## Introduction to Complex Biological Systems Professor Dibyendu Samanta and Professor Soumya De Department of Bioscience and Biotechnology Indian Institute of Technology, Kharagpur

## Lecture 25 Cancer

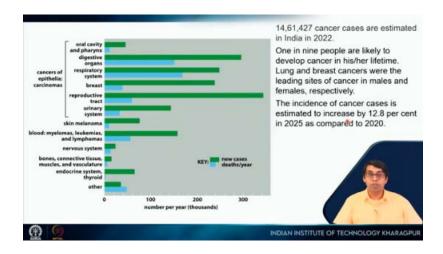
Welcome to this course introduction to complex biological systems. So, this will be the last lecture of week 5. In this lecture, I am going to discuss cancer. Cancer is a collection of diseases.

So, all of these diseases are concerned with cells, different types of cells. So, you can think of that different there are so many different cell types. So, there are so many different cancer types. Some of the most predominant cancers are listed here on the y-axis, and the x-axis tells you the number of people who are infected by this disease per year, and this data is taken from the United States. We can see that some diseases, some cancers are more prevalent, and some cancers are less prevalent.

So in this case, what you see is in green shows the number of people in thousands who are affected each year, who are diagnosed with this particular cancer each year, and in blue shows the number of deaths. So it turns out that cancer has become one of the major killers. So, this is in the context of the United States, and here is some data for India. So, in the year 2022, 14,61,000 people were diagnosed with different forms of cancer. Again, considering the population that we have, this number is underestimated.

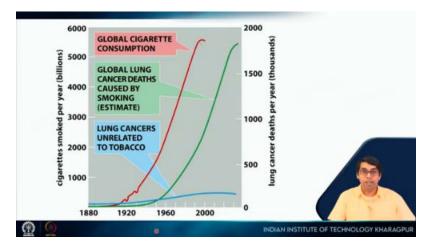
So, there are many people, most probably from rural areas, who are not getting diagnosed. So, one in nine people is likely to develop cancer in their lifetime in our country. It turns out that lung cancer is most prevalent in males, and breast cancer is most prevalent in females. The incidence of cancer cases is estimated to increase by 12.8% in 2025 compared to 2020. So, that's a big jump.

So, with a more aging population and a large population, this has become a real issue for our country. So, one thing that I would like to point out is that lung cancer, which is most prevalent in males, has a very good correlation with smoking. So, in this plot, what you can see is this is years, so this is time, and these are the number of cigarettes smoked per year by people. The red line shows you global cigarette consumption, so how the global cigarette consumption has increased over the years.



So, now with all governments have put in a lot of effort to educate people about the problems of cancer. So, this number you will see is slowly coming down. The green line shows the global lung cancer deaths caused by smoking. So, you can see there is a direct correlation. So, of course, it takes some time for cancer to develop.

Therefore, this green line lags this red line, but they look very similar. So it is following exactly the same trend as that of cigarette smoking. And this blue line shows lung cancer that is unrelated to tobacco, and that is very less. So this is something that we will see that there are causes of cancer which are sporadic, but then there are also causes of cancer which are directly correlated with our lifestyle. So, this slide shows you some of the basic events that happen during the formation of cancer.



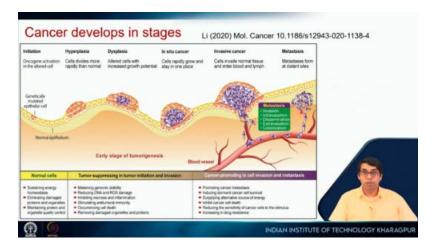
So, the cells that you see here, these are called epithelial cells. So, these are the ones that are on the top. Now, let us say one of these cells gets mutated. Now, this mutation can happen during cell division. So, in the last lecture, we have seen that stem cells are there, and there are cells that die.

So, stem cells have to divide to replenish the to replace the dead cells. So when that happens, when DNA replication is happening, there can be errors. So that can lead to mutation. Also, errors can come from some other agents.

So, errors can come from the internal process of DNA replication. Errors can come from external agents. We will see some of these agents in the later slides. So, let us say for some reason a mutation is caused, and that mutation happens in some critical protein, so that the cell can divide more rapidly than a normal cell. So, this is called hyperplasia.

So, here these cells look very similar to normal cells, but they can divide faster. and while this happens, since you are getting so many cell divisions, so many DNA replications, again more errors can occur. So, some cells will gain additional mutations in some critical proteins or genes, which will lead to this increased growth potential. So, now you see these extra cells which are different from these cells and those cells will divide rapidly, and what we get is a tumor. So, at this stage, this will be a benign tumor because this tumor is localized in a particular site.

So, at this point, you can actually cut it out. So, if this is diagnosed in time, a surgeon, a good surgeon, can cut it out so that they can remove this tumor. Whole mass. So, no more cancer cells. However, if detection is not done in time, there is a probability that these cancer cells will gain some more functions due to more accumulating mutations, and they can break this tissue and start moving.



Once they get into the bloodstream, they can go to different parts of the body, and these cells can accumulate in different places and they start growing as tumors. So now the tumor will spread throughout the person's body. So this is invasive cancer, and the last stage is metastasis, which is very difficult to treat. So if you are interested in cancer biology, then

these are some of the terms that you will come across, and it is a good idea to familiarize yourself with these terms.

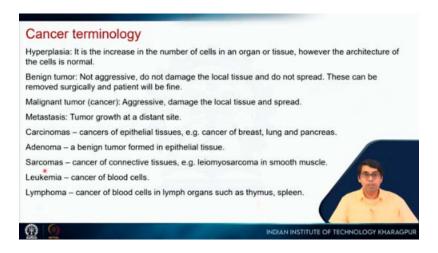
So the first one I have already talked about is hyperplasia. It is the increase in the number of cells in an organ or tissue. So an increase in number beyond what it normally should be. However, the architecture of the cells is normal. So it looks very similar to normal cells.

Then we saw a benign tumor. It is not aggressive, it does not damage the local tissue, and it does not spread. So it remains localized in a particular place, and this can be removed surgically, and the patient will be fine. A malignant tumor, which is now cancer, is aggressive, it can damage the local tissue, and it can spread. So we already saw that in the last slide.

And metastasis is when these cells get into the bloodstream, get transported to different parts of the body, and start growing as a tumor at a distant site. So, this is metastasis. Now, it turns out that there are different types of cancer, and depending on the cell type, they get different names. For example, you will hear of carcinomas. So, carcinomas are cancers of epithelial cells.

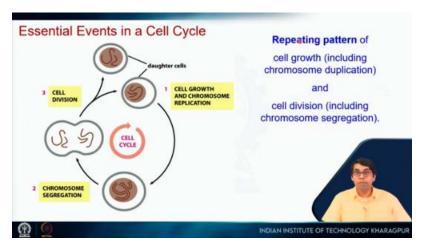
So, for example, breast cancer, lung cancer, and pancreatic cancer. So, these are all carcinomas. An adenoma is a benign tumor formed in epithelial cells. So, if it is localized in a particular place and if it is a benign tumor, it will be called an adenoma. Sarcomas are cancers of connective tissues.

For example, in muscle cells, when that happens, it is called leiomyosarcoma. Cancer of blood cells is called leukemia. The same blood cells, most likely the white blood cells, when they accumulate in lymph organs such as the thymus or spleen, then that will be called lymphoma. So, these are all cancers, but depending on the cell type and the location, they are given different names. So, just a reminder about the cell cycle we have already seen: that a cell decides to divide.

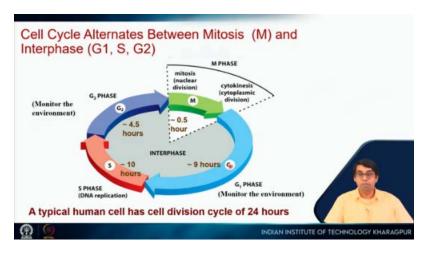


So, there is cell growth and chromosomal replication that happens in the *S* phase, then in the metaphase, chromosome segregation happens, and then we get the cell division. So, two daughter cells are formed, and we have seen in the case of stem cells that it can be symmetric division or asymmetric division. Now, many things can go wrong. So, when the chromosome is replicated, mutations might occur because of errors in DNA replication. When chromosome segregation is happening, errors might occur if the mitotic machinery, for some reason, does not work properly in that particular cell.

So, you might actually, instead of getting two chromosomes here and two chromosomes here, get three chromosomes here and one chromosome here. So, those will also create problems. So, the repeating pattern of cell growth, including chromosomal duplication, and cell division, including chromosomal segregation, can lead to cancer. So again, we have seen that there are several checkpoints. So, this is your typical cell cycle: *G1* phase, *S* phase, *G2* phase, and *M* phase.



There are checkpoints that monitor these phases and ensure that no problem has occurred. But when there are mutations in proteins that are involved in these checkpoints, then the probability of getting cancer increases. So, what happens when this cell cycle regulation goes wrong? Then, uncontrolled proliferation or cell division happens, which can lead to the formation of a tumor. So, cell division without checking for DNA damage.



So again, when the DNA is being replicated, DNA damage can occur, and mutations can happen. So there are proteins, there are enzymes which are called DNA repair enzymes, which check, identify these mutations, identify this damage, and correct them. But if those corrections are not made, it can lead to cancer. So DNA replication errors are not corrected. DNA damage can happen during replication, and DNA damage can also happen due to some chemicals which are called external mutagens.

So, again, if those mutations are not corrected, they can lead to cancer. But again, remember that not just one mutation is enough; there have to be multiple mutations in critical proteins or critical genes. When that happens, then the chances of cancer increase. So here are some numbers. So the error rate of DNA polymerase is 1 in  $10^5$ . So now our genome is 6 billion base pairs long.

So you can see how many mutations will happen. It turns out that in one round of replication, there will be 120,000 errors. So that's a huge number of errors. However, I told you that there are DNA repair mechanisms. So, there are proteins which will detect these errors, detect this mismatch in the DNA that has happened, come, and correct it.

So, all of these reduce this from 1 in  $10^5$  to 1 in  $10^9$ , which means that 3 to 6 mutations will happen during one round of replication. And then again, by chance, most of those mutations will happen in places that are not critical. Again, mutations might happen in a genome, in

a gene, but the mutation might not be such that it changes the characteristic of the amino acid. So, when I talk about evolution, I will talk about this in more detail. So, you will see that in that case, what happens is that the mutation will not have any effect on the structure or function of the protein.

But in some cases, this mutation can create problems. So, the chances of that are less, but considering the number of cell divisions that happen, if you add that, then you will see that the chances of getting a problematic mutation increase. But in most cases, all of those are taken care of by mechanisms that are present in the cell, like DNA repair. But when mutations happen in those enzymes, which are responsible for DNA repair, then we have a problem. So, during replication, DNA polymerase has proofreading activity.

So, it will reduce this error. After replication, a well-designed DNA repair machinery repairs the damage. So, that also reduces this number. So, it ultimately reduces it to 1 in  $10^9$ . But, if DNA damage is beyond repair, then the cell dies by initiating apoptosis.

So, there is something that is also there. So, I have not talked about this, and you will see this in the later lectures that if the repair cannot be performed. If the DNA damage is beyond repair, then it will trigger certain events in the cell, which will result in the death of that cell, which is called apoptosis. So, this is regulated cell death. Now, when this DNA repair or apoptosis, when both these fail, then we have a chance of cancer.

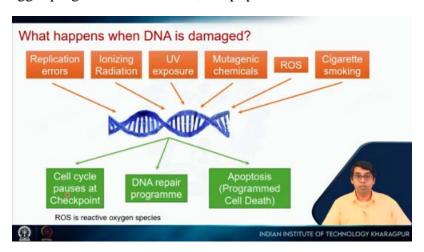


So, there are many checks and balances present in the cell which prevent cancer from occurring. So how does DNA damage happen? So, this is DNA. There are several factors. I talked about replication errors. You can also get cancer from ionizing radiation. Mutations can also happen from UV exposure. So, we will see that UV light can cause certain reactions. So, it can result in DNA damage or mutations.

Now, this is something that is very interesting. So, in next week's lecture, I will talk about vitamins and you will see that there is a particular vitamin, vitamin D, which is synthesized only when you get UV light. So, UV light triggers a particular reaction which results in the formation of vitamin D and vitamin D is very important for us. So, you need certain exposure to sunlight, but overexposure can cause unwanted mutations in the DNA, which can lead to cancer.

So, there is always a balance, and this is something that you will see in biology. If the balance is upset, then problems happen. There are also chemicals which are mutagenic, so mutagenic chemicals can cause mutations in the DNA. There are reactive oxygen species which happen due to certain reactions that occur in our body. They can also create mutations, and of course, there are proteins and enzymes which deal with this type of reactive oxygen species. Cigarette smoking, which is a lifestyle factor, involves tobacco that contains certain mutagenic chemicals, which can cause DNA mutations and we have already seen examples of data on cigarette smoking and how it is related to lung cancer.

So, these mutations can cause these problems. So, they can create pauses at the checkpoint. So, mutations happen. So, the cell cycle pauses at the checkpoint and it can try to fix it. If that does not happen, there is a problem. If there are mutations, the DNA repair program kicks in. It will remove that mutation and correct the DNA. So, if that happens, then the cell will proceed through the cell cycle, and everything is fine. If that does not work, then the cell will trigger programmed cell death, or apoptosis.



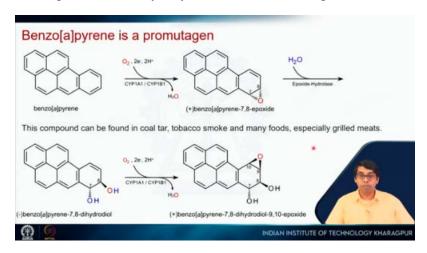
So, I talked about mutagens. This is one example, called benzopyrene, benzo(a)pyrene, and it is a pro-mutagen. So, what is a mutagen? A mutagen is a chemical which will cause mutations in the DNA. So, how do you check whether a chemical is a mutagen or not? There is a test called the Ames test.

So, one can take these chemicals and put them in bacteria. So, there are certain ways you will do this experiment and then you check for mutations in the bacteria. So, if mutations happen in the bacteria because of these chemicals, you say that chemical is a mutagen. Now, benzo(a)pyrene is a pro-mutagen because if you do this Ames test with benzo(a)pyrene, you will see there are no mutations. But people have seen that this is present in tobacco and it causes cancer.

So, what happens? It turns out that this molecule itself is not a mutagen. However, when it is inside our body, our body tries to flush out this type of molecule. So, it goes to the liver and the liver will do certain reactions so that it can be easily flushed out. One of those reactions is these oxidation reactions.

So, these types of epoxides are formed and it turns out that these epoxides are cancerous. So, these are the actual mutagens. So, if you add these epoxides in bacteria, you can see mutations happening. So, that is why this molecule is called a pro-mutagen; initially, it is not a mutagen, but once it enters our body, it goes through the liver, it gets oxidized into these types of molecules, into these epoxides, which are mutagens. So, this type of compound can be found in tobacco smoke, coal tar, in some food types, and especially in grilled meat.

So, if you take any meat that is grilled or slightly burnt, then in those burnt portions, you will get this type of molecule. So, based on what we are discussing now, it seems that a major factor in cancer is mutation. So, in that sense, you can consider cancer as a genetic disease, and there is much evidence of that. So, this is how chromosomes look in a cancer cell. So, if you take a cancer cell and look at all the chromosomes, if you paint them with different colors using fluorescent dyes, you will see something like this.



Now, there are several abnormalities in these chromosomes. So, if you think about our chromosomes, we have 23 pairs. So, 1 to 22, and then the last one is the sex chromosome, so X and Y and we have two copies of each chromosome, one coming from each parent. But if you see here, there are four copies of chromosome 1, six copies of chromosome 2, and three copies of chromosome 3.

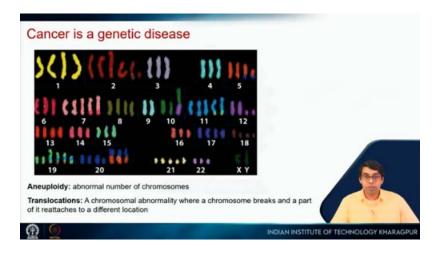
So, the chromosome number is completely random. It is not maintaining these two copies for each chromosome. This is something that is called an euploidy. So, an abnormal number of chromosomes and this happens in, and you will see this in many cancer cells.

So, it means that when the cell is dividing, mitosis happens, and somehow the mitotic machinery is not working properly. So, the number of chromosomes is increasing and this can create a problem. Suppose there is a particular protein, you need only two copies of that protein, but if you have more copies of that protein because there are more chromosomes, then there are more copies of the gene, so it will produce more proteins. So, the protein dosage increases, and that can upset the balance. So, this is one problem.

Another problem that you will see if you look at this chromosome number 20. So, there are many copies. Now, if you just look at these, there are two different colors. So, blue, which is actual chromosome 20, and there is also orange, which is most likely coming from chromosome 5. And again, in chromosome 5, you see there is orange and there is blue, which is coming from chromosome 20.

So, there is a problem here and this type of chromosomes are called translocations. So, this process is translocation. So, these are chromosomal abnormalities where a chromosome breaks and part of it reattaches to a different chromosome. So, you can imagine that there are two chromosomes. These are two different chromosomes. This chromosome breaks; the top part attaches here, and the bottom part attaches here. So now you have a primary chromosome where the top part is from one chromosome and the bottom part is from a different chromosome, and then, of course, the reverse will be there in the other chromosome.

Now, proteins which are at this junction region will also get snapped off. So, if there is a gene here in chromosome 5 and if there is a gene here in chromosome 20, those genes, part of those proteins, will get swapped. In many cases, this type of translocation causes abnormal signaling events and leads to cancer. So, here is sort of a summary of what I am discussing so far that this is a normal cell. It will accumulate mutations due to different events.



So, it takes in one mutation which might result in hyperplasia. So, the cell is growing more, and then it gains some more mutations which can most probably damage the DNA repair mechanism. It can gain some more mutations which can damage the programmed cell death or apoptosis. It can gain some more mutations and things like that.

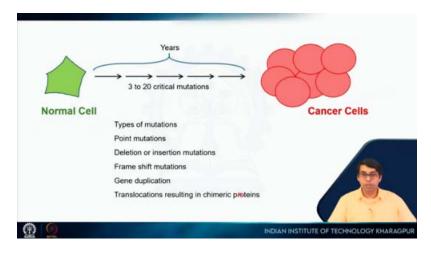
So, these mutations will accumulate to finally turn it into a cancerous cell. So, this whole process can take years and it turns out that there are 3 to 20 critical mutations which are needed. The least is 3, and it can be as many as 20 mutations in critical genes not in any normal random places. They have to be in certain critical genes, and we will see what types of genes those are those and that can lead to cancer.

So, what are these types of mutations? It can be a point mutation. Just one nucleotide changes to another nucleotide, which can change the characteristic of the amino acid, which can change the function of the protein. It can be a deletion, so a part of the gene is deleted, or it can be an insertion, where some other genomic part is inserted into this gene so that new amino acids are encoded. It can be a frameshift, so maybe one nucleotide is inserted, and the whole frame is shifted.

It can also be gene duplication. So, a particular gene, let us say, helps the stem cells to divide. Now, that particular gene is duplicated due to errors in replication. So, instead of one copy, you have maybe 3, 4, or 10 copies. So, then the number of proteins that will be produced will be more, and that can create a problem.

Translocations, as we saw in the previous slide, occur when different parts of the chromosomes get chopped up and attached to other chromosomes, and that can result in chimeric proteins. So, chimeric proteins mean there is a protein where one part of the protein comes from one gene and another part of the protein comes from another gene and

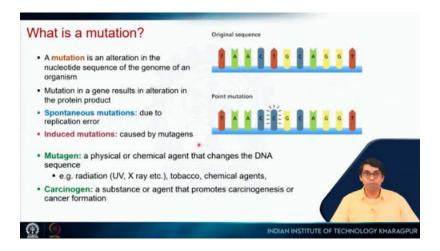
that can have a mixed function which is not present in any of the normal proteins. So, here is a more detailed description of a mutation. So, what do I mean by mutation?



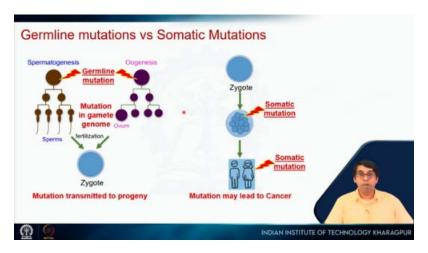
So, let us say this is the original sequence and you can see that these nucleotides remain the same. However, this T gets changed to C. Now, if this T gets changed to C, if this is my codon, which is CTG, that will get changed to CCG. So, it can code for a completely different amino acid.

In many cases, it will code for the same amino acid because of codon degeneracy, but in some cases, it can actually code for a different amino acid. And if that different amino acid is very different in characteristic, for example, let us say a hydrophobic residue gets converted to a charged residue, it can change the structure of the protein and the function of the protein. So, a mutation is an alteration in the nucleotide sequence of the genome of an organism, and this mutation can result in an altered gene product, which is the protein. Spontaneous mutations occur due to replication errors, and induced mutations are caused by mutagens. So, we saw an example of a mutagen.

So, what is a mutagen? It is a physical or chemical agent that changes the DNA sequence. So, it can cause this type of mutation. So, it can be UV, X-Ray radiation, it can be chemicals like benzopyrene, which can come from tobacco chewing or tobacco smoking. So, when these mutations cause cancer, it will be called carcinoma.



Now, we have seen that there are two types of cell division: one is mitosis, and one is meiosis. So, somatic cells undergo mitosis. So, in this case, if a mutation happens, then that can cause cancer, but it will be in the same person. But if this mutation happens during meiosis in the germline cells, like the sperm or the ovum, in that case, it will be passed on to the next generation. So, in this case, the mutation will be transmitted to the progeny, that is, the next generation.

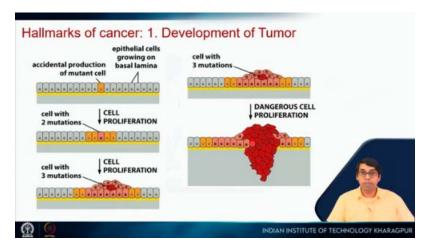


So, what are the hallmarks of cancer cells? There are these four major hallmarks. The first one is uncontrolled mitotic division. So, uncontrolled cell division is called tumorigenesis. The second one is evasion of cell death signals.

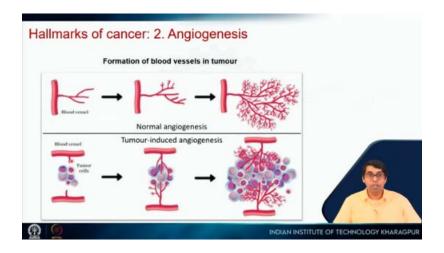
So, it avoids apoptosis, which is immortalization. So, these cells will keep on dividing, dividing, and dividing. Now, since the tumor is formed, it will need nutrition so, it will need the formation of blood vessels. So, these cells will induce the formation of blood vessels.



So, that they can get nutrients, oxygen, and all these other materials to divide faster, which is called angiogenesis. Finally, they will break open the tissue, damage the tissue, and invade into other places, which is metastasis. So, normal cells will not do that; they stop growing once they have reached a certain stage. So, they will not grow on top of each other but the cancer cells will do that, so they can invade tissue and move from one place to another. So, the first one is the development of a tumor. We have already seen this: there is a mutation that results in hyperplasia. These cells are growing.



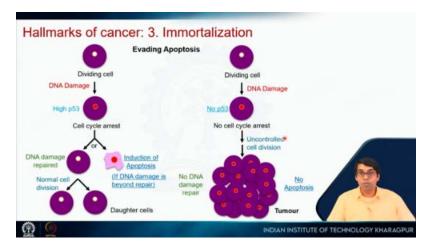
Then, more mutations accumulate and then, the tumor grows but still, this is localized. This is present in one place. The second one we saw is angiogenesis. So, they will induce the formation of these blood vessels. So, you have a tumor that is growing here. Now, there was this blood vessel, but since the tumor is growing, it needs more nutrients.



So, it will induce the formation of these blood vessels. The third one is immortalization. So, let us say this is a normal cell. So, it is a stem cell, let us say, it divides and while that division is happening, there is some form of DNA damage.

So, in that case, it will induce this DNA cell repair mechanism. So, one of the proteins that is involved in this is called p53. We will see p53 is also called the guardian of the genome. So, what it will do is it will first arrest the cell cycle. It will stop the cell cycle.

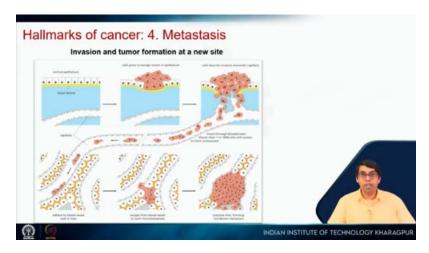
It will recruit the DNA damage repair proteins. So the DNA damage can be repaired and once that happens, it will allow the cell to proceed through the cell division. So daughter cells are formed. If it cannot repair the damage, the damage is such that it cannot be repaired, then it will recruit the proteins which induce apoptosis.



So, it will start the programmed cell death and result in the killing of the cell. So, this is what happens in a normal cell, this is happening every day in your body. But during cancer, this normal process does not happen. So, let us say DNA damage happens because of

intrinsic or extrinsic signals. Let us say there is a mutation in p53 or for some reason p53 is not recruited.

So, then there will be no cell cycle arrest. So, and also no recruitment of this DNA damage repair mechanism. So, there will be uncontrolled cell division and that can result in a tumor. And finally, there is metastasis, which means that this tumor grows and then it damages this tissue so that it breaks open and gets into the bloodstream, and then these cells travel, they go to some other place, and then they start growing a new tumor there, so here the tumor spreads throughout the organism. What causes cancer?

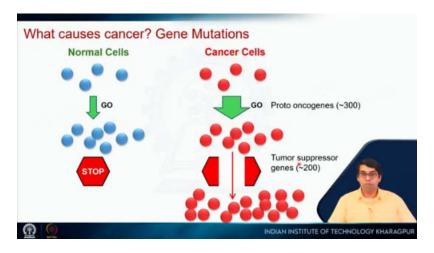


We saw that gene mutations cause cancer. Now, in normal cells, let us say these are stem cells, and we have seen examples that our body requires the formation of new cells, like we have seen in the case of red blood cells, we have seen in the case of the gut cells, epithelial cells. So, you need the formation of new cells, but this formation of new cells is highly regulated. So, when new cells are needed, there are signals which will say, 'OK, go divide.' So, they will start dividing. When the required number of cells are produced, then there are stop signals, and the cell division will stop.

However, in the case of cancer, this 'go' signal is hyperactive. So, they will keep on dividing, and this stop signal is also damaged, which means that they are not going to stop. So, they will keep on proliferating; they will keep on dividing. Now, the proteins which are involved in this step, which tell a cell that now it has to divide, there are more requirements; this type of protein is called proto-oncogenes. So, their normal function is to signal the division of a cell, and it turns out that there are 300 such proteins. When there are mutations in these proteins, they become oncogenes.

So, they can lead to cancer. There will be mutations in some of these proteins. Proteins which give this stop signal are called tumor suppressor genes. Now, if there are mutations in some of these tumor suppressor genes, then in those cells, of course, the cell division will not stop; there is no stop signal, and it turns out that there are almost 200 such genes. So, in total, there are 500 genes which are the critical genes, and mutations in some of these genes can lead to cancer.

So, I told you that there has to be some critical mutations. So, those critical mutations have to occur in these types of genes. So, what are proto-oncogenes? Their normal function is to regulate cell growth, to regulate cell division, and to regulate cellular communication. So, cells talk to each other through their extracellular matrix, and they can tell that there is some gap here, and more cells are needed.

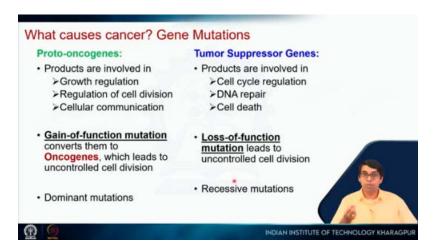


So, they will signal the stem cells that we need some more cells, so please divide. So, all of these things are regulated by this. However, if there is a mutation in them, so that they themselves become hyperactive. So, there is no signal that is telling them that, 'Okay, now tell the cell to divide,' or 'Now do not tell the cell to divide. So, if that signal is gone, and they are always telling the cell to divide, in that case, that is a gain of function and that results in uncontrolled cell division.

These types of mutations are dominant mutations. For example, if you think about this particular gene, there are two copies of this gene, one from each parent. Now, one is a normal copy, and the other one is the mutated copy. If there is a mutation in the mutated copy, then that mutation is enough to overcome the function of the normal protein or the normal gene. So, it will be a dominant mutation.

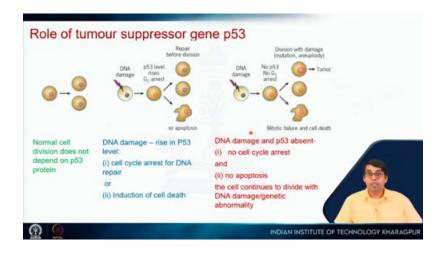
The other type of cells that we, proteins or genes that we saw, are the tumor suppressor genes. They are involved in cell cycle regulation, DNA repair, and cell death, which is apoptosis. So, if a mutation happens in this, then the stop signal is gone, and that will again lead to uncontrolled cell division. Now, these types of mutations are recessive mutations. So, it means that again you have two copies.

So, one copy is corrupted, but the other copy is still there. So, it will still give the stop signal, and the cell division will stop. So, mutations have to happen in both copies. So, that is why this is a recessive mutation. So, I talked about p53, which is a tumor suppressor gene.

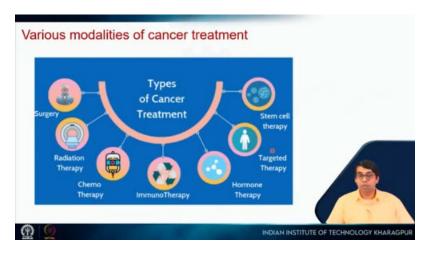


So, in normal cell division, it does not depend on the p53 protein because cell division is happening; there is no issue; everything is working fine. However, if there is DNA damage that happens, it can happen because of the replication process itself, or it can happen due to some chemical agent, UV radiation, or some other radiation. So, in that case, p53 levels will rise and the first thing they will do is stop the cell cycle. So, arrest the cell cycle, bring in the DNA damage repair proteins.

If all of these things work, then, of course, go through cell division. If that does not work, then it will induce cell death. If DNA damage happens and p53 is absent. So, for some reason, there is a mutation in p53, or p53 is not activated because of the proteins that recruit p53. Then, of course, none of these things will happen: no cell cycle arrest, no DNA repair, no apoptosis. The cell will continue to divide, and that can lead to tumor formation and eventually cancer.

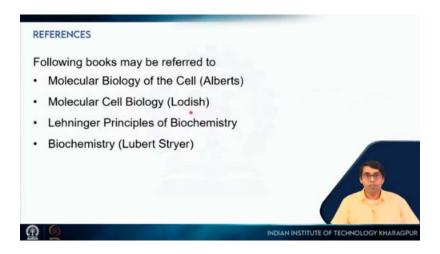


So, p53 is called the guardian of the genome because it protects the genome. It ensures that the correct DNA is present and that cell division happens correctly. So, there are all these types of cancer, and now, over the past 40 years, all sorts of therapies have been developed. So, I talked about surgery. There is also radiation therapy, where you use radiation to kill these cancer cells. There is chemotherapy, where you use medicine to target the cancer cells and kill them. We will talk about immunotherapy when we talk about the immune system. There are targeted therapies.



So, we saw some of that in the last lecture when I talked about stem cells, and then, of course, there is stem cell therapy. So, I talked about these two in the last lecture a little bit in the last slide.

So, these are the books you can refer to for this course.



Thank you.