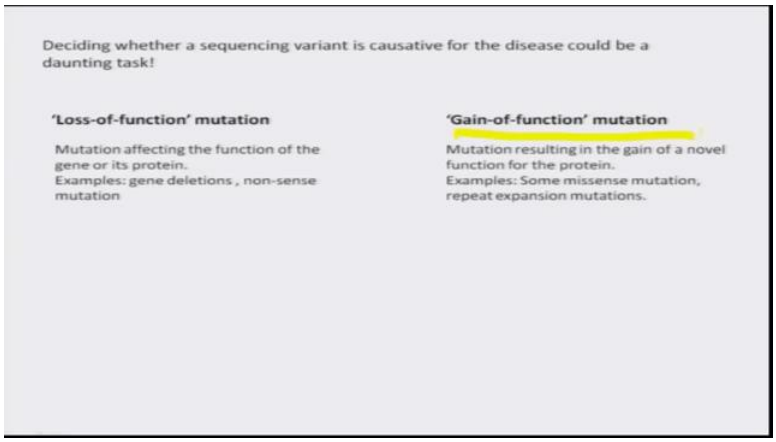
Human Molecular Genetics
Prof. S. Ganesh
Department of Biological Sciences and Bioengineering
Indian Institute of Technology, Kanpur

Module - 03
Lecture - 10
Mutation and instability of human DNA – Part-II
From pedigree to molecular pathology

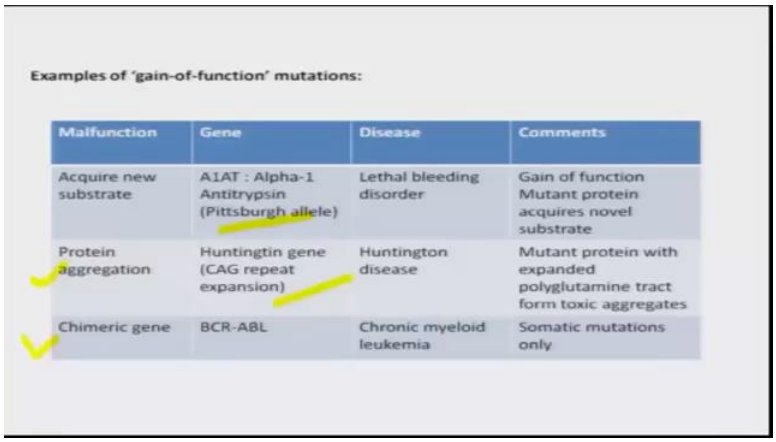
So, welcome to this second lecture of the third week of the course human molecular genetics. In the previous lecture, we looked into how mutations leads to, what do you know as loss of function defect, how a change in the DNA can result in a loss of the function of the gene or protein sequence. So, today we are going to discuss about what is known as a gain of function mutation. How change in the DNA give some novel function to the gene or the protein. Therefore, that event results in a disease.

(Refer Slide Time: 00:56)



So, as we have discussed, the gain of function mutations often results in gain of a novel function. The mutant version of the protein now gains a novel function and the novel function could be not good for the cell. As a result, you have the disease. For example, some missense mutation may lead to a kind of property, wherein the protein do not fold properly and they aggregate and these aggregates could be toxic. This is an example I had given.

(Refer Slide Time: 01:23)

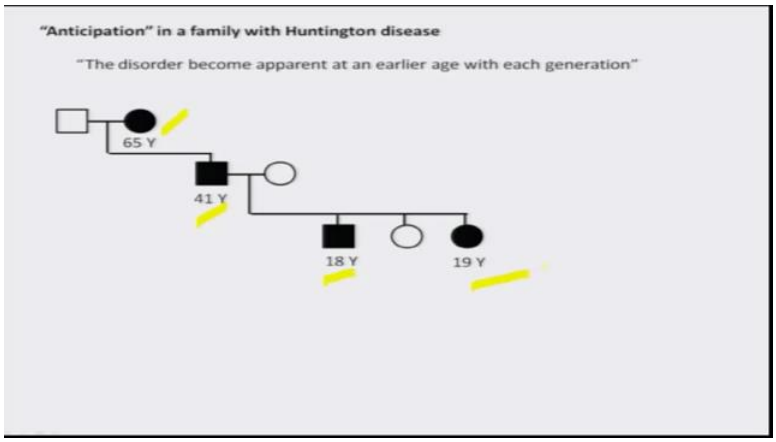


So, let us look into some of these, examples and how variations there can lead to so called gain of function. I have given you three examples. So, you are going to look into the second and third examples in detail. The first one is very interesting. So, this particular protein, called A1AT, which is Alpha-1 antitrypsin, is involved in a very severe disease called as lethal bleeding disorder. So, if you have a particular mutation of this particular gene, now that would result in a condition, wherein the individual will bleed to death, right and this mutation results in what is called as a gain of function. So, the mutant, now because of the change acquires a novel substrate which is not a substrate for this protein, if it is a wild type and as a result, when it acts up on the novel substrate that results in the heavy bleeding, right?

So, this is an example. But, what we are going to discuss is two other conditions which are very very unique and will give you an idea as to how unstable your DNA can be, because this particular Pittsburg allele, which results in the lethal bleeding disorder is a rare form. So, we are going to look into two examples that are relatively more common and happens again and again and we are going to look into why does it happen. So, let us first look into a condition called Huntington disease and this disease is caused by a gene which codes for a protein called Huntingtin and this happens, because of, you know, a change in the DNA resulting in expansion of certain unique sequences, repeats, right and as a result, these repeats are present in the coding sequence. As a result, the protein also acquire certain novel repeats of certain amino acids, leading to their aggregation in the cell, for example here in case of neuron and that could be toxic

to the neuron. The neuron die, as a result the individual also develop neurological conditions. The second one, the third one we will see little later.

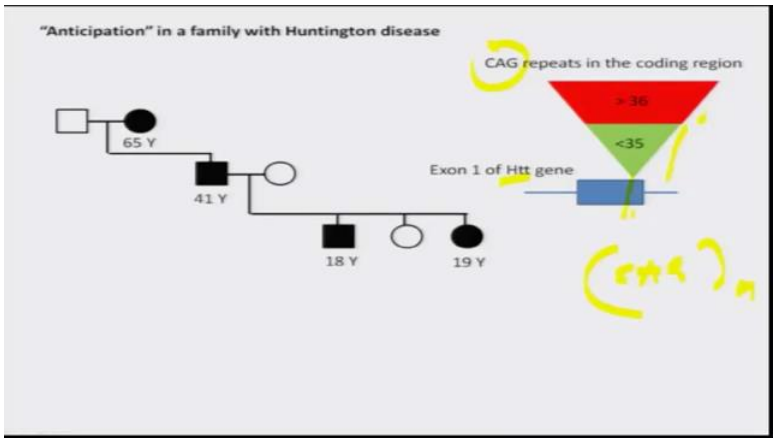
(Refer Slide Time: 03:45)



This is an example of a pedigree. So, what do you see? So, we discussed already that in the exceptions to Mendelian pattern, we looked at a condition called anticipation. Classically, the anticipation is characterized as a disorder, which, that becomes apparent at an earlier age with each generation. As shown here, this individual developed the disease that around 65 years. In her next generation, her son developed around 41 years and his son or daughter, you know, developed it much earlier, when he or she is around 20 or earlier, right and this is called as anticipation. For certain diseases, you can anticipate that in the next generation the disease onset is going to be earlier than what you have seen in the current generation. As a result, they will have more severe effect, symptoms, right?

Now, why should, a disease invariably express such kind of anticipation, meaning should become earlier at onset, you know, every generation why should it be earlier So, that is, it remained a paradox for a long time, until they identified the gene that codes for this particular protein which is involved in the disease called Huntington.

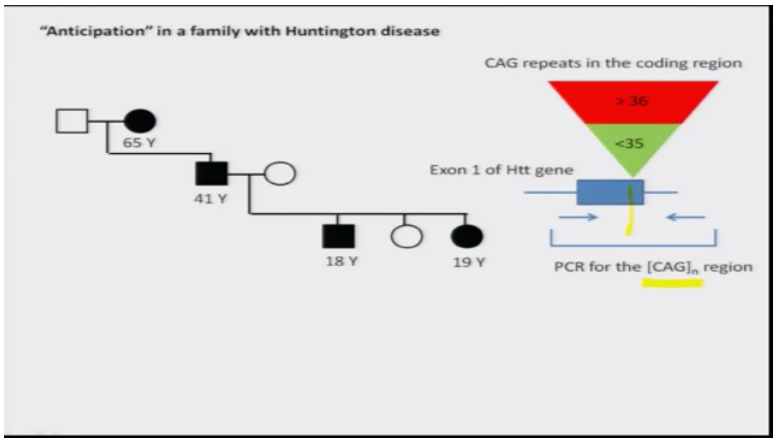
(Refer Slide Time: 05:07)



So, this is what the schematic of the gene, which is, short form is Htt or Huntington, Huntingtin. So, what is interesting here is that this particular gene in its first exon, it has certain repeats. These are CAG repeats and this CAG repeats are polymorphic, meaning if, if you look into the normal population, the individual could, some of us could have 12 repeats, other person could have 20, other person could have 12, 13, 15. So, it varies. Majority of the normal individuals, the number of repeats that you have in this particular site could be less than 35. So, generally it is denoted as, you could, that is CAG and then you put it as "n", "n" which is, could be 12, 13 whatever it is. But, these are continuous repeats;CAG, CAG, CAG and so on.

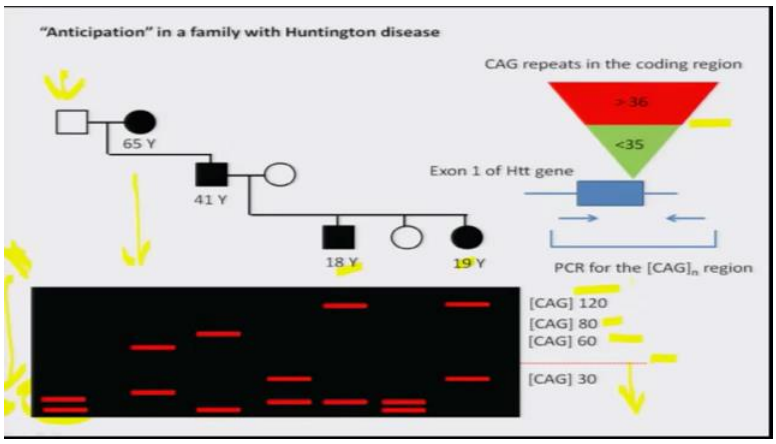
But as and when the repeat number exceeds 35, it becomes 36 plus, then individuals show symptoms. So, what is shown here in red color is pathogenic allele or the mutant allele, allele that becomes pathogenic meaning causing the defect. When it is less than 35, it is normal; all of us have less than 35 number, right? So, why does it happen?

(Refer Slide Time: 06:26)



So, we can really search that, but just before getting into that, let us look into how do you really use this information to tell whether in this family that have been shown here, whether there is a correlation, right? So, what you can do? You can, this is where the CAG repeats are; I can design a primer and you can do a PCR. The primer is flanking this CAG repeat region and I can amplify, right? So, when I do a PCR, this is going to amplify the region that has got CAG repeats. So, if I have a larger repeats of CAG, my PCR product would be larger. If I have fewer repeats of CAG, my PCR product will be smaller. So, I can do that. So, let us see how does it do?

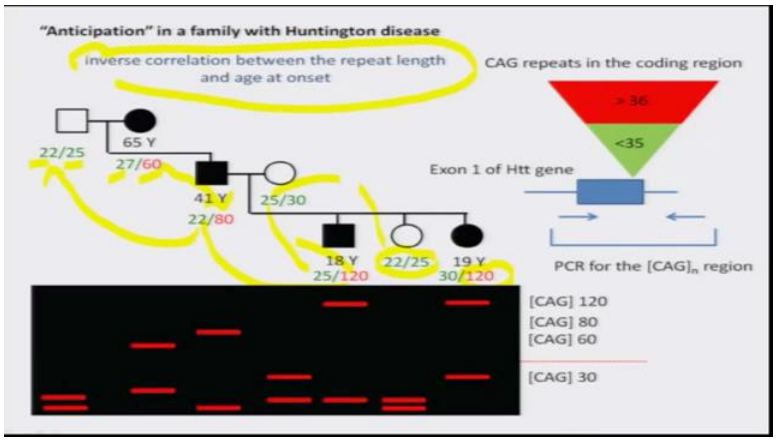
(Refer Slide Time: 07:08)



So, this is a gel. I have done a PCR. I have separated the fragments in agarose gel and this is how it separated. So, we are going to have high molecular weight fragments on the top and low molecular weight fragments on the bottom, because they migrate faster. What do you see? I see

that this individual who is asymptomatic, he has got two repeats, two alleles, two bands representing two alleles, one that he got from his father, other one is from mother, which both of them are below 35. So, this is the red line that shows the danger zone that is denoted here. So, if you have above 35, the repeat number is more than 35, you are going to have the disease and the mother, this lady,you see that there is, one of her allele is below 35, the other one is above that is about 60 repeats. What happened to her son? Her son have a much higher molecular weight repeat and that could be probably 80 and then in the next generation, both of them, these two individuals have much larger repeat that is 120.

(Refer Slide Time: 08:21)

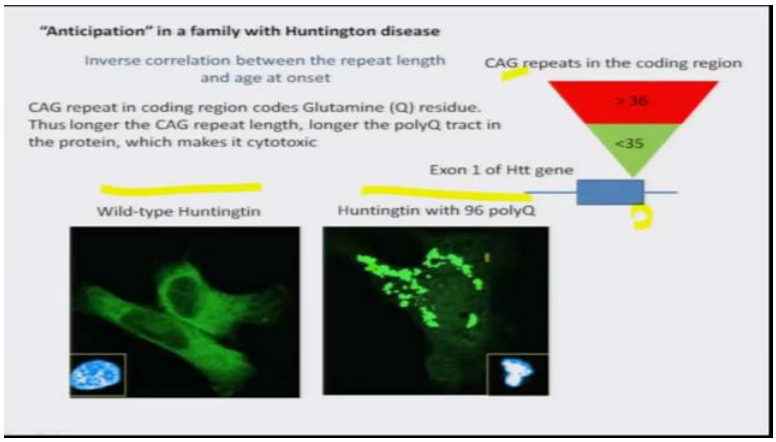


So, if you plot it, this is how it looks like. Here the red numbers denote the disease associated allele, the green one represent the wild, wild type. As you can see here, 22 and 25, these are the unaffected individuals. This lady carry 27, which is a normal allele, 60 which is affected and this 22 has come from father, the 60 had come from mother. Obviously, 60 did not stay as 60, it expanded, the number increased and in the next generation, we can see the 25 had come from mother, but the 80, you know, gone up to 120, whereas this individual was lucky, she got the 25 from mother, 22 from father, nothing happened, stable. But again, you see that 120, this allele that is 80 that when it came from her father, it became 120.

So, the moment the repeat length exceeds 35, it becomes unstable. In every subsequent generation it increases, expands. That is why it is called as dynamic mutation and such mutation, you can see.. There is an inverse correlation between the length of the repeat and the age at

onset; larger the repeat, earlier would be the onset of the disease.

(Refer Slide Time: 09:40)



The question is, why? Why should be the disease appear much earlier if you have a larger repeats? What is interesting is that, as I told you, this CAG repeat is present in the exonic region of the G, meaning they are coding for a protein. So, it so happened, the CAG itself is a codon. So, when you have 25 or 30 or 12 such repeats, so your protein will, also will have as many glutamine repeats in your protein. So, if it exceeds 35, it becomes 60, 80, 90, we are going to have as many glutamine in that particular protein. So, it is affecting the way the protein is being made, as to how many glutamines that is there. So, what happens is that, what is known is that, when the glutamine number, the repeat, if it is below 35, it really does not affect the protein functions. We have many such proteins that have got, you know, 15, 20, 30 glutamine repeats.

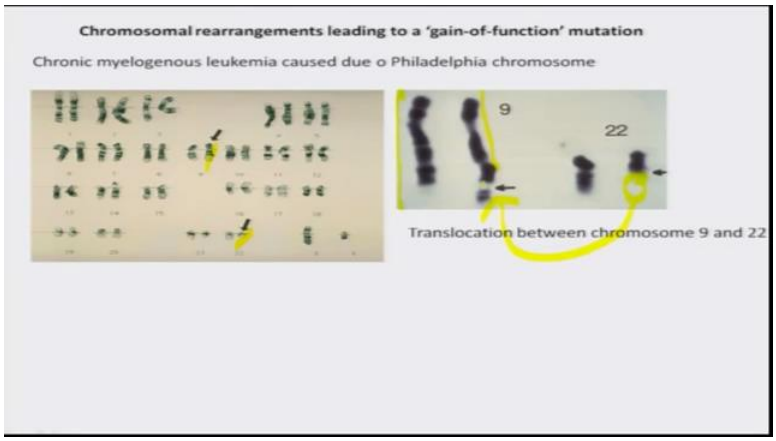
But, when it exceeds 35, 40, now the same repeat, when it is in, in longer form can form the protein to aggregates. This is what is shown here. This is a cell, human cells, expressing either the wild type Huntingtin, where you have a normal repeat range, which is, could be, 25 or so and you have a other cell in which have 96 glutamine repeat. You can see how this signal is seen. So, in this on the left side you see that there is a green signal all over the cell that is inside of cytoplasm. So, that is the normal distribution of this. But on the other hand, on the right side, you see that there are large aggregates that are seen and these aggregates can be toxic to the cell. So, that is what we believe that when such protein having longer glutamine expressed in the neurons and that could lead to the death of the neuron and as a result, these individuals developed what

do you call as ataxia, chorea. This is one of the phenotype of Huntington disease. This is a neurodegenerative disease. To begin with, the individual cannot walk and he will have every other symptoms that are normally associated with neurodegenerative conditions.

So, why should, you know, repeats expand, is a mystery still. So, there arehypothesis, which either supports that it could be because of like what you have discussed in non- allelic recombination. Because they are repeat containing alleles, they may align identically in a way expected or it could be misalign. As a result, you could have one allele having a larger repeat, the other one is having a shorter repeat or there could be, you know, there is a defect in the replication process. When the DNA replication takes place, somehow, you know, one of the two strands gain more repeat units and so on. So, these are hypothesis. If you are interested you can go and read more. But what is, you know, established, that such repeats, when transmitted through generation, meiotic process, invariably expand, results in what do you call as anticipation.

So, you can really predict as to what would be the age of onset, if you are able to look into the number of repeats for a given gene for a given disease. So, that is a gain of function. So, whatever protein function it is supposed to do, because of glutamine repeats, it gains a novel function that is toxicity. Now, you are going to look at one more such bad guy, meaning a mutation, gain of function mutation giving a novel function to your protein, again resulting in a disease condition..

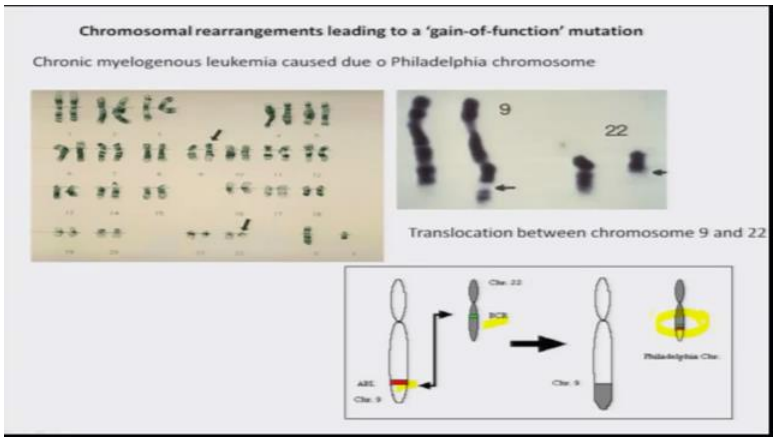
(Refer Slide Time: 13:24)



This really, here in this case it is does not happen, because of certain repeat element. It happens because of some instability between two chromosomes, right? So, what we are going to look at is a condition called as Philadelphia chromosome, a particular form of chromosomal translocation that we discussed in the first week of lecture, where we are talking about structural abnormality. That is a recurrent event; more often you will find in many individuals a translocation between chromosome, involving chromosome 9 and chromosome 22 and this was first identified in the city of Philadelphia by a geneticist and since then it has been named as Philadelphia chromosome that is fusion between 9 and 22. So, that is, you know, shown here.

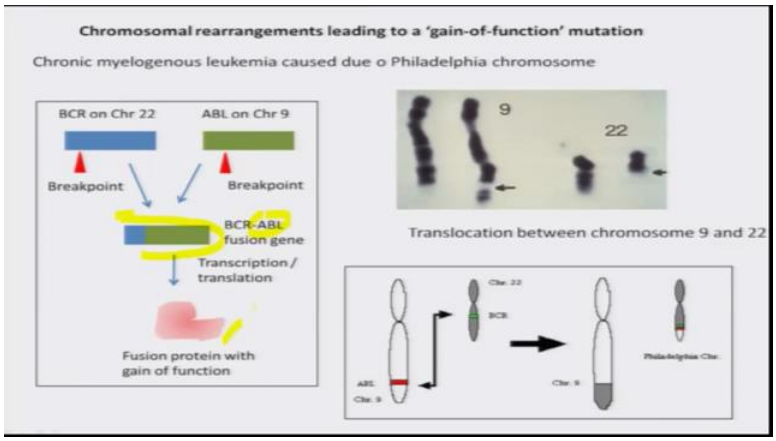
We can see that these are classic plates, karyotype prepared then, wherein they are able to show that this is a normal chromosome 9 and you have a chromosome 9 in which part of chromosome 22 came and translocated to chromosome 9. So, when such kind of translocation happens, how it can lead to the so called gain of function mutation?

(Refer Slide Time: 14:34)



This is what shown here in this schematic. You have in chromosome 9, a gene called ABL and chromosome 22, you have another , gene called BCR. So, when there is translocation, what happens? This translocation happens such a way that these two, the coding sequence of these two genes are fused together, right, like what you see here. So, now one of the gene comes under the control of the other gene. So, maybe a one gene should not be expressed in a tissue. Because of the fusion, now what happens, it starts expressing and exactly that is what happens in this case, leading to a very very severe form of cancer called as leukemia or the white blood cell cancer, where this is a recurrent event. It can happen; this is a somatic mutation. It does not get transmitted from one individual to other, it happens in a somatic cell. If it happens, then you would have the condition called leukemia, right?

(Refer Slide Time: 15:36)

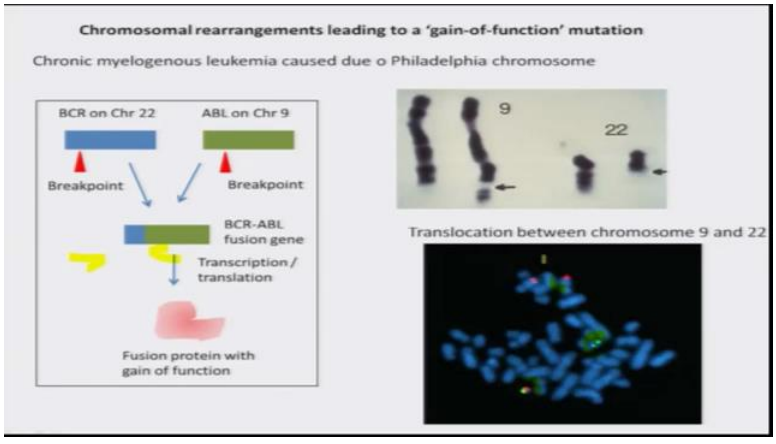


So, how does it happen? So, this is what shown on the schematic on the left side. So, you have

the BCR gene on chromosome 22 and ABL gene on the chromosome 9. So, when the break point takes place, again I told you that these are, the break points are very precise; always it happens over there. So, again there are some structural elements that are present in the chromosome, resulting in recurrent events. So, you may have an individual in which it happens. It is a somatic cell, but it could happen again in another individual. So, it is not transmitted. It happens again and again. So, it is a recurrent event and happens because of certain structural features of the chromosome. It happens between chromosome 9 and 22. But, when it does happen, what happens, you have gene, you know, which has got the fusion, meaning the BCR is, is driving the ABL gene.

So, when the protein is being coded, you are going to have part of a BCR and large part of ABL together is expressing and this protein, you know, has got a function. It is a gain of function. Normally it should not express, now it expresses because of the fusion and when it expresses, it drives the cell to divide and divide and form cancer, right? So, the translocation results in the cancer. So, that is how people diagnose you. So, for example if somebody is having this leukemia and they undergo a treatment, chemotherapy and then people look into, after the chemotherapy is over, whether they still carry, the cell that has this kind of translocation, right, structural changes. If they do not see, then they would believe that this individual is alright. But, even if few cells show such kind of, you know, changes, then you would tell that the, the treatment, therapy is not yet effective. So this is what shown here.

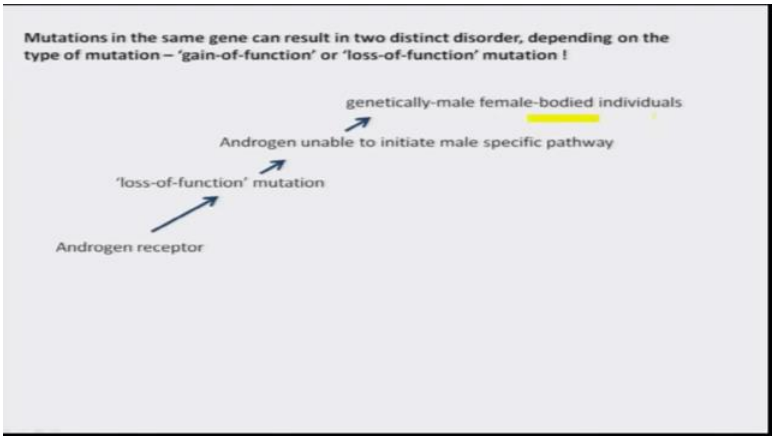
(Refer Slide Time: 17:34)



So, you can culture cells and either you can go and do a PCR to identify such kind of fusion. For

example, you can design a primer on either side of it and try to amplify. If still there are cells that are, you know, floating in your body, then if you extract DNA from a good number of cells and do a PCR, if you can get a product that means still you have fusion gene or you culture the cell and do what is called as FISH or fluorescence in situ hybridization, which would tell you whether the fusion chromosome is there. You can seen here; you have red and green here, which denotes that 9 and chromosome 22 came together for the translocation and these are normal chromosomes, right. So, this is a way you use this molecular biology approach to even diagnose or to understand the disease process, right.So this is one of the ways by which you do.

(Refer Slide Time: 18:30)



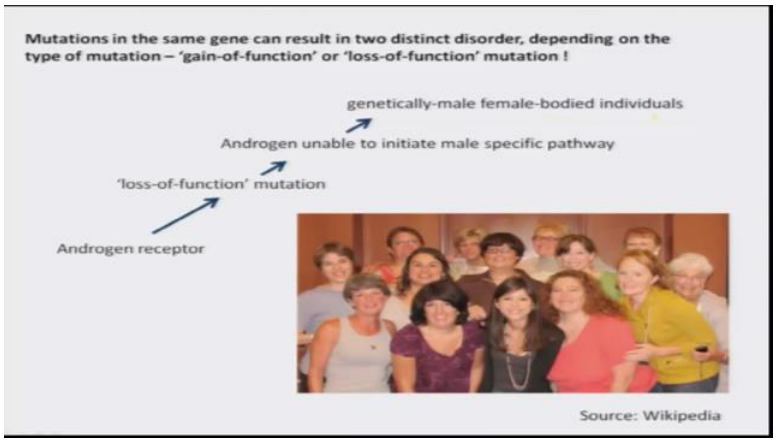
So, that are the two example that we discussed about gain of function. How mutation can offer a novel function to the gene product which is not good. It is not good for the organism, not good for the cell, as a result you have the disease, right. So, but there are conditions that are very unique. At times, the same gene depending on whether you have, what kind of mutation, whether it is a gain of function mutation or a loss of function mutation, you could have entirely two different kind of conditions, disease. I am going to give one example. This example is of androgen receptor. So, what is androgen receptor? Is a transcriptional factor; it goes and binds to certain DNA elements and activate the gene.

When would it be active? Whenever there is androgen, the sex hormone, the male sex hormone, if it is there in the body it will go and bind to this androgen receptor and make it active. Therefore, it will go and transcribe all the target genes. But, you could have a mutation, which is

loss of function mutation, meaning, for example it is a transcriptional factor, it should have a protein domain called DNA binding domain. So, it helps the protein to go and bind to the DNA. If there are changes in that coding region, so you may have, a protein will not be binding to the DNA. Therefore, the cell would assume there is no androgen, right. So, that is a loss of function.

So, when there is no androgen receptor, even if the body secrets the androgen, now the cell, the body would assume that there is no androgen, because there is no effect of androgen, because it needs the receptor to be active to go and transcribe and the androgen is very very critical for male development. Thedifference that you see between male and female, the voice, the growth of the hair, the body, how it is built and many such, even behaviour are regulated by the hormones. So, in an embryo, a child is a genetically male and if he is having a, receptor, you know, mutation in the androgen receptor, then the child would develop as such there is no androgen. So, the body would develop as a female and that is called as, genetically male female bodied individual. These are also called as testicular feminisation syndrome. They may have a testis which is undescended inside the body.

(Refer Slide Time: 21:06)

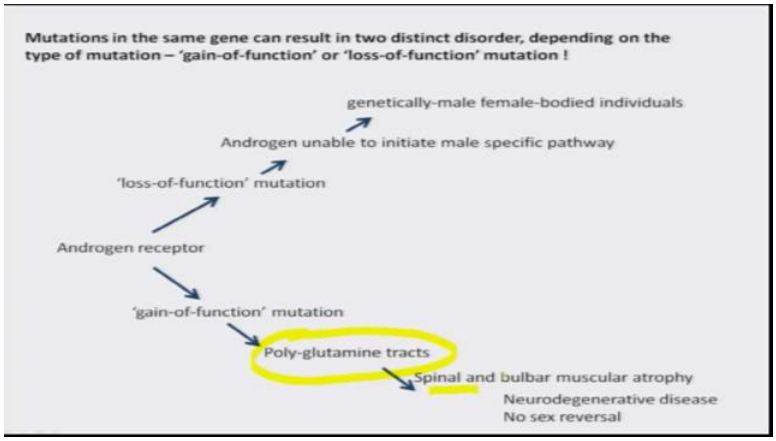


Otherwise they are, you know,, they develop like female, they feel like female and they behave like female and I am just giving one such image from Wikipedia and this is a group of individuals, all of them are genetically males, they are XY. But, they have a defect in their gene that codes for androgen receptor and all of them are, you know, sex reversed, meaning phenotypically they are female; they grew up as female, they feel like female and this is an

example of what is called as loss of function mutation, right?

So, this is interesting aspect, because these are the group of individuals who came together and said we have a mutation, but it is not our fault. Why we should not show our face? Because, often when you go and, and look into many of human genetic literatures, you would have photographs displaying a phenotype of individuals, but their face will be darkened. This is one of the ethics, because ethical consideration, because you do not want to reveal the identity, it may be, you know, a social taboo. So, but these individuals said that I do not want that kind of label, because I am fine. Whatever I have is not because of my fault, I am what I am. So, that is why you have this photograph. But, the same gene could also have what is called as a gain of function mutation. For example, it can have repeats.

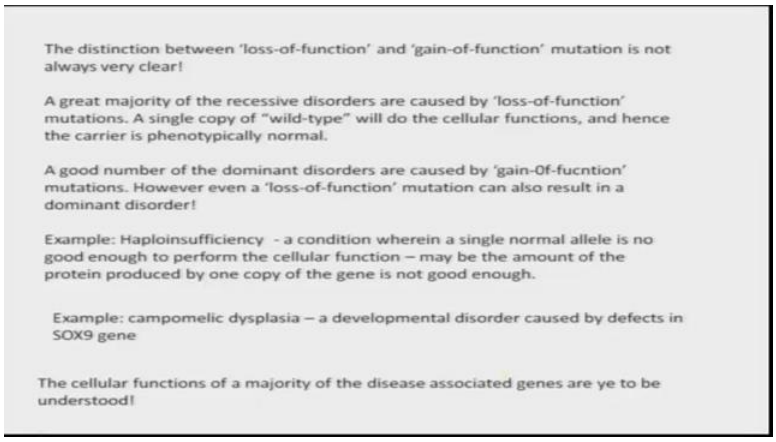
(Refer Slide Time: 22:39)



Exactly the way we discussed for Huntington disease, these gene which codes for the androgen receptor also has got, you know, CAG tracts, repeats and vary in the population, but when it exceeds 30, 40, you know residues, repeats, then it could gain a novel function, which is toxicity. Now, the protein becomes toxic. As a result, you may have neurodegeneration. Especially here in this case, this disease is called as spinal and bulbar muscular atrophy; you do not need to worry about the disease name, but what is,that it is a neurodegenerative disease. Mainly it affects the motor neuron, the neuron that help you to coordinate your body while walking or standing up or whatever,, all the functions and they lose their function, because of a gain of function mutation.

But, an individual who has got this disease that is spinal bulbar muscular atrophy or SBMA, they are not like the other group that is they are genetically male, but phenotypically female. That is not the condition here. They are genetically male, they are phenotypically male. But, they still could be having the other condition that is the neurodegeneration, because of gain of function, the polyglutamine, the toxicity. But, these proteins did not lose its original function that is binding to androgen and activating a male specific gene expression. So, that was alright. So, you know, you can now see, a same gene depending on how the mutation affects the proteins function can have altogether very, very distinct phenotype, right? That is an example.

(Refer Slide Time: 24:07)



So, when you have an example, when you discuss it looks so nice, but it is not the case for every gene and, and every mutation. So, the distinction between loss of function and gain of function is not always as clear as example that we have given; androgen receptor, testicular feminisation or SBMA. It is much more, you know, a grey zone. So, one thing, of course we can make a statement. That is a majority of the recessive disorder like what you have seen, Lafora disease or many other disease that you have discussed are caused by loss of function mutation, because you have deletions, you have nonsense mutations, missense mutations. So here, it is often you can sort of infer that recessive disorder could be because of a loss of function mutations. Because, if you are a carrier, you have one copy which is defective, lost its function, but you have other copy which is wild type, still doing a function, therefore you are normal. That is good enough. Even half the amount of, you know, protein of a given gene is good enough for you to do, do a normal function. So, therefore you are carrier, still phenotypically normal.

But, if you look into the other aspect that is the gain of function defect, a good number of the dominant disorders are caused by gain of function defect, like you said, Huntington disease or SBMA in case of androgen receptor and you can go on discussing many, many. But what is interesting here is that the dominant disorder can also be caused by loss of function mutation. For example, a deletion; a deletion of a gene in heterozygous condition, meaning of the two copies one is deleted, other one is intact, but can still result in dominant phenotype. Why? Because the, half the amount of protein that is made by the wild type copy is not good enough for doing a normal function, right? So, that condition wherein one wild type copy is unable to provide you the normal function is called as haplo insufficiency; haplo meaning half, the haploid that is what you refer to and insufficiency is that that it is not good enough, right?

So, there are many conditions. One of them is, we also discussed earlier, called as campomelic dysplasia, a disorder wherein, the growth of the bone is also affected and that is caused by defect in SOX 9 gene. So, this gene again codes for a transcription factor and if you have a mutation, in which even a deletion, the gene is not expressed, even then just one copy wild type, other one is deleted, still you would have the disease, because the half the amount of protein is not good enough for a normal biological function, right? So, that's where it becomes difficult. You cannot make a statement that all the dominant disorders are because of gain of functions, need not be, right?

The more challenge is that for majority of the genes that have been identified to be associated with many disease, we do not know what is the functions of the protein. You may know it is a transcription factor, you may know it is an enzyme, you know it is a receptor protein and so on, but still what is the function of that protein at the cell level, at the tissue level, at the organism level that is yet to be understood. It takes many many years to understand the functions of the protein. Therefore, even to model them as to how a particular mutation, you know, affected the function, it is going to be challenging question, because I do not know what is the function of the normal protein.

So, how would I even understand what is the function of a defective protein? So, that is, that is a

challenge and the many many labs have been working on it; possible that some of you who are listening to this course may eventually become a leading scientist, may solve some of this mystery. That is, you know kind of brings end to the second lecture of this week, wherein we discussed how the so called gain of function mutations, meaning a mutation changing the function of the protein, as a result that protein has acquired a novel function and that novel function could be harmful to that individual, as a result you have the disease and you also looked into how a same gene results in either a gain of function defect or a loss of function defect and each one can lead to independent disorders and some exceptions and so on. So, with that, we come to an end to the second lecture of the third week and we will see you in the next lecture.