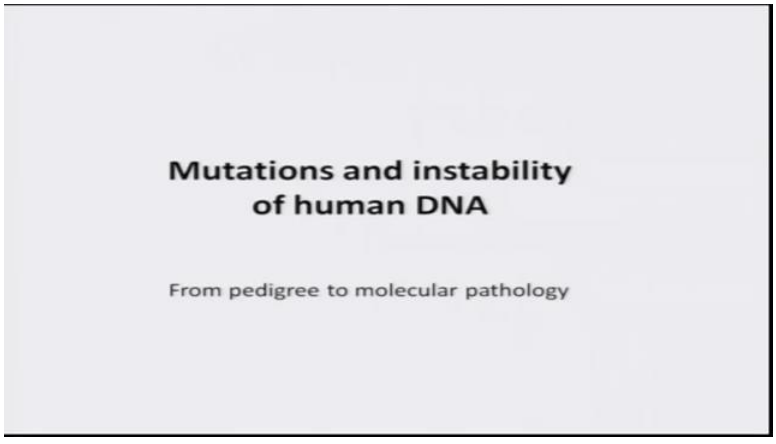
Human Molecular Genetics
Prof. S. Ganesh
Department of Biological Sciences and Bioengineering
Indian Institute of Technology, Kanpur

Module – 03
Lecture – 9
Mutations and instability of human DNA – Part I
From pedigree to molecular pathology

Welcome back to the NPTEL course, human molecular genetics. This is the third week; this is going to be the first lecture in the third week. In the last two weeks, we have looked into the central dogma; how the information is processed within our biological system, then we try to understand what is the gene and how the mutations do happen. Then, we looked into the pedigree as to how you look into the phenotype and try to trace as to whether the phenotype is dominant, recessive, X-linked, Y-linked and how you would use that information for predicting, the future generation whether they are at risk of having a given phenotype or not and then, we moved on quickly to some of the molecular biology techniques that we used to identify the gene, characterize the mutations and look into the changes in the DNA. So, with this background, we are going to move into a new concept that is, we are going to look into how changes in the DNA can alter the gene function? How does it really, the sequence difference in your DNA can modulate the way the gene functions, therefore you end up having phenotype.

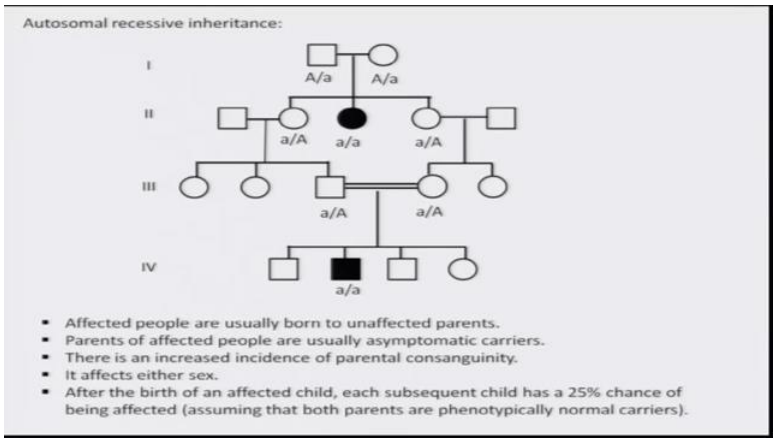
(Refer Slide Time: 01:38)



So, that's the title of this particular week that is"mutation and instability of human DNA". In this 3-4 lectures, this week we are going to look into the various different types of mutation that we have, why do we have the mutations, how it affects the gene function and finally what kind of methods we use to characterize how a defect in the DNA, bring about some changes in the protein and cell and so on. So, we are going to, towards the end of the, 4 lectures of this week, we are going to look into, also how you will use model systems to model human disease, disorders, genetic disorders. So, that's the focus of this week. So, let's start with the concept that is from pedigree to molecular pathology. So, that's the theme for this week's lecture.

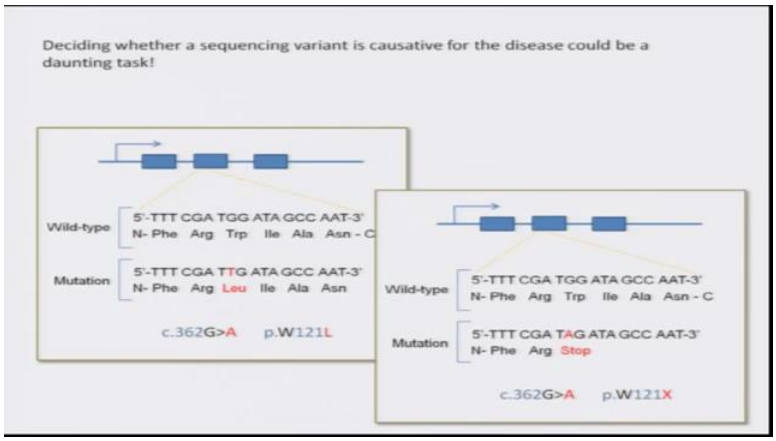
So, this is something a recap. So, I have discussed with you the previous week as to how you can use the information in a family that we are able to put together in a pedigree by talking to the family members and with that information you are able to tell whether a given phenotype or the disease, if you are looking at a disease, whether it is a dominant character or a recessive character and what would be the genotype you sort of predict.

(Refer Slide Time: 02:59)



So, here what we are giving is, again we are putting back the recessive disease in a family, something that we already discussed.So, it is when you look into this analysis, you really do not know where is the gene and what kind of changes that are present there. So, it is much easier. But, if you look into mutations, at times it becomes difficult.

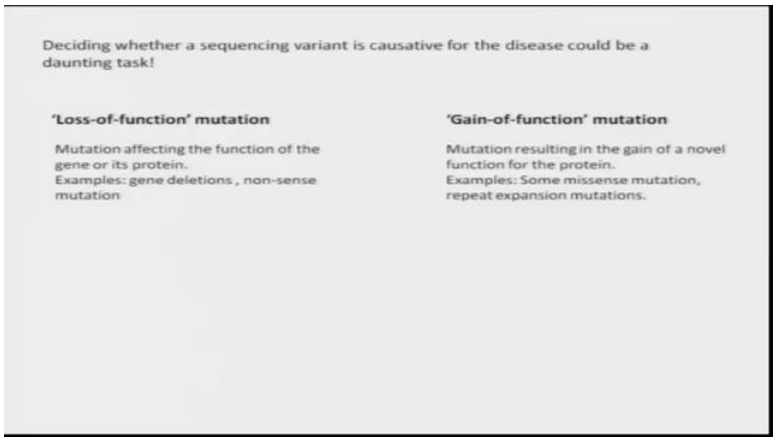
(Refer Slide Time: 03:23)



For example, I am showing here again one of the earlier mutations, that arehypothetical genes we are showing here, a nonsense mutation wherein, a codon which codes for a amino acid is converted into a stop codon. Therefore, the protein will not be made, because there is a premature termination. By looking into this mutationyou can say here at least what kind of effect you have. But, look at on your left side. What is the mutation? It is a, the samecodon, you could have different change in the nucleotide base, which results in altogether a different amino acid.

Now, by looking into these two variants which are likely that you could see in any disease, would it be easy for you to say whether which one of the two variant is causative for the disease, because you see a change that is there in an affected individual, but that does not mean that change is the one that is causing the disease. So, one has to really look into the function and how a change affects that function. That is very important. So, that is a verydemanding task. It takes years to understand what is the function of a gene and second, how a change in that particular gene, the change that you have seen in a family is the one that is causing the disease. So, it is extremelytough question to ask.

(Refer Slide Time: 04:54)



So, if you look into the type of mutation that you see in genetic disorders, they can be broadly categorized into two groups. These are the two large groups. One group of mutation you call as loss-of-function mutation, which is very self-explanatory, meaning that the change that you see in a gene, the DNA segment of a gene, results in the loss of the genes function. So, the mutation affecting the function of the gene or its protein; either it may affect, the gene itself may not be expressed, RNA may not be made or it does make RNA, but it does not make protein. There are many examples for this. For example, you could have deletions; an entire gene is deleted, so obviouslythe gene has lost its function or it could be the first exon which is deleted, the rest of the regions of the gene exists, but still it is useless, because most often you have the first exon carrying the promoter element, so if you have a deletion of first exon, the gene may not express or one particular exon is deleted, so, it probably has got a coding sequence for some very important functional domain of the protein, so you loose that coding region, so, obviously even if the protein is made, it is not going to function. Therefore these aresome examples. The other subtle change that can again lead to what you call as loss-of-function is nonsense mutation that is just now we discussed in the previous slide. So, you have a change. It could be as simple as one nucleotide change that modifies a codon into a stop codon. So, it will not make any protein; even if the RNA is made, it will not make any protein.

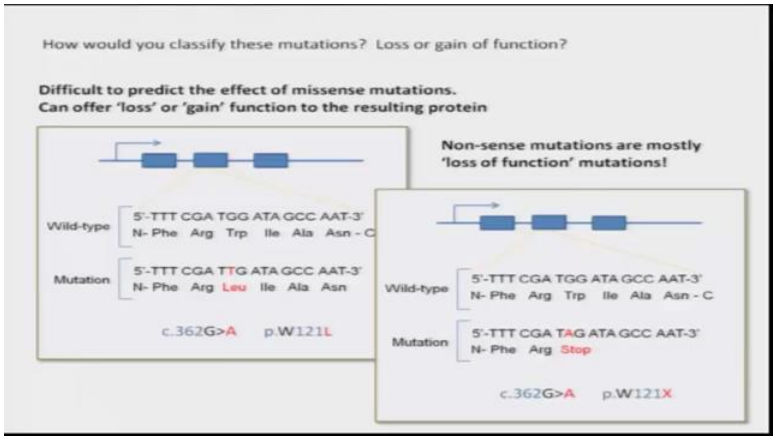
The other group that you are looking at is called the **gain-of-function mutation**. It is the most challenging group of mutations, becauseyou have to understand the gene function, you have to characterize every mutant allele that you see, only then you will be able to “infer” that it is a gain-of-function mutation. What do you meanby gain-of-function mutation? Here the mutation,

the change that you see in the **DNA could result in the gain of a novel function** for the protein. So, because of the variation, the protein acquires a novel function. So, this is going to do something more than what it is expected to do. So, that is called a gain-of-function.

Some of the examples are missense mutation.It is not necessary that every missense mutation would have a gain-of-function effect, but in some it does. We will see some of the examples. For example, missense mutation could alter the way the protein forms, its structure and that could be an abnormal structure which may make the protein to aggregate with each other and make what you call insoluble aggregation in the cell, which may result in toxicity to the cell. So, this is just one base change in the DNA, can lead to such kind of toxic form of protein, which could be detrimental for thatcell.

So, in several conditions and also you have anew group of mutation that we knew, about in last 10, 12, 15 years. It is called as a dynamic mutation. So, we will be talking about them little later in this lecture. As you can guess from the slide, it is called as the repeat expansion. You have certain repeat and these repeats when inherited from one generation to the other, that is they are transmitted from one generation to the other, and the repeats expands in successive generations. As a result, it can change the way they function, the gene functions and that could be again a function, so you will see that.

(Refer Slide Time: 04:54)

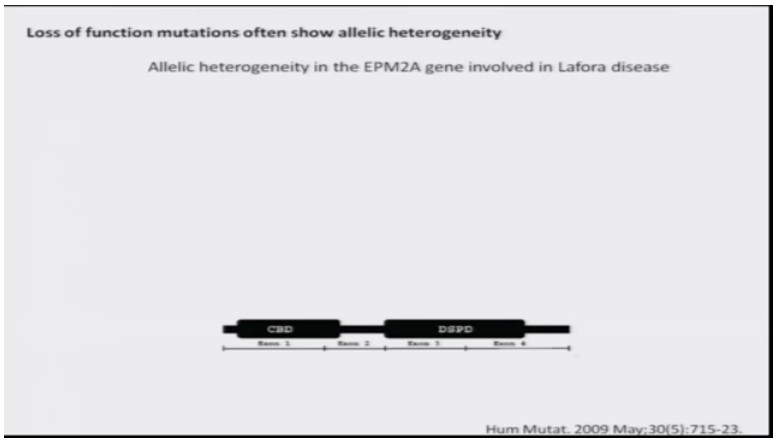


So, we are going back to these two examples that we discussed earlier, one missense mutation and one nonsense mutation; the missense mutation on the left side and the nonsense mutation on

the right side. So, by looking into we will be able to tell, for example the nonsense mutation is likely to be a loss of function mutation, because it is going to truncate the peptide, polypeptide, so you will have a shorter peptide, than normally the gene should make and this may be nonfunctional. So, it is loss-of-function and if you look into the other example, which is called as missense mutation, wherein you have a change in amino acid, it could either be your loss-of-function mutation, we have changed the residue, therefore the protein is no longer able to function as it should.So, it is a loss-of-function.

But, sometimes it can acquire a new function. So, it also can be a gain-of-function mutation. So, by looking into change in the amino acid, you will not be able to tell whether it is a gain or loss of function or it could be silent, meaning, the change in amino acid may not really impact the protein, the way it function. Therefore, still that could be present in the normal individual and if you happen to see an individual who is having a disease, you cannot call that as a mutation. It could be a normalvariant that is present in the normal population. So, it is extremely difficult toassign whether it is a gain-of-function or loss-of-function, if the changes are at the protein level, where one amino acid replaces the other.

(Refer Slide Time: 09:59)

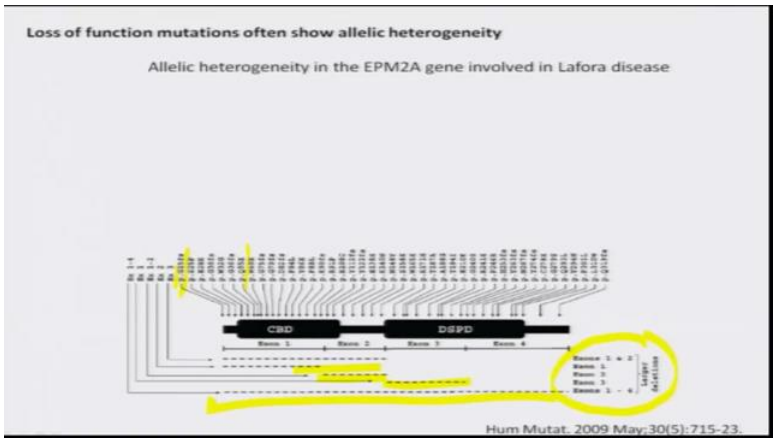


So, let us look into one class that is loss of function mutation. Firstly, let us look into what are the, some of the signature that you see in loss of function mutations and then we move on to go and discuss the gain-of-function. Here, I am going to describe one particular gene, which is named as EPM2A causing a disease called lafora disease that our lab has been working for long

time. So, one of the classic signatures of a loss-of-function mutation resulting in a disease is that the mutation show what is called as allelic heterogeneity. What does it mean? The same gene would have a large number of different mutations, hundreds of different mutations. Some of them could be deletions, some of them could be frame shifts, some of them could be insertions, some of them could be missense mutations and you have clinically very similar phenotype, therefore you can categorize them as loss-of-function mutation, because not that any mutation and every mutation cangive a proteina novel function. So, it cannot be a novel gain-of-function mutation. So, this is a signature, a gene defect there in results in a recessive disease, most likelyshould have a loss-of-function mutation and they show large number of allelic heterogeneity, large number of alleles.

Let us look into this example of this gene EPM2A. It has got 4 exons as shown in the slide and the protein is a phosphatase, let us not worry about the name, it has got two functional domains as shown in the slide. One that is shown on the left side, amino terminal is CBD that has the function to bind to glycogen, the other one on your right side that is carboxyl terminal is the DSPD which is phosphatasedomain.

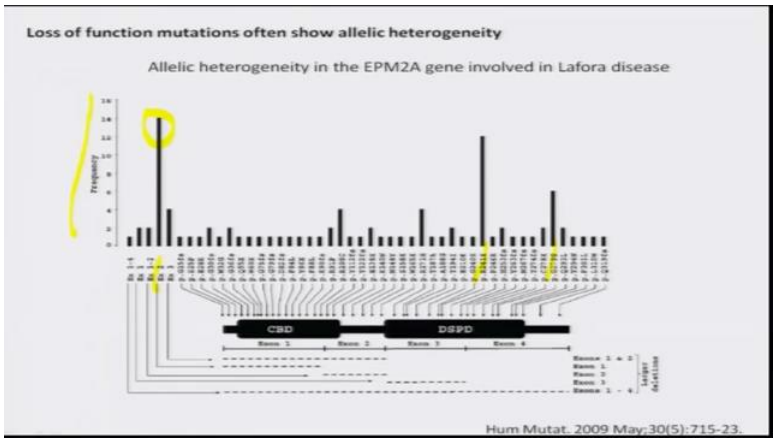
(Refer Slide Time: 11:50)



Now, if you look into the mutations, you can see that it is spread all over the gene, right from exon 1 to end of it, you will find all sorts of mutations.You have deletions, as you can see here, All that are shown here are deletions, the deletions spanning either the entire gene or a part of the gene, for example exon 1, exon 2 or exon 1, 2 and together these are deletion mutations. You

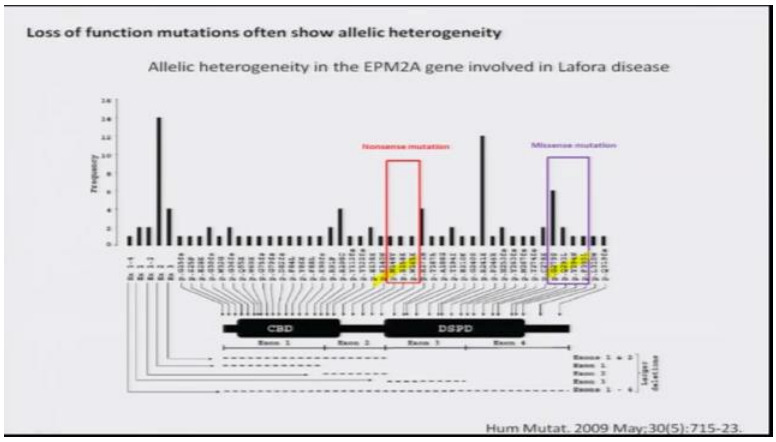
also have other mutations, for example frame shift, missense, nonsense,you name any mutation that you know and would find in this particular gene. So, that is what explains what is known as allelic heterogeneity. You have50 different alleles, mutant alleles that are associated with the disease in this particular gene. So, if you have an allele that is a deletion allele and if you have an allele that is missense mutation, both of them are resulting in the same disease, likely that the missense mutation that you found also is a loss-of-function mutation.So, that is with regard to the type of mutations or what you call as locus, allelic heterogeneity, sorry. So, but if you look into the frequency of the mutation, then you see variation. Some of the mutations are more frequent in the population as compared to the other.

(Refer Slide Time: 13:14)



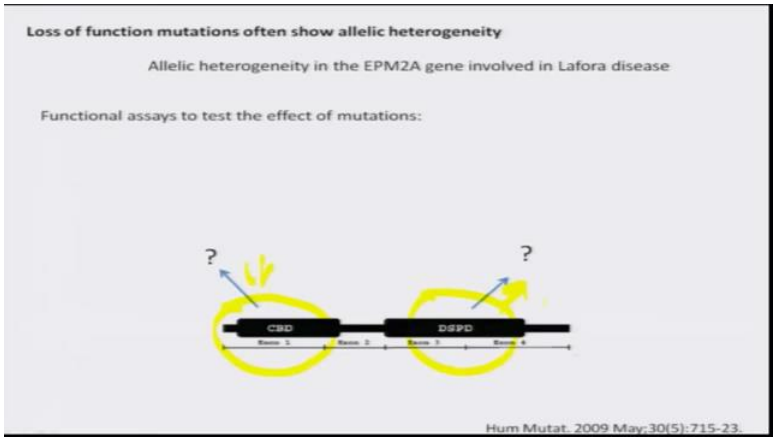
Let us look in here. For example, what is shown on the y-axis ishow many families,
independent families, that are not related to each other show a particular type of mutation. As you can see here, a large number of families show exon 2 deletion. Likewise, you have a large number of families showing this nonsense mutation that is R241X, which is a nonsense mutation, followed by a missense mutation and so on. So, there are some alleles that are more frequently involved in the disease as compared to the other.So, that is a classic signature for a loss-of-function mutation.

(Refer Slide Time: 13:57)



Let us look into these two classes. You have, for example three different mutations, all of them are nonsense mutations and for example here, all of them are missense mutations. So, what we are trying to say is regardless what mutation you see, they wouldlead to what you call as loss-of-function effect.

(Refer Slide Time: 14:18)

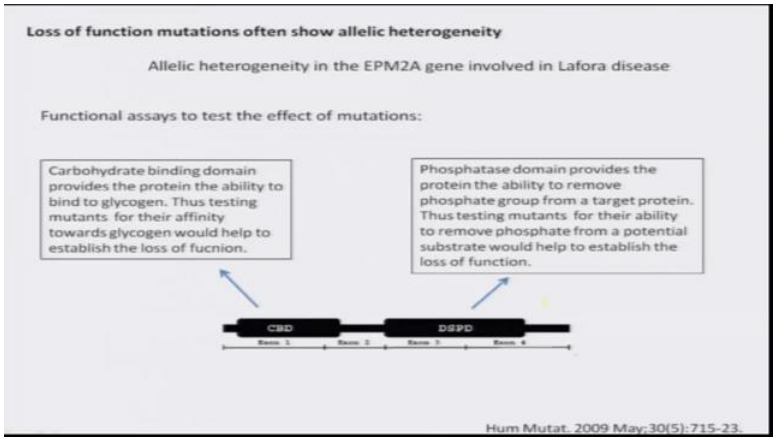


So, how do you test them? For example, if you have a missense mutation and how do you know that indeed it iscausing the disease? One way to look at is that you go and look for that particular variant in the normal population and if you do not find in normal population that means this is not a common and normal allele. That is one way. The second thing you want to look at is, whether a given amino acid residue that is mutated, is it functionally conserved, meaning if that residue is very very critical for that protein function, then you would find that residue to be present inthe proteins coded by the closely related animals. You look into the mouse, rat and

monkey, as manyanimals that you want to sequence and see whether this amino acid is present in all the protein, the equivalent of this protein in thatgenome. So, that would tell you, whether it is critical. But, it is more important to do some functional assays. So, what functional assay would you do?

An example is shown here on your slide. So you have, as I said you have two domains. One is the CBD, which binds to the carbohydrate. So, for example, it can bind to, the protein can bind to glycogen. The other one is the DSPD, which I said is the phosphatase domain. So, you can ask the question. If I have a mutation in this domain, would the protein bind to glycogen or not? If it does not bind that means it has lost its function. Likewise, you can ask a question, if I have a missense mutation in this domain, whether the proteinhaving this, the new amino acid because of the mutation, still show the kind of activity, that is phosphatase activity?

(Refer Slide Time: 16:11)

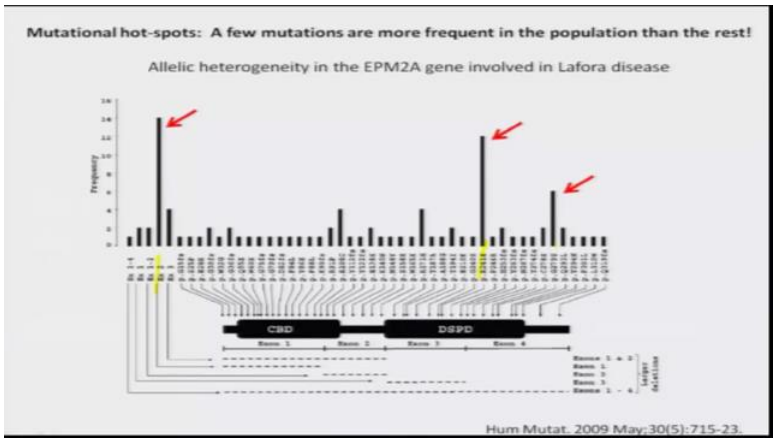


This is what I have shown here. So, carbohydrate binding domain, so it gives the ability for the protein to bind to glycogen. So, you can express the mutant and test whether it binds to glycogen. Likewise, if you have mutation in the phosphatase domain, you can express these mutants and do some assays with substrate and see whether it removes the phosphate group. So, this is, routinely people use this kind of approaches to characterize a given mutation to see whether that change has altered the function of the protein. So, in all these, the prerequisite is that, you should know what is the function of the protein? So, that itself, is time consuming work. So, you can

understand now, how difficult it is to carry outthis kind of experiments. But, that is how people do it.

What is interesting that we have looked at is that you have all sorts of mutations, because the mutations normally are random. I gave an example earlier that you are in a room and there are, is a wall. You must have gone to some trade fair and so on. They have a huge screen, balloons, stuck to that and they will give you a gun and then, you can shoot. So, it is very easy, because there are, everywhere in that board you have balloons, so you do not need to aim, as long as you aim at theboard you will be able to hit one of the balloons. But, if a large screen has got only one balloon, if I give you 3 bullets to hit, using one of the 3 bullets, it is going to be extremely difficult for you. So, mutations are normally like that, you have the residues all over in your genome, the error happens more or less randomly and some errors, can cause a disease, therefore it does not get to the next generation. Some, it does not really affect the way your genome functions, therefore you carry these alleles. That is what you call as randomness. But there are, at times some mutation that are more frequent than the other, that really does not explain that the mutation could be a random process. So, let us look in here.

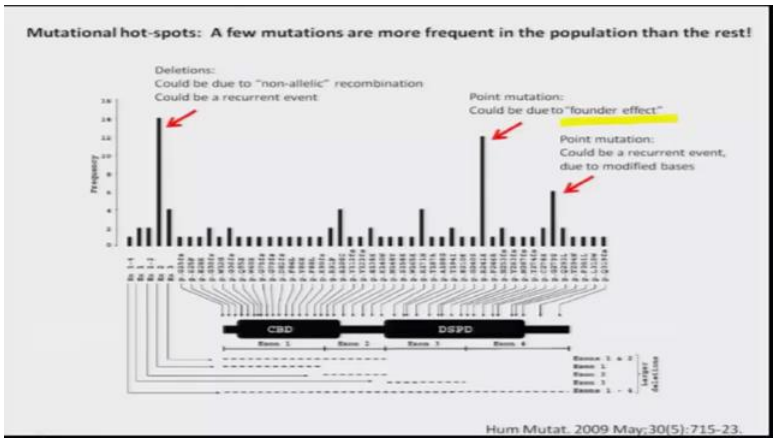
(Refer Slide Time: 18:19)



This is something that we have already discussed that the bar diagram here shows the frequency of these mutation in independent families. We are not talking about multiple members in a family, but we are talking aboutmultiple families. What is interesting? So, you find 3 tall buildings, each one representing, the alleles, the mutant alleles that are more frequent as compared to the rest. So, let us look into what they are. There is one on the extreme left, which I

said is a deletion allele, the one here somewhere in the phosphatase domain is the nonsense mutation and you have another, which is missense mutation. So, what could be the reason? Why these alleles are more common than the other?

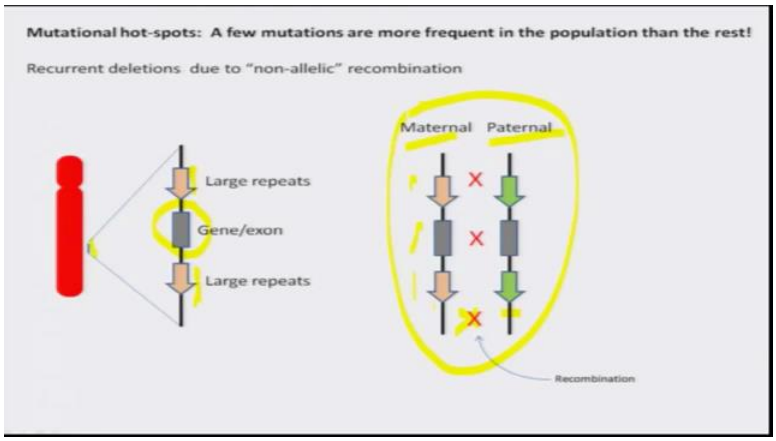
(Refer Slide Time: 19:02)



I am giving some example; for example, deletions. The deletions could be due to a process called as non-allelic recombination. These are illegitimate recombination, the error during the mitotic process that can result in deletions, right and this process depends on your genome sequence, so it can be a recurrent event, meaning it can happen again and again during meiosis. Therefore, you have more chance of this mutation coming into the population. We will discuss that. The second point could be what is called as the "founder effect". What is founder effect? This is, it depends on the few individuals who seeded the population in an isolate, for example Ireland. It was not having any human population before. A group of 15 individuals go and settle there. Now, it all depends on what kind of good and bad allele that carried that is what is called as founder effect; we will discuss again.

The third possibility is that often these are point mutations. It could be because of a modified base. These are some changes that happen in your DNA, because of certain requirements and these changes in the bases can lead to mutations and that could be recurrent. These are hot spots, what they called as. So, let us see what they are.

(Refer Slide Time: 20:22)



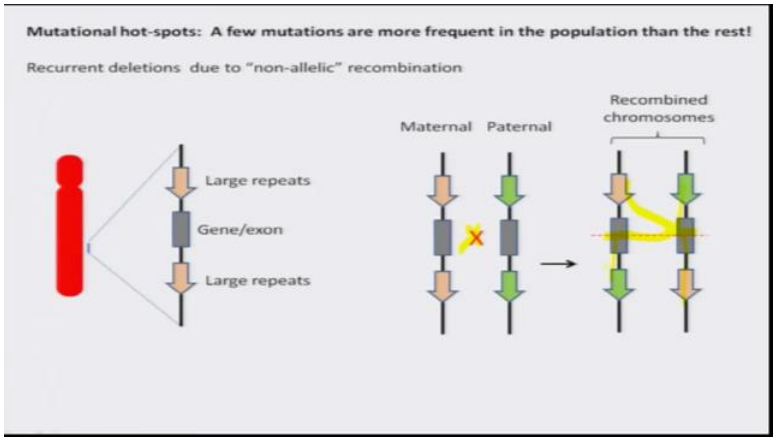
Let us first take what is called as a non-allelic recombination leading to large deletions. More often when such kind of deletions happen, the break points are well defined. Exactly in a given region, the DNA is cut and they are joined together, you have lost a particular segment. Remember, we discussed an imprinting disorder, wherein we said that you have deletion leading to Prader-Willisyndrome. So, how such precise breakage at the DNA takes place? So, let us have a look at it.

So, this is what it is. So, what you are looking at on the left side is a chromosome, right and you are, you are just amplifying a small segment of the chromosome. Let us say, this segment has got a gene that you are interested in or an exon and it may so happen at certain regions, your genome, a gene may be flanked by a large repeated segment, identical sequence, one on the upstream and one identical sequence on the downstream of the gene that you are looking at. When you have such kind of genomic architecture, such complexity, then more often, such regions are more prone for deletion, right? In other words, whenever you looked into mutations that are happening recurrently, such deletions, you find that these hotspots are flanked by such large repeats, right and that is why we believe that these repeats result in such kind of deletion. So, what could be the mechanism?

So, we are looking at a meioticprocess, the recombination in meioticcells. So, this is what is shown here. So, what you have is in a cell, for a given segment, for example the same segment that is shown on the left side, we are looking at the paternal and maternal copies. Obviously, of the two copies of autosome, one you received from mother, the other one from father. Now, both

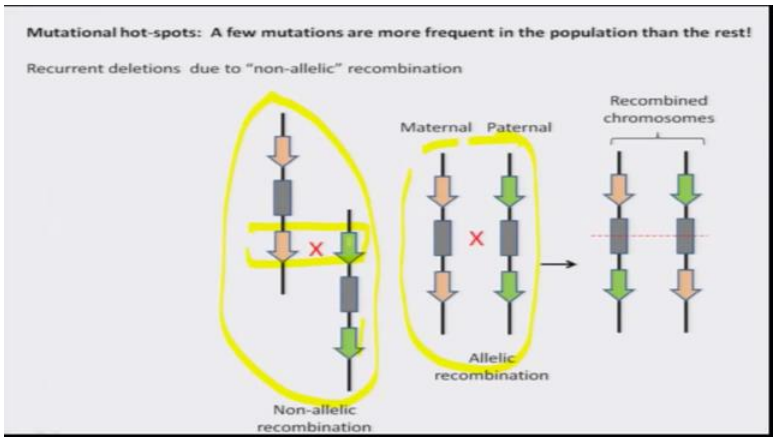
of them have similar architecture that is you have a gene flanked by the repeats, right? So, just to distinguish the paternal and maternal repeats, so I have given a different colour, but nonetheless, assume that sequence wise they are identical. So, paternal is green, maternal is kind of pinkish. So, what happens? The cross denotes the recombination. Recombination meaning the DNA will be broken and will be joined with this. The recombination is random. It can happen here, here, here, here, anywhere.

(Refer Slide Time: 22:55)



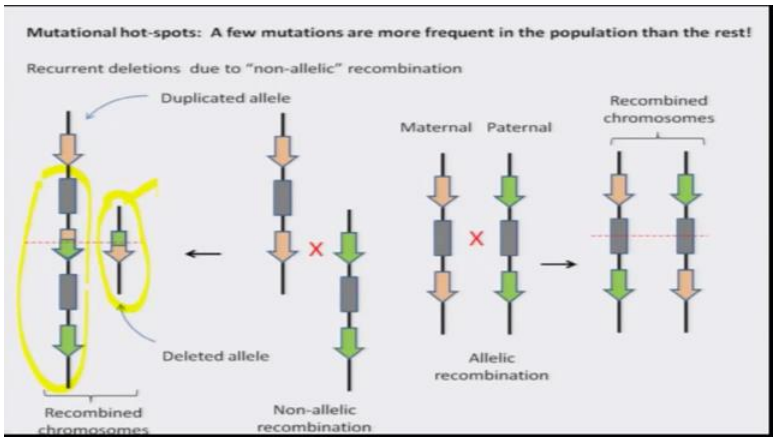
So when, when the recombination happens, this is how it happens. For example, the recombination happens within the gene. So, what happens? There is a DNA double strand break here and you see that the paternal one got joined to maternal and here the maternal one got joined to the paternal. So, this is how it is. So, exchange took place. That is what happens in your meioticcells, what you call as recombination and this is so precise mechanism, it exactly cuts at the same base, even if it is within an exon, by changing this allele, you do not alter the coding region. So, this is a normal process. It happens in every meioticcell and we all carry this kind of recombinant chromosomes and we are all normal. But, at times it can create problem.

(Refer Slide Time: 23:35)



So, this kind of recombination, where you have identicalexchange resulting inwithout any net gain or loss, you have precisely the chromosome broken and joined together, it is called as **allelic recombination**. But, you could have a situation like this, because your recombination machinery basically looks at the sequence of our DNA and try to bring these homologues together and cut and join the chromosome. So, here this is a peculiar or very intriguing condition, because you have two repeat elements, which are almost identical in sequence. So, either it can align like the one that is shown in the allelic recombination that is a perfect homologue coming together or there could be what is called as misalignment, which you call as **non-allelic recombination**, meaning theses two repeats are identical, so there is a recombination taking place. So, if it does take place, what will happen?

(Refer Slide Time: 24:34)



So, you are going to have two novelalleles that is coming out of the recombination. In one allele, you have lost the gene; in the other allele you have duplicated the gene, right? Now, it depends, whether you see any one of the two alleles in the next generation. So, especially if this deleted allele is going to be a **loss-of-function**, because you have lost, assuming that a loss-of-function allele results in a recessive disease, you may still see a next generation, because that individual could be a carrier. He is heterozygous for that mutant allele, so you have in the population and if it so happens that he marries another individual, who is also a carrier, you are going to see the baby, the next generation there is a probability that, that individual would have the disease.

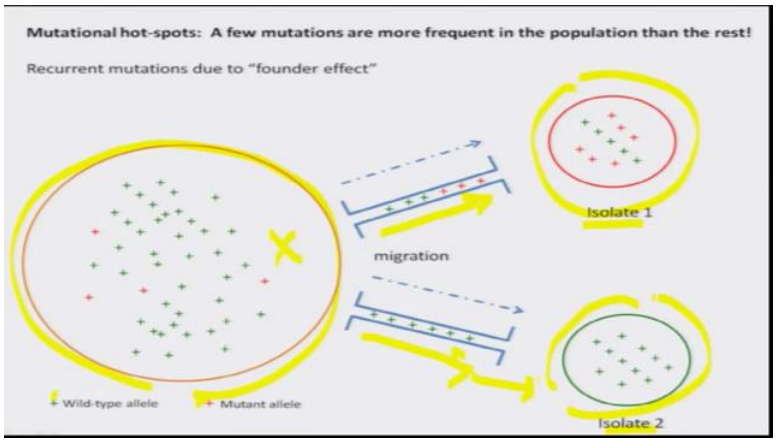
But, the other allele, wherein you have two copies of the chromosome, may or may not be seen in the population. Because, if there is a increased dosage, for example this contributes to the formation of a, either a sperm or a oocyte, then if it fertilizes, for this gene we are going to have three copies. So, that could be at times, causing some embryonic growth arrest. Therefore you never see or the embryo may survive and you may have a condition, like for example Down syndrome. There are regions that are, three in number. It need not be three copies of chromosome 21, but regions of, the long arm of chromosome 21. If there are three end copies, then you still have Down syndrome. So, these are the conditions by, by which you could have that.

So, since I am talking about repeat flanking a gene, more often resulting in what is called as non-allelic recombination, you would often find the breakage happens where the repeats are, right? So, therefore it is a recurrent event. It happens because, not because of the gene, but because of the repeats that are flanking a given gene. So, that genomic structure predispose that region, more prone to have such deletions. So, that is the hotspot. That is how more often you see deletions.

We are going to look into another aspect as to why you have more often a particular mutation more prevalent in a population. So, we call it as **founder effect**. I gave an introduction that it all depends on the few number of individuals who seed the population in a place that was not, , you had human settlement. So, it is a new set of individual go and occupy and then the population

grow and it all depends on what kind of allele they carry. So, this is something like what is shown here.

(Refer Slide Time: 27:11)



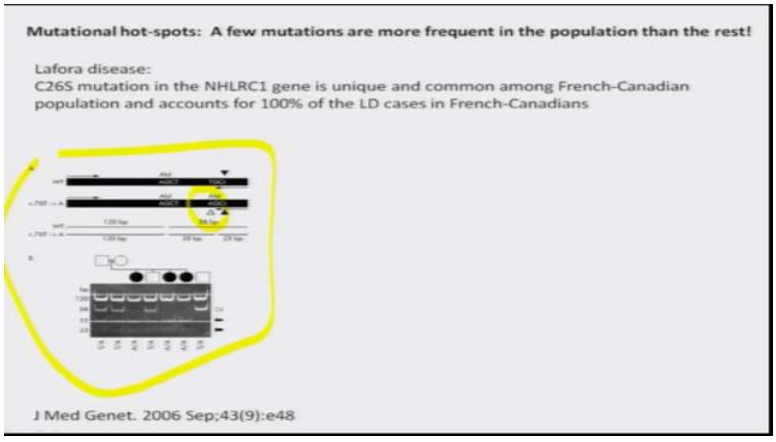
So, you have a large circle that is this population, let us say and you have individuals that are having the wild type allele for a given gene that are shown with, the green colour plus and also a few individuals who are, let us say carrier for the mutant allele, right? The carrier, they do not show symptoms, so they are carriers that are there, but the frequency if you calculate it is very very small. Now, let us assume a condition that there is some disturbance in this population.It could be drought, it could be flooding, it could be something or a war or ethnic clash, whatever it is. So, as a result the population is forced to migrate. So, not everyone go to the same place. The different groups migrate to different place, wherever they find, the place suitable for living. It may so happen that the place they go and settle is a place where nobody lived before.It could be challenges, but people go, because they believe that could be a better place than they have been living so long.

So, it all depends now, who are the individuals that go together as a group. I have shown here two examples. One is isolate 1, isolate 2. We call it as isolate, because we assume once they migrate to the two different place, they do not mix with each other. They are geographically separated and the, the population grows; therefore the eachpopulation would have its own signature.

Let us look into the one on the top.This is a migration. So, it so happened here, almost equal number ofcarriers and individuals having wild type migrated to thisgeographical location and they seeded the population. So, what do you expect? You will expect here that the chance that a carrier will marry another carrier is very high, because 50% of them are carriers. So, one in two is likely to be a carrier for the same mutant allele. So, as a result, you are going to have very high incidence of a given disease, if that mutant allele is causing a recessive disease. Look at the other scenario. Again, a different, small population migrated to this ata different place, which you call as isolate 2. It may so happen here that individuals all that migrated to the isolate 2 arecarrying the wild type allele. There is no recessive allele. So, as a result isolate 2 will not have the disease, unless a new mutation comes in which is extremely rare condition.

Therefore although the original population from where these two isolateoriginated, but you will find depending on the individual that form these two population, kind of bring the risk of having a disease. Isolate 1, higher risk of having the disease, because few individuals who founded the population had this recessive allele. Therefore, you call them as founder effect. It is because the founders, right, whereas isolate 2, the founders were all carrying the normal allele, therefore this population is not, with the burden of having the disease. So, this is what you call as founder effect.

(Refer Slide Time: 30:29)

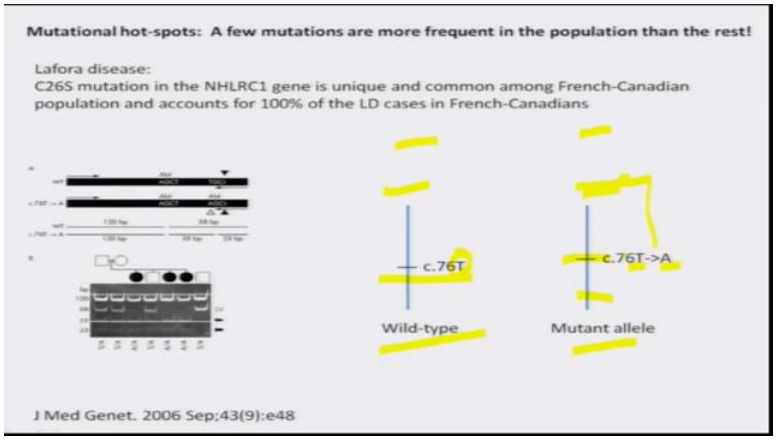


There are many examples. Again, I am going to discuss about one example of the disease that our lab has been working on that is lafora disease. This disease is very unique in the sense, you

have it in India, Pakistan, many Asian countries, you have it in the European population, like the French, Italian, the Spanish, all you have, but you do not find an individual who is having the disease who has descended from the Britishpopulation. As you know, the North American continent, was populated bythe British and if you consider Canada, again these are seeded by either the British population or the French population.

So, if you look into the lafora disease patients in the white population of Canada, you would often find that the French Canadian or the Canadians who migrated from France, they have the disease, not the British and what is interesting is that all the French Canadians who develop this disease, have the same mutation. That is C26S that is cysteine mutated to serinein this particular protein of the gene, which you call as NHLRC1. So, that is again an example of founder effects. So, this is something that I have already shown when we were talking about diagnostics that you can use the restriction fragment length polymorphism. So, because of the mutation, you may gain a site or you may loose a restriction site and you can use that as a diagnosis to identify an individual is carrying a mutant allele. So, what is interesting, how do you know that a given mutation is because of founder effect? It could be recurrent, as I said there is a hot-spot, again and again the same mutation comes into the population, how do you know it is because of founder effect.

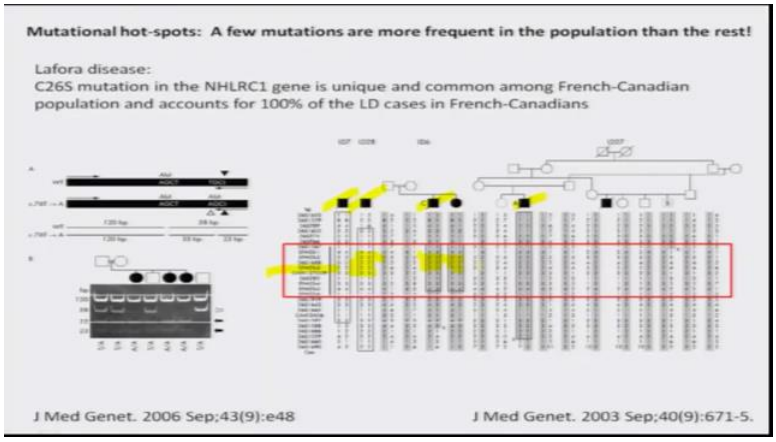
(Refer Slide Time: 32:27)



The way the people look at is the neighbouring regions of the mutated region. As is shown here, this is a wild type allele. It carries residue T at position 76. where is a mutant allele, the T is

converted to A. So, what we look at is that what are the other variants that are normally present in the population that may not be causing the disease, ok. But there, you know, they are in the same DNA, therefore an allele that is present here, which is not linked to disease being separated by a mutation that causes a disease by a recombination is less likely if they are present very close to each other. So, we look into the proximity and tell whether it is a founder effect or it could be an independent event.

(Refer Slide Time: 33:20)

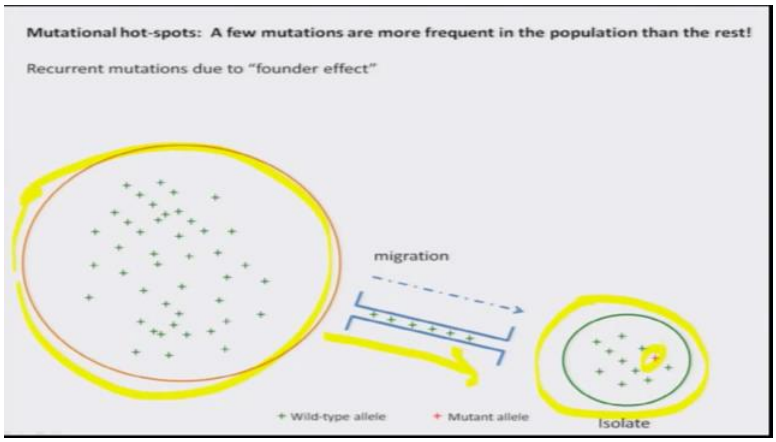


Let us say this. So, we have an allele for a different marker, not necessarily a coding region, but some other marker A and B on either side of thewild type allele carries T. But, in the mutation, in the mutant allele, it could be a different allele C and D. So, you may not see the C and D in a normal population with the wild type. It could have happened after the mutation took place. So, invariably it will be present along with the mutantallele. So, we use this kind of approach to look into a large number of families having the same mutation. That is shown here. For example, here is somewhere the gene having this mutation that is C26S and you can see that all these families, 1, 2 independent, , affected members, family 3, family 4.

In all these, if you look in this region that is shown within red, rectangle, you will find they carry the same allele, 2 2 2 2 2 2 3. So, you find that, all affected members, having the same allele. That shows that all the mutation must have originated from the same individual, who may be one of the individuals who migrated from France to Canada and he was a carrier and he seeded the entire population and the chance that the same mutant allele coming in combination

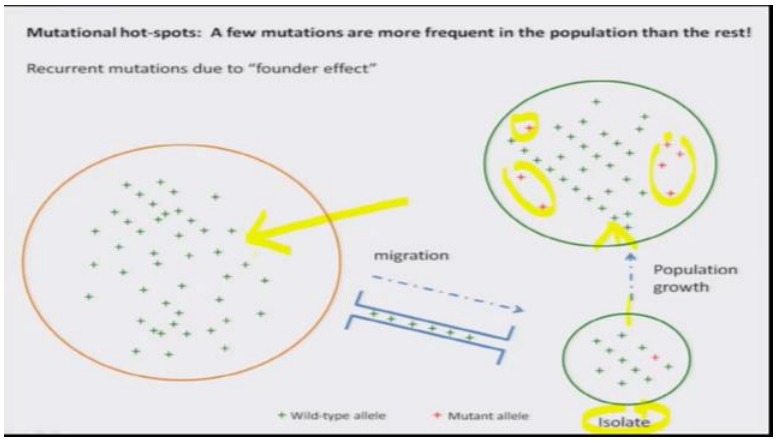
with other, because of another carrier marrying this person is highly likely and therefore, he would have the disease. So, this is what you call as founder effect and that is how you normally characterize and see whether it is a founder effect or it could be because of some other reason.

(Refer Slide Time: 35:04)



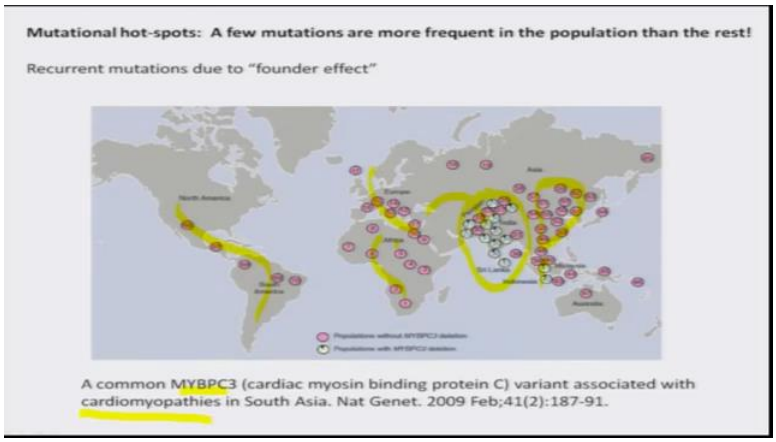
So, it could may happen. For example, the person who migrated from France to Canada may not be a carrier. But after, you know, he migrated to Canada, in one of his germ cells there was a random event that changed this, you know, mutant allele. As a result, you know, that so happened that that germ cell happen to fertilize the egg, who is now his son and then, he carried that mutant allele and likewise, it could have come in the population. So, this also can happen. That is what is shown here in this diagram here. So, you have a population, all of which could be normal, meaning they may not have a defective allele. All of them are normal, but there could be a migration process, wherein a small number of individuals, because of some reason migrated out and formed what you call as isolate. Now, in this isolate, there is a mutational event that took place, because of some error and one individual carried this.

(Refer Slide Time: 36:20)



Because of this small population you have one individual had this mutation, because of a random process. Now, the population grows and you carry on when you migrate. Now, what you will find that this mutation now increase in the population and now, this mutation is unique to all the population that migrated from this isolate and you may not find this in this population, so because it happened much later after they migrated from the main land.Thisis another example as to why certain alleles, mutant alleles are present only in certain population, not in the other, like I told you that British do not have and so on.

(Refer Slide Time: 36:57)



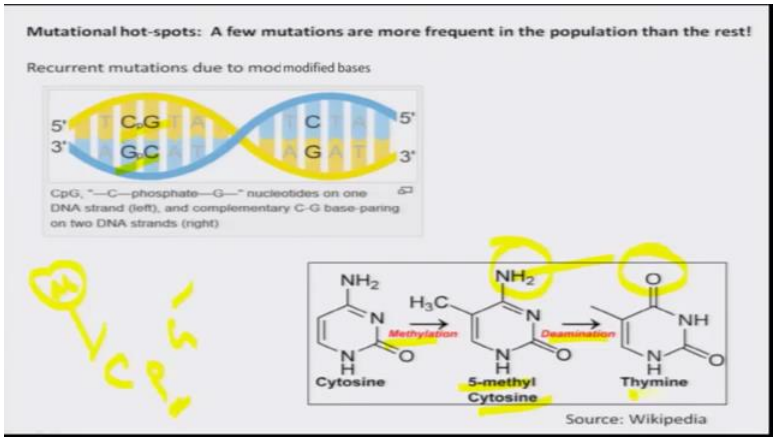
So, that is a very good example for Indian population, because there is a, there is a gene called MYBPC3. It codes for a protein that is involved in the cardiac function, meaning your heart muscle function and this gene has got a variant. There is a deletion in this gene, few base pairdeletion and this deletion is present even in the normal population. But, if you have a deleted

allele, you are at higher risk of developing a condition, what is called as cardiomyopathy. Your heart do not function normally, because of some changes in the structure of the heart muscle and that may result in death, premature death.

What is interesting is that this particular mutation is present only in the Indian population. You see here, so that is where you see. You do not see in the Americans or Africans or Europeans or Chinese, Japanese, but you see in some part of Malaysia and Singapore, because these are again people migrated from India. So, this is what I explained to you. You know, from here somewhere the humans migrated to India. The individual who migrated to the Indian subcontinent did not carry the mutation and then the population grew and that is when somebody probably, , had a mutation. Now, they seeded the so called Indian population. Now, all of them have this. The main like, for example Africans do not have, others do not have, but whoever migrated from India may carry this allele. So, likewise you see in Malaysia and Singapore and that is how you see this allele. So, this is example of how the founder effect mutation can lead to the increased frequency of certain alleles.

You are going to go to the third example as to how, you have mutation hot-spots. Why certain alleles are more often present, right? So, here we are going to discuss about the modified bases. What is that?

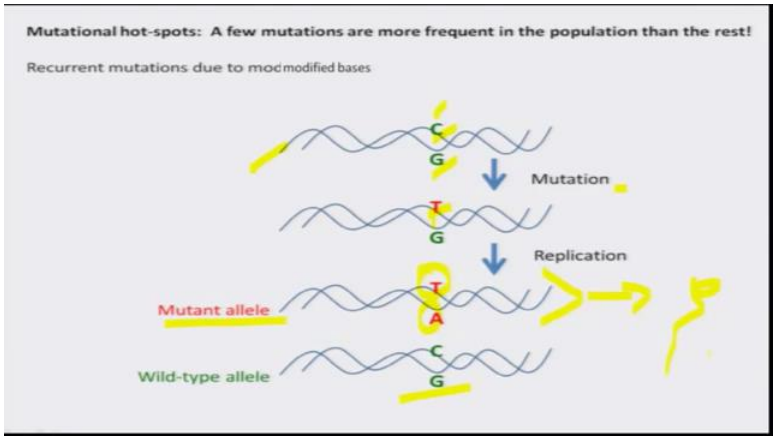
(Refer Slide Time: 39:00)



As you know, your genome has got several sequences. One of the sequences which gets modified, it could be a base onto which a methyl group could be added. That is one of the modifications which you call as methylation. So, more often, such methylation happens in what is called as CpG islands. So, when you have C base and C, C and G together, the p here refers to, the CpG, the p refers to the phosphate and of course, C and G are the two bases. Now in such cases, the methylation, the C base gets methylated. So, when that happens, this is one way your chromosome or chromatin what you call gets some message as to how that region of the genome should function. It could give a signature that whether the gene should express or not express or some other signature, right? Now, it can, it does not change the base, but the way it functions, that stretch of the DNA functions, depends on this kind of modifications. But, at times that modification can also put a risk, risk of having some mutation like what you see here. Cytosine gets methylated.So, you have the methylation that you see here.

What happens in this condition is that you could have a process called as deamination. So, this region, you can see that that is being removed and as a result, when you remove this region in the methylated cytosine, you would find that this residue now becomes thymine.So, this change, it is rare. It does not happen always, but at times if it happens, then you could end up having a new base in that place.

(Refer Slide Time: 40:47)

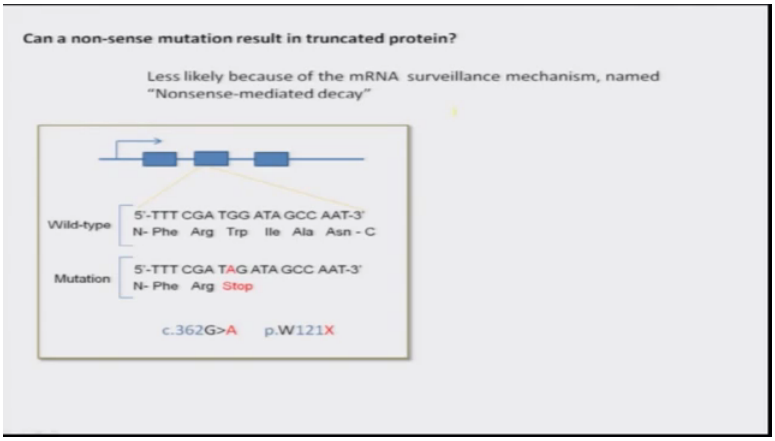


So, that is what is shown in this cartoon. So, to begin with, you have C and G as a complementary in a double stranded DNA and the cytosine here gets methylated and that is

leading to the mutation that you see here, T here. There is a mismatch.So, there is a mechanism by which this mismatch is removed, but if it does gets, what happens? When the DNA replicates, now each strand serves as a template. Now, if the T serves as a template, you are going to have A as complementary , so that is your mutant allele, which is very different from the parental DNA and of course, you are going to have the wild type allele. So, now if this results in the formation of a sperm or egg, and that is going to carry a mutant allele, if it affects a function.

So, such change is going to be recurrent, because this cytosine getting methylated possibly a signature for the gene to function or that segment of the genome to function, it is going to happen in every individual. So, such residues are at higher risk of getting methylated. That is why you have such hot-spots, wherein a given mutant allele more often comes into the population, not because of founder effect, but because of this kind of changes.So, what we are going to ask a question is can a nonsense mutation, meaning a mutation that gives a stop codon as, as the message instead of an amino acid can result in a truncated protein? So, if you have a change, so that is going to tell the protein synthetic machinery that you have to stop here.

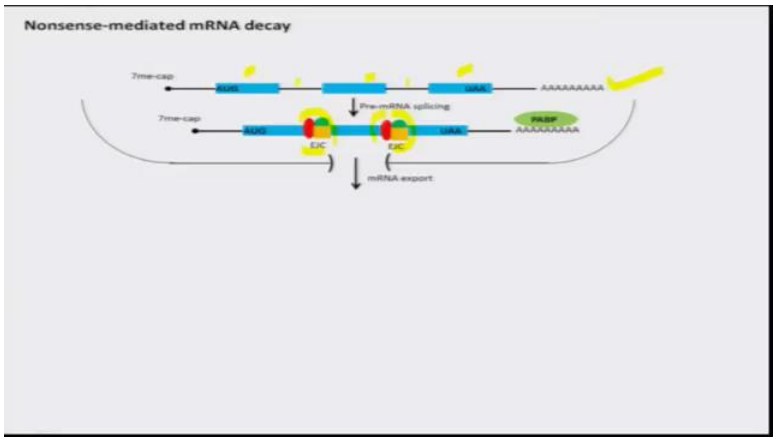
(Refer Slide Time: 40:47)



So, like what is shown here in the screen, so would it result in truncated protein? So, less likely, because in your cell, you have a quality control mechanism, which is called as mRNA surveillance mechanism and the technical name for this mechanism is nonsense-mediated decay. In the sense if there is a premature stop codon in your mRNA, you have a mechanism in your cell that scans through the each and every mRNA that is being made and if it has got a premature

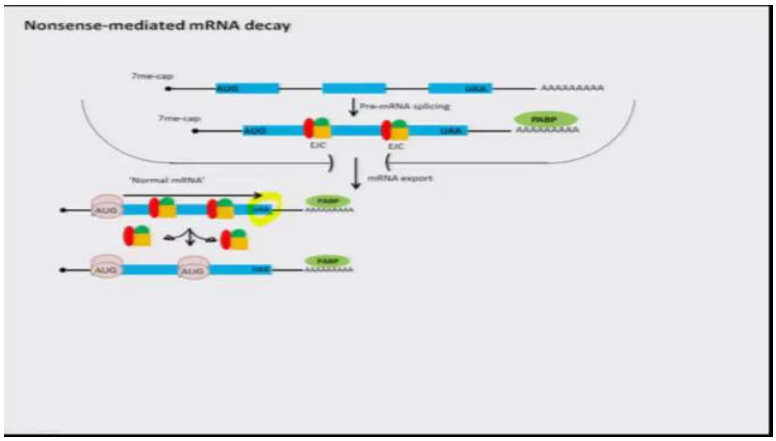
stop codon, then the cell tries to degrade that mRNA, because when the DNA is being copied to RNA, there could be error and if that error is not detected, you are going to make proteins that are defective and you do not want such kind of protein, because one, you waste energy in making such protein, two, such protein can be harmful to the cell. So, you try to get rid of the mRNA. Therefore, they do not make, protein that are not desirable and such mechanism even identify transcripts that have such premature nonsense codon, premature stop codon. So, let us look into how this nonsense mediated mRNA decay, this mechanism operates in the cell and how does it really remove transcripts that carry premature termination codon.

(Refer Slide Time: 43:40)



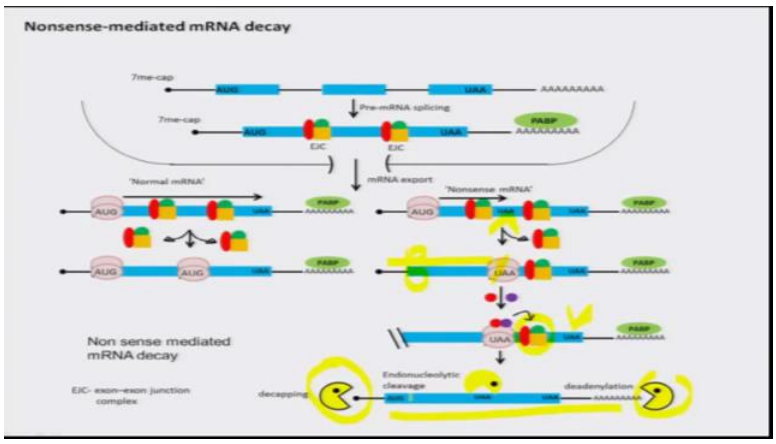
So, what is shown here in this schematic is a transcriptional process, wherein you have the so called primary transcript, which retains both exons and introns. So, these are the exons and these are the introns. So, the RNA, as we discussed, undergo a splicing process. During the process, you have a set of protein that come and bind to the exon boundaries or the junctions

(Refer Slide Time: 44:05)



So, these protein which are bound to the exon junction, now along with this, the entire complex is transported to the cytoplasm. So, what we are seeing is a normal mRNA, wherein you have a stop codon that is present in the last exon. Therefore, you really do not have theexon junction complex present downstream towards 3 prime end of the stop codon. So, when, a RNA comes into the cytoplasm, you have these ribosomes that comes and binds and scans for whether this RNA is good enough for translation. So, what it does is, it pretty much scans and then looks at the RNA and then codes for, may be a peptide. But, during this process, what happens is, it displaces all the complex protein. The exon junction protein are removed as the ribosome complex is, scanning the mRNA.

(Refer Slide Time: 44:57)

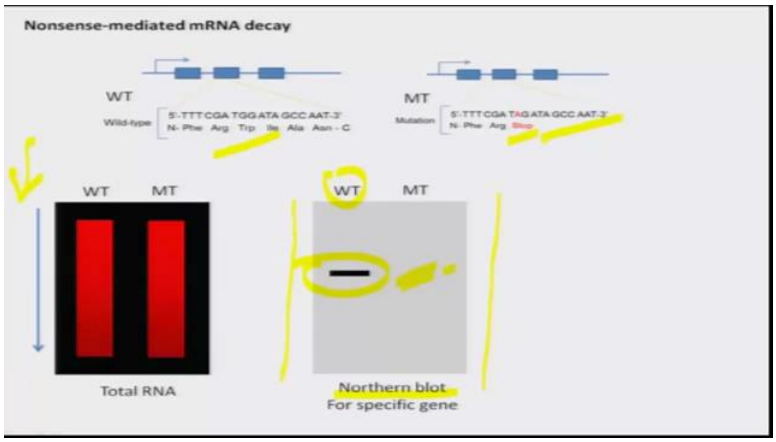


On the other hand, if you have an RNA which has got premature termination codon, like what you see, because of the mutation, and that could be present in a exon, which is not the last exon.

So, in this process what happens, the ribosome comes and binds, of course at the 5 prime end of the RNA and scans through and the ribosome is going to fall off as and when it finds a stop codon. So, in this process, you have a stop codon here. It arrests its translation over there. But while doing so, it is unable to dislodge some of the junction complex that are still present, because the actual, the original stop codon is present downstream to that, exon complex, the junctional complex.

So, in this kind ofcondition, there are set of proteins that activates the complex and initiates a process, wherein you have an enzyme which try to degrade such mRNAs. So, the mRNA is now degraded, therefore they are not allowed to code for the peptide, which may be truncated, which may be abnormal and therefore they are removed. So, this is a mechanism, a pretty much like a proof reading mechanism to make sure that you do not have RNA that have a mutation, which has the potential to make abnormal proteins. So, that is why when you look at individuals that carry these mutations, you may not really see the RNA present in the cell in a detectable level or they may not even code for the truncated protein. So, that is the mechanism what we call as nonsense mediated mRNA decay.

(Refer Slide Time: 46:37)

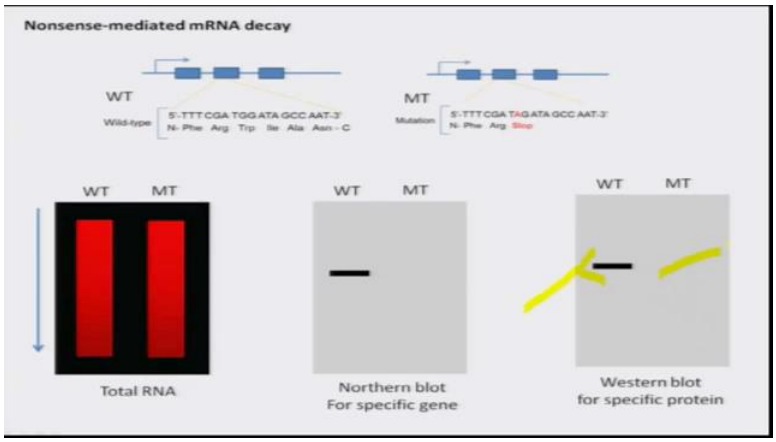


What you can do is, you can, create cell lines out of individuals. So, for example, a normal individual and an individual who carries the mutation.So, what you do, you can generate cell lines or we can collect the blood cells from them or some tissue from them if they undergo some

surgical procedures and then you can look at the RNA. So, you can extract RNA and do what is called as agarose gel electrophoresis. On your left I am showing a gel electrophoresis pattern, where the red smear represents the total RNA. They are different sizes, therefore, they are size separated. Now, how would I know, whether my RNA of interest, the RNA that is made from a given gene is expressed or not expressed?

So, I will take a small region of that gene as a probe, label them with certain radioactive moieties and then try to hybridize, because the DNA sequence are complementaryto the RNA. So, they will go and bind to, if the RNA has got a sequence. This is what is shown here. Just the way we discussed **Southern blotting**, you can now probefor an RNA which is there on a membrane, with a DNA probe and this is called as **Northern blot**. So, if in the wild type I have a normal sequence, therefore the gene is expressed, I get a signal like this. But, the mutant, because of the nonsense mutation, now the system tries to remove the RNA. Therefore the RNA is not at all degraded as they are synthesized. So, you will not find any signal in this place, which indicates that RNA is not present in the tissue, meaning they are removed, they are degraded and you can ask the same question, whether the protein is being made.

(Refer Slide Time: 48:25)



So, I can use an antibody which recognizes the protein. Like I separated RNA or DNA, I can also separate protein according to their molecular weight in the gel and then transfer to a membrane, just the way we described for Southern blotting and then use an antibody which would recognize, for example a peptide. The peptide could be your protein of interest and if it is expressed, you are

going to get a signal and if it not expressed, you are not going to see signal. So, this is the way we sort of quantify and then establish thatnonsense mutations results in the degradation of the RNA. Therefore they are not, therefore even the truncated protein, peptide is not being made.So, this is the example of, the hot-spots and other such mutations that we have looked at.

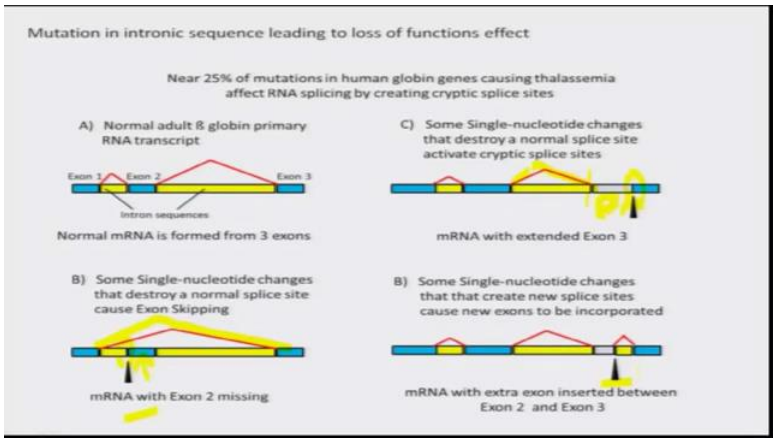
All these that we discussed are in the coding sequence or the entire gene. So, what if there are changes in the non-coding regions, for example intronic regions? The segment of DNA, when copied into RNA, anyway they are spliced out, removed. So, if there are changes, would it affect? The question is it could.

(Refer Slide Time: 49:42)

If the change affects some functional property of that, gene, the transcript and so on. An interesting example is beta thalassemia. This is a condition, wherein, the individual having this kind of a mutation in a gene that codes for the globin genes, what happen is their capacity to carry hemoglobin, the oxygengoes down. So these are, normally deficient for the oxygen. Now, majority, nearly 25% of the mutation that are known to be associated with thalassemia are mutation that is present in the intronic site that affects the splicing, the way the RNA is spliced. As a result, you have deficiency of hemoglobin and this is one of the very common conditions in the Indian population.

Let us look at how does it really do? So, a change can make a signal for a splice site which is not present earlier. So, that could lead to aberrant splicing. Let us see here.

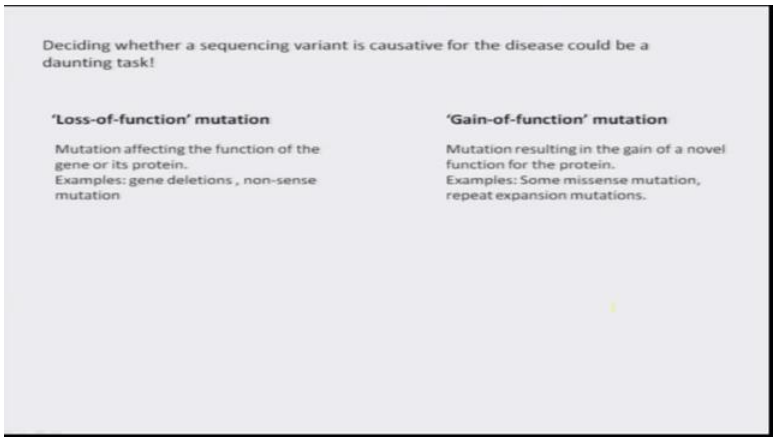
(Refer Slide Time: 50:45)



This is a normal beta globin gene in a primary RNA.Exon 1, exon 2 , exon 3, all the introns that are shown in yellow colour are spliced out and they are joined together to form a normal RNA. So, how various mutations can lead to aberrant splicing. For example, here you could have a mutation around this junction, but in the intronic region, whereby the RNA, now unable to give the signal that this intron should be removed. As a result what happens? You know, you have exon 1 and 3 combine together, you have lost exon 2. This would result in, loss of the coding sequence, which is present in exon 2 or you could have changes that destroy a normal splice site.

For example, you have a splice site here, which, because of mutation here, now the transcript has lost the splice site and therefore, there is some aberrant splicing happens, leading toinclusion of certain sequence that is normally present in intron. So, this would again shift the reading frame and you will not have the desired protein or you could have sites that are altered in the intron, but leading to what you call as a novel exon. Now, it is because of the, it functions as if it is an exon that is retained in the mature transcript. Again, it would change the way the protein is being coded, because after all these are nucleotides and open reading frame is inserted in a new sequence, it is going to affect the way the protein being coded. So, these are called as splice site mutations; can affect the splicing pattern and in certain conditions these are very common like what you see in beta thalassemia.

(Refer Slide Time: 52:38)



So, that is, pretty much covered about how various mutations can lead to what is known as loss-of-function mutation. So, these are the four distinct mechanisms that you see that results in such kind of hot-spots and, and mostly we discussed about what is called as the recessive loss-of-function mutation. So, that pretty much brings an end to this lecture and we will be continuing our discussion on the mutation, especially on the so called gain-of-function defects in the next lecture. See you then.