its promoter methylation is affected.

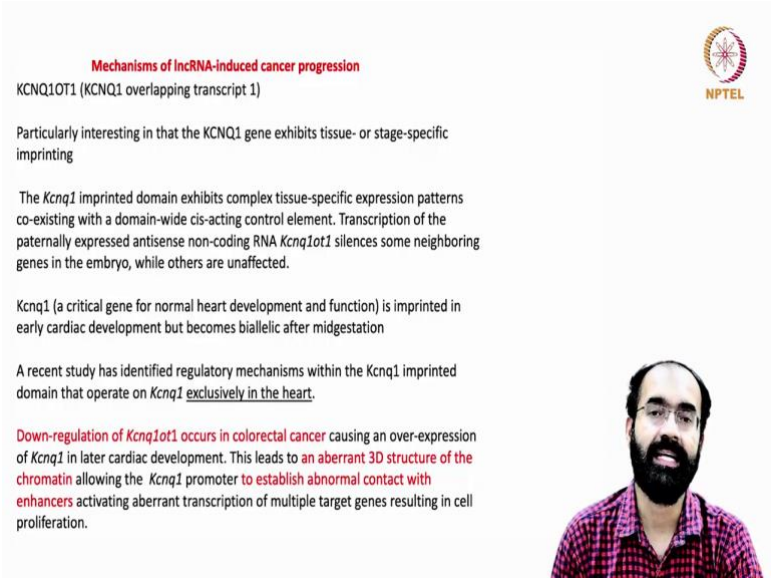
So, we should understand one non-coding RNA not only influence a promoter region via LSD and PRC2, but it can also influence the DNA methylation. DNA methylation in turn high DNA methylation. DNA methylation means it is in the cytosine of the CpG island not; we are not referring to any random DNA methylation. So, DNA methylation often leads to reduced gene expression.

Now, let us think about another example of ZEB2 antisense 1. So, ZEB-2, we know that it is an important gene in the development and also contributing to the cancer formation. It overlaps, the ZEB-2 antisense 1 overlaps the 5 prime UTR intron of the ZEB-2 gene which contains an internal ribosome binding site required for the translation of ZEB-2 mRNA, which is prevented by the splicing of the 5 prime intron because of this overlap region which it has got.

So, the over expression of ZEB-2 antisense 1 in epithelial cells prevents the 5 prime UTR splicing. 5 prime UTR splicing do not happen means the ZEB is affected. So, this can eventually lead to the up regression of ZEB-2 protein because the 5 prime untranslated region controls the controlled elements.

So, ZEB2 can directly inhibit the E-cadherin expression because it is a negative regulator of E-cadherin expression. And this can bring down the E-cadherin levels and that in turn affects it can facilitate the epithelial to mesenchymal transition. Another way of facilitating the metastasis in cancer or secondary tumour formation in the cancer. So, antisense RNA influence in multitude of ways the gene expression events that can eventually lead to cancer progression.

(Refer Slide Time: 17:35)



So, now let us see some mechanisms of long non-coding RNA mediated cancer progression in another example, apart from what we saw already. One example is KCNQ1 overlapping transcript 1. It is written as Kcnqlot1. So, it is very interesting RNA non-coding RNA that it can inhibit tissue or stage specific imprinting. We know imprinting can come in multiple ways, it can be tissue specific or it can be which gene originated from what parent.

And some of them are imprinted for life, like we know we saw bar body the entire chromosome is imprinted for rest of the life of that organism and based on whether it is a paternally is imprinted or maternally is imprinted, it is a random. Some genes it is maternally imprinted, in some genes it is paternally imprinted.

So, it is very important this Kcnqlot1is very important in affecting the stage and tissue specificity of imprinting. So, the Kcnq imprinted domain exhibits complex tissue specific expression patterns coexisting with a domain-wide-cis-acting control element. So, it comes with a cis regulator along with this non-coding RNA.

So, the transcription of paternally expressed antisense non-coding RNA KCNQ1OT1 silences some neighboring genes in the embryo while others are unaffected. So, wherever the Kcnq1 is expressed just like similar to the x inactivation centre, where Xist is coated, it spreads right. So, the neighboring genes are affected. So, the rate of

expression of Kcnqlot1 can influence the neighboring cell because it spreads, the gene repression even it spreads.

So, Kcnq is a important gene for the normal heart development and its function. So, if Kcnq1 O; Kcnq1 is an important gene for the normal expression. So, the Kcnq1 overlapping transcript presence of that can tweak the expression of Kcnq1. So, this is also an imprinted gene. And; that means, maternal and paternal copy it is a parental preference is exhibited.

And in early cardiac stage development because the Kcnq1is very important and detrimental and, but it becomes later on biallelic. Early stage in cardiac development it is imprinted whereas, later the gene switches to biallelic gene expression by around mid-gestation period.

So, what we should understand? The damage what Kcnqlot1 can do on to Kcnq1gene will already come to effect in the early stage because it need to influence only one copy of the gene because that time it is imprinted. Later on, Kcnq1imprinting is lost. But so, the Kcnqlot1 expression on Kcnq1expression is quite detrimental especially in the cardiac development.

So, a recent study has identified regulatory mechanism within the Kcnq imprinted domain that operate on Kcnq1 exclusively in the heart. So, what we should understand this overlapping transcript of Kcnq1 can influence the dynamics of Kcnq expression exclusively in the heart, not in the rest of the body. So, this can cause a major impact in the heart development.

So, down regulation of Kcnq ot that is a non-coding RNA. Kcnq1 is an mRNA, Kcnqlot1 is a non-coding overlapping transcript occurs. So, if you down regulate down regulation of Kcnqlot1 occurs in colorectal cancer that is post cardiac development, like colorectal cancer can come in any adult. Causing an over expression of Kcnq1means Kcnqlot1 is a negative regulator of Kcnq1. So, it will cause an over expression of Kcnq1in later on cardiac development in the embryos also.

So, this leads to an aberrant 3D structure of the chromatin allowing the Kcnq1 promoter to establish abnormal contact with enhancers activating aberrant transcription of multiple target genes resulting in cell proliferation. So, what we should understand, Kcnqot role in

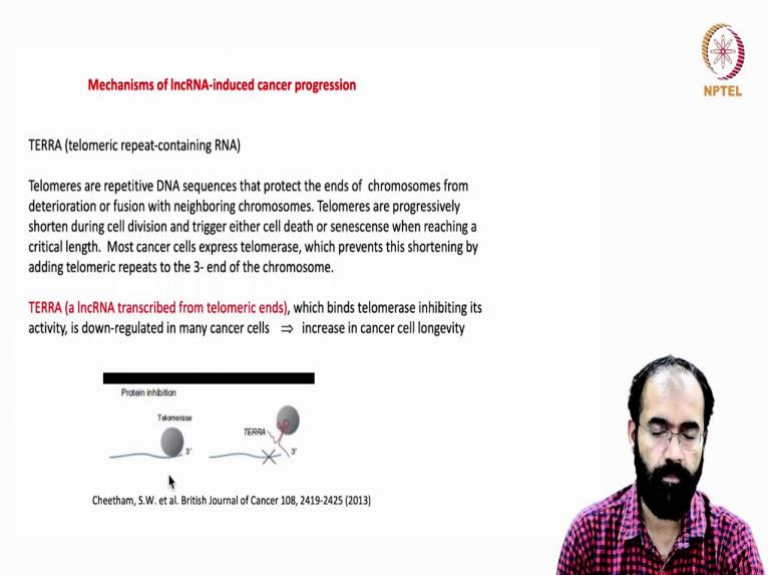
regulating the Kcnq can have a negative impact on the cardiac development, but this can also cause a aberrant 3D structure of the chromatin. Chromatin usually can influence various cis and trans regions in the genome. Cis means same chromosome, trans means another chromosome.

So, this 3D structure can go aberrant. So, that it can it is just like instead of putting food into your mouth, you are putting into your nose. So, they are close by places, but do is it acceptable it will go to your lung and you will die, you will choke and die, although mouth and nose are only one inch in distance.

So, the 3D structure an unwanted chromatin comes in contact with a active chromatin then that unwanted area also will turn on and that can leads to lot of complication. So, that is what happened. The aberrant 3D structure of the chromatin allow the Kcnq promoter to establish abnormal contact with enhancers. Enhancers means gene facilitators.

Enhancer elements activating aberrant transcription of multiple target genes resulting in cell proliferation and needless to say cell proliferation can lead to cancer.

(Refer Slide Time: 24:00)



And another mechanism of long non-coding RNA mediated cancer progression. One example is TERRA, telomeric repeat containing RNA that short form is TERRA. So, telomeres we know it is structures in the tips of chromosomes. And they are repetitive

DNA sequences that protects the ends of chromosome from deterioration due to end replication problem.

And you can genes can be lost; the chromosomes are eroded the genes located in the tips of the chromosome can be lost. So, you do not want that to happen. So, telomeres are progressively shortened during cell division and trigger cell death. If the telomere length is below certain limit that cell will be marked for cell death.

If it is not marked for cell death, then that can lead to cancer because telomerase activity is seen only in germ cells and also in cancer cells. Cancer cells revive it, and lead to a active proliferation and that time cancer cell do not want to undergo apoptosis. So, they revive the telomerase.

So, most cancer cells express telomerase, so that they will not die cellular senescence prevented, so which prevents the shortening by adding telomeric repeats to the 3 prime end of the chromosome. So, telomerase have got a this T2AG3 sequence is there in an RNA form and it will convert it into DNA and attach onto the chromosome end if telomerase is there. Usually, in a healthy cell it is missing. Only in germ cells like testes and ovary cells have that.

So, what a TERRA will do? TERRA is a long non-coding RNA transcribed from the telomeric ends, which binds to telomerase inhibiting its activity and it is down-regulated in many cells. So, cancer cells, many cancer cells you do not want this TERRA to be present because if TERRA is present telomerase activity is inhibited, and absence of TERRA leads to a enhanced functioning of the telomerase and that can increase the cancer cells longevity.

So, TERRA which binds to the telomerase inhibiting its activity the usual case, but it is down-regulated in cancer cells because of which telomerase is active and it can leads to cancer longevity because the telomere is revived in those situations. So, that is what happens. You have the telomerase enzyme, and the TERRA prevents its activity, absence of TERRA the telomerase is allowed to function properly.

So, we will continue to learn more about the antisense RNA and regulating various developmental and disease conditions with examples in the next class.

Thank you.