Cell and Molecular Biology
Prof. Vishal Trivedi
Department of Biosciences and Bioengineering
Indian Institute of Technology, Guwahati
Week 08
Concepts of Genetics (Part 2)
Lecture - 28
Genetic Recombination

Hello everyone, this is Dr. Vishal Trivedi from the Department of Biosciences and Bioengineering at IIT Guwahati. Now, in today's lecture, we are going to discuss the application of molecular biology in the context of correcting genetic diseases. So, what you can see here is that we have different types of genetic diseases that are found in the human population. So, we have Huntington disease, hemophilia, Parkinson's disease, colon cancer, and many different types of diseases that are being found in human populations. And how is the human population managing these diseases? They have had different types of therapeutic options.

One of the options is that you are actually going to generate the drugs, and these drugs are actually going to overcome the deficiency, or he is going to take care of the detoxifications and other kinds of things. So, the mechanisms of these diseases are very different. In some cases, you are actually overexpressing a particular factor, or in some cases, you are not able to provide a crucial component, and that's how you cannot run a particular metabolic pathway. So in either of these cases, you are supposed to provide that particular factor that is responsible for it.

The correction of these diseases. For example, we have thalassemia, muscular dystrophy, cystic fibrosis, Huntington's disease, and so on. All this is happening because one of the crucial factors is either being produced in large quantities or is not available for the human body. And the only way you can actually provide this is if you take this factor and put it into the cloning factor. So let's see how some of the diseases are being managed for human welfare, and we will take a few examples of these diseases.

So we have HIV infections, leukemia, thalassemia, sickle cell anemia, and Parkinson's disease. In HIV infections, we have the responsible mutated gene called CCR5, which is a chemokine receptor, and CCR5 is the receptor responsible for the replication of the HIV virus. And the target cell for this is the CD4-positive T cells because CD4-positive T cells are the cells that are responsible for immune function in the human body. And you are actually producing the CCR5 gene, and that's how you are using the adenovirus as well as the messenger RNA as a delivery vehicle. And that's how you hope that if we do that, you will be able to provide the CCR5, and that's how you can overcome the

mutations, and that's how you can overcome the HIV infections.

The other example is leukemia. What we are discussing is acute myeloid leukemia, or AML. And in that case, you actually have the mutated gene that is present in CD123 or Tdac. So, TRAC is a T cell receptor alpha chain, and the targeted cell therapy is through the CAR T cells. So, you will be able to generate the chimeric antigen receptor T cells, and you are actually going to have the messenger RNA, which will be delivered into a particular T cell type, and that is how you are going to generate the chimeric cells, and these chimeric cells are going to cure leukemia.

Beta thalassemia and theta thalassemia involve the BCL11A, which is known as the B-cell lymphoma factor 11A, and here the target cells are CD4 HSCs, or the hematopoietic stem cells, and you are actually going to use messenger RNA to deliver the clone factors. Then we have sickle cell anemia, and in sickle cell anemia, you have BCL11A, and that's how you are going to have CD4-positive HSC-removing stem cells. Then we have Parkinson's disease. So in Parkinson's disease, you are actually going to have LRRK2, or Parkin, or PINK1, or DJ-1. These are some of the proteins and genes that are responsible for Parkinson's disease, and here you are actually going to cure the IPCs, which are called induced pluripotent stem cells, ESCs, or embryonic stem cells, and mesenchymal stem cells.

And the virus you are going to use is the adeno-associated virus, the adenovirus, or messenger RNA. What is the general approach to treating genetic disorders? So, for example, if this is the patient who is actually suffering from a particular genetic disease. So, what you are going to do is collect the desired cell types. For example, if we take the case of a T cell, you have collected the T cell, and suppose this patient has AML, right? So, what you are going to do is collect the target C cells and then, using genetic engineering tools, whether they are CRISPR-Cas or TALEN or all those kinds of tools, you are actually going to make corrections in the genome of these cells. So you are going to take these cells; you are going to correct the errors.

So you are going to change the mutated genes. And then you are actually going to grow the transformed cells in the in vitro fatty dishes, and then you are actually going to put these mutated genes back into the patient. What will happen is that these transformed cells, which are actually going to have the correct form of the gene, will work for this particular patient, and ultimately it is actually going to overcome the particular deficiency. Now, if we want to do this right, we have our main goal: to correct the genetic information, right? So genetic information is going to be, as you remember, when we were discussing the genome, the genetic information will be present in the form of a genome, right? So, in this particular case, what you are going to do is take the genome,

and within the genome, a particular gene may actually be the target DNA, right? So, this particular gene has some problems, doesn't it? It has some mutations, or it has some kind of abnormality. So, that gene has to be corrected.

Now we have two different types of approaches that people are using: one is called the traditional approach and the other is called the modern approach. The traditional approach depends on the gene sequence. And it has mostly been using the PCR-based method. So where are you actually going to clone a particular gene, and how are you going to use that? Where in the modern approach are you actually going to use the nucleases and all other kinds of sites? Here, you are going to use the ring finger nuclease. You are going to use CRISPR-Cas, and so forth.

So let's first discuss the traditional approach. So the traditional approach is mostly based on the genome sequences. And it also depends on one of the basic phenomena of recombination. So recombination is the core of genetic engineering. Recombination is the genetic process involving the breaking and recombining of DNA segments, resulting in the creation of new combinations of alleles.

This mechanism operates at the gene level, fostering the genetic diversity that mirrors the variation in the DNA sequences among the organisms. So, the genetic combination is one of the basic phenomena through which the two particular DNAs are recombining with each other, and that is how, in that process, they are sharing or exchanging the DNA content, and that mixing mechanism is always being used in the traditional approaches to generate or correct particular types of genetic disorders. So, as far as genetic recombination is concerned, recombination could be of homologous recombination or non-homologous recombination. Within the homologous recombinations, you can actually have bacterial recombinations or eukaryotic recombinations. Whereas in normal homologous recombination, you have multiple approaches or multiple options, right? You have the CSSR.

Within the CSSR, you have the SIR type, or you are actually going to have the tyrosine type. Then we have the NHEJ. We have the transpositions. Then we DJ actually recombinations. Within the transposition, we have DNA transpositions, we have retroviruses, and we have poly A retroviruses.

So, let us first start with the homologous recombinations, and we will first understand the mechanism of the homologous recombinations, and then we will see how the homologous recombination can produce genetic recombinations and how it can, in turn, actually change the genome. So homologous recombination. Homologous recombination is a genetic recombination process characterized by the exchange of genetic information

between two closely related or identical molecules of nucleic acid. These molecules can be either double-stranded or single-stranded and are typically composed of DNA in cellular organisms. Although in viruses, RNA may also be involved in the recombination process.

Mostly, it is involved in the repair mechanism and meiosis of prokaryotes and eukaryotes. So you can understand that you have duplicated parental chromosomes. So these are the two different types of chromosomes, which are homologous chromosomes. So you can imagine that if this is the chromosome from the male and if this is the chromosome from the female, then what will happen is that these homologous chromosomes are actually going to go through the process of recombination, and they will actually do this because they have a particular DNA sequence that is common, right? As a result of this recombination, there will be a genetic exchange of material between them. Two chromosomes undergo the process of crossing over.

Remember that when we were discussing mitosis and meiosis, it was said that crossing over is one of the phenomena responsible for genetic diversity because during crossing over, one part of a chromosome may go to another chromosome, and the other chromosome may also change its content, right? So, it is actually about being responsible for genetic diversity. And the same phenomena can be used to produce the genetic modifications. So, once there is a crossing over between the chromosomes, you will see that the pink portion is being given to the blue portion and the blue is going to be given to the pink portion. So, this portion is actually a part of this chromosome, and this portion has actually been a part of this chromosome, and that's how this portion is now altered in both chromosomes, and as a result, what will happen is that you are actually going to have the separate features of the Offspring. So, and this is just a simple chromosome, you can actually have genetic recombinations even in this region.

So, depending on what kind of sequence similarity you have, it is actually going to do the crossing over. Now, in the homologous recombination, you can have two different types of pathways within the prokaryotic as well as the eukaryotic pathway. So, within the prokaryotes, you can have the RecBCD pathway and the RecF pathway, whereas in the eukaryotes, you can have the DSMR pathway as well as the SDSR pathway. So, first discuss prokaryotic recombination, homologous recombination in the prokaryotic system, and then we will discuss homologous recombination in the eukaryotic system. In the mechanism of prokaryotic homologous recombination, this is the DNA of homologous recombination where you are actually going to have different types of factors.

So, in the first is the initiation. So, rec BCD binds to the blunt or nearly blunt area of the double-stranded DNA break. Then it is actually going to have the unzipping done. So

Reg B and Reg D helicases work together to unzip the DNA. So this is actually going to happen here.

So Reg B and Reg C are actually going to unzip the DNA, which means they are going to have helicase activity. They will unwind the DNA. Then the Reg B nucleates domain cuts the emerging single strand. So then you are going to have the Reg B, which also has the nucleus activity, and it's actually going to cut one of the strands. Then we have the chaiside encounter.

So unzipping continues until encountering a chai side, and the chai side has a sequence called GCTGGGGG. And then we have the chai-side recognitions. So DNA unwinding pause briefly and then resume at a reduced rate. Rec BCD cut the DNA strand with chai. And ultimately, multiple Rec A proteins will be loaded onto the single standard DNA with the newly generated three-prime end.

And then there will be a homologous search. So this is the strand that is going to be available for making homologous recombinations. And then, when ReqA-coated nucleoprotein filaments search for similar DNA sequences on homologous chromosomes, the search induces stretching of the DNA duplex. And then the strand invasions occurred. So, the nucleoprotein filament moves into the homologous recipient DNA duplex, forming a dual D-loop.

And then we have the resolution options. So, if the D-loop is cut, further strand swapping forms a Holliday junction. Resolution by the RUB, ABC, or XG can produce two recombinant DNA molecules with the recipient's reciprocal genetic type, with the interaction of the DNA molecules differing genetically. Alternatively, the invading 3' end chi can initiate DNA synthesis, forming a replication fork, and this type of resolution produces only one type of non-reciprocal recombination. So, this is exactly what is shown here, right? You have the RecBZ system; it is actually going to unwind the DNA, and then it is going to truncate one of the strands.

It will reach the chi site, and then from the chi site, the DNA is going to break. This break will cause the DNA to be coated with the ReqA proteins, and then it will actually participate in the recombination with closely related sequences. As a result of this recombination, it will produce two different types of DNA. Then we have another pathway that is called the REC-F pathway. So, in the REC-F pathway, bacteria employ a homologous recombination repair mechanism for the single-strand gaps in the DNA.

When mutations inactivate the REC-BCD pathway and additional mutations disable the SCCD and XO1 nuclei, the REC-F pathway can also repair the double-strand DNA

breaks. So in this, you are going to have the initiation, and then RECQ helicase unwinds the DNA; RECJ nucleates and degrades the strand with a 5' end, leaving the strand with the 3' end intact. Then we have a REC-A binding, so the REC-A protein binds to the strand with a 3' end, and then the REC-F, REC-O, and REC-R proteins aid or stabilize RecA in this process, and then there will be strand invasion. So RecA nucleoprotein will search for the homologous DNA and exchange places within the identical or nearly identical strands in the homologous strand. The strand invasions will occur, and then there will be branch migrations.

So, similar to the REC-BCD pathway, it involves the movement of the Holliday junction in one direction, and then there will be a resolution. So, similar to the REC-B pathway, holiday junctions are cleaved apart by the enzyme in a process of resolution, and both pathways may undergo an alternate non-reciprocal type of resolution. Despite differences in the proteins and specific mechanisms in their initial phases, both the RECB-CD and RECA pathways share similarities. They both require single-stranded DNA with a 3' end and RecA protein for strand invasion. Additionally, the pathway exhibits similarities in the phases of branch migration and the resolution of Holliday structures.

So this is the pathway of the other pathway, and the mechanism of this pathway is also going to be the same. And then we have the mechanism of eukaryotic homologous recombination for repair. So remember that the homologous recombination mechanism is used only to repair damaged DNA because the information from the damaged DNA is missing. So you can actually bring in that information from the neighboring residues. So, with the help of the neighboring residues, you search for the homologous DNA and then actually copy the damaged DNA.

So in the normal cell, you are going to have double-strand DNA breaks, and that is how you are going to have the activation of homologous recombination in the eukaryotic system. So ATM recognizes the DSB, phosphorylates H2X, facilitates MDC binding, and then we have the MRN complex, which includes MRE11, RAD50, and NBS, localized to the DSB. Then we have the CT; IP creates a C prime overhang where there is exonuclease activity, and then the RPA binds to the C prime overhang. Then we have RAD51, BRCA, BRCA1, and BRCA2 replacing RPA to form the filament of DNA to proceed with homologous recombination. So this is exactly what you have: a double-strand DNA break that activates the ATM, and then ATM recognizes the dual-strand phosphorylated H2AX and facilitates the MTCB binding, ultimately leading to the homologous-directed repair of the DNA strands.

Then we have the two different types of proposed pathways for eukaryotic homologous recombination-mediated repair. You have the classical double-strand break repair

pathway or the synthesis-dependent strand-annealing pathway. Both of these pathways operate from one organism to another. So, in the classical double-strand break pathway, you have the three-prime end invade an intact homologous template, then the formation of the double-strand Holliday junctions, and then junction resolution results in crossing over or non-crossing over. Similarly, we have the synthesis-dependent strand annealing pathway, which is conservative and results exclusively in non-crossover events.

So, the key factors involved in their role in eukaryotic homologous recombination are. So, we have the MRN complex, and its job is to initially stabilize the double-strand breaks. Then we have the ATM, or ataxia telangiectasia mutated. It is recognizing the double standard that exists. Then we have BRCA1 and BRCA2, which are breast cancer-associated genes, and these are involved in checkpoint activation and DNA repair.

Then we have a CTIP, C-terminal binding protein, interacting with proteins, and it interacts with BRCA1 and D-phosphorylated 53BP1. Then we have RPA1, or replication factor A, which stabilizes the single-stranded DNA. H2AX, a member of the H2A histone family, is responsible for the recruitment and accumulation of DNA repair proteins. And then we have a RAT51; RAT51 forms the filament on the DNA strands. Now, if you summarize the homologous recombination in both prokaryotic and eukaryotic systems, what you will see is that there are proteins or processes involved that are common between the two.

So you have the introduction of double-strand breaks (DSB) that are not present in the E. coli system but are present in the eukaryotic system, where you have SpO1 and HO that are going to be involved in the introduction of double-strand breaks. Then we have the processing: the REC BCD complex, the nucleus, or the helicase and nucleus systems. And here we have the MRX complex, which involves the RAT 50, 58, 60, and nucleases. Then we have the assembly of complex formations and filaments.

REC BCD or REC F pathway, which is going to be RAT 52, RAT 59, and BRCA2. Then we have the pairing and a strand exchange, so REC A. which is involved in the prokaryotic system, whereas here you have the RAD51 and DMC1 exclusive to the mucus. Then we have the branch migrations for branch migration; here you have the REC-UV-ABE complex, whereas it is unknown in the eukaryotic system.

Then we have the resolution of the holiday junction. So RUVC, here you are actually going to have the RAD51C, XRCC3 complex, WRN, and BLM. So this is the comparative study of the summary of homologous recombination in prokaryotic as well as eukaryotic organisms. Let us move on to non-homologous recombination. So, in non-homologous recombination, the recombination involves the physical exchange of DNA

segments between chromosomes or DNA molecules. When this exchange occurs between the structures of DNA with no extensive sequence homology, it is termed nonhomologous recombination.

Unlike homologous recombination, it does not require a double-strand break in the DNA for initiation. It is also relatively imprecise and error-prone and often leads to the insertion or deletion of nucleotides at the site of recombination. NHEJ is a DNA repair mechanism that involves the direct ligation of the broken ends and thus does not require a homologous template strand. Whereas in the transpositions, the CSSR or VDJ recombinations and the phenomena of transposition also do not require extensive sequence similarity between the strands of DNA involved in the recombination process. So, normal homologous recombination does not require sequence similarity.

It actually happens abruptly, and it does not require the double standard breaks for initiation, either. So basically, the non-novelized recombinations will be an exchange of DNA between them with no extensive sequence similarity. There could be sequence similarity, but that would be very, very minor, and in these kinds of cases, you are actually going to have multiple examples, such as transpositions and VDJ recombination, which are responsible for the generation of different types of antibodies, and so on. And in all of these examples, there will be no sequence similarities involved. So we have the site-specific recombinations, which are called SSR.

So-called conservative site-specific recombination, or CSSR, is a combination of the two defined sequence elements of the DNA. This process is carried out by the protein known as recombinase, which brings together the specific ends of the DNA, forming the synaptic complex and resolving into three outcomes. So you can have the insertions, inversions, and deletions; these are the recombination sites. You are actually going to have the insertion, inversion, or deletion. The type of recombination in the CSSR is based on amino acid sequence homology and the mechanism followed; most recombinations can be classified into two types.

It can be a serine recombinase or a tyrosine recombinase. Serine recombinase cleaves all strands involved in the exchange, whereas tyrosine recombinase cleaves one strand at each site, leading to the formation of the Holliday junctions. So, tyrosine recombinase and serine recombinase are involved. The difference between tyrosine recombinase and serine recombinase is significant. So, the recombinase is going to be a cry, whereas the recognition site is going to be the lock site, and it is actually going to be utilized for the circularization of FASP1. We have the lambda integrase, which will be at the P and B sites, and it is an integration of the phage lambda.

Similarly, we have the XRD-CD resolvase complex, which is going to be different in the nucleoid or the ser of plasmids. It is a resolution of the dimers in E. coli. Similarly, for the serine recombinase, we have the H-in invertase, which is an HX site, and the inversion of the promoter in the salmonella. And then we have TA3 and Y sigma resolvase, which are rest sites, and they are going to be resolved during the replicative transpositions.

Mechanism of the serine recombinase. So one molecule of recombinase catalyzes the cleavage of a single strand. Thus, a total of four molecules of recombinase is required for the recombination process. The fifth hydroxyl group of DNA at the three prime acts as a nucleophile to attack the recombinase-DNA complex at the five prime to generate free recombinase and recombined DNA. The slippery hydrophobic part of the top and bottom halves of the recombinase dimer rotates by 180 degrees, and that's how this portion goes here and this portion goes there. So all four strands are first going to be broken by the serine recombinase.

So you are actually going to have four strands. One strand will have the 3' prime OH, the other will have the 5' prime OH, and so on. Then there will be a dimer that will rotate; in that case, this portion will rotate on this side, and this portion will rotate on this side. And that's how it is actually going to have recombination, and that's how you are actually going to have the exchange of material between the two strands. Then there will be electrostatic interactions stabilizing the initial and rotated states of the recombinase, and ultimately, there will be an exchange of DNA material between all four strands.

Then we have the mechanism of tyrosine recombinase. So, the recombinase cleaves one strand at each site of recombination. So they are going to cleave this particular site, and this is going to cleave the other site. So you're going to have the two strands that are actually participating in the recombination process. The cut strands are exchanged between the DNA molecules.

So this portion will go into this, and this portion will go into that. And that's how this will now be a part of this molecule, and this is going to be part of this molecule. Then we have the formation of the holiday junctions, and then we cut at the second strand on each side. Strand exchange occurs, resulting in the resolution of the Holliday junction. So once you have a cut on one strand, you will have a cut on this strand and then this strand, and then there will be an exchange between the cut extent; the cut extent will exchange between the two strands. And then there will be a formation of holiday functions, junctions, and you are going to have the cut at the other strands, and there will be an exchange of genetic material, and that is how you will have the recombination.

So, you see that here you have the blue, and then you are going to have the yellow, which is coming from this strand, and then you are actually going to have the red instead of the blue right in this particular strand. Similarly, for the lower strands, the lower strand is actually going to have blue, and it is also going to have red on the other side. Similarly, for this one, it has red on both sides. So, it is going to have red on both sides.

But on this side, it is actually going to have blue instead of red. And that is why both strands are now going to have the new DNA molecules or the recombined DNA molecules. What is the biological significance of CSSRs? So it has actually been involved in DNA repair. It is also involved in gene regulation. It's involved in genomic rearrangements.

It's also involved in the development of horizontal gene transfer. So in DNA repair, it allows for the precise cut and replacement of the damaged DNA strands, maintaining genomic integrity in E. coli. Whereas in gene regulation, the phage lambda uses a CSR mechanism to switch between the lytic and lysogenic cycles. Then the genome rearrangement, the CSSR causes gene rearrangement, leading to various combinations such as inversion, insertion, or deletion of the gene sequences. For instance, switching a flagellar component of a flagellum in Salmonella is mediated by the CSSR mechanism.

Then we have horizontal gene transfer. So bacteria often use this mechanism to acquire new traits from other bacteria or mobile genetic elements, leading to bacterial adaptation and evolution. Then we have the development where many organisms use the site-specific combination method to regulate tissue-specific gene expression and cell differentiation during development. Then we also have tracking of cell lineage during the development of Drosophila that was done using the FLP-FRT system. And then, CSR is widely used to manipulate and engineer the genomes of organisms, such as the ablation of gene function, induction of gene expression at a specific time during development, and so on.

Then we have another example, called VDJ recombination. So, VDG recombination is responsible for the production of antibodies, right? So, it is a specialized set of DNA recombination mechanisms that imparts enormous diversity in the B cell and T cell receptors. This recombination process occurs at specific sites on the V, D, and J segments in the genes for the generation of immunoglobulins and TCRs. These recombination sites are called recombination signal sequences (RSSs). There are two different types of RSS: 12-base pair RSS and 23-base pair RSS.

Recombination cannot occur at the same site of the RAG. So, this is a 23-base-pair RSS, and this is a 12-base-pair RSS. And this is the T-cell receptor. So what you can see here

is that the V, D, and J are recombining with different types of modules. So you have the different variations of the V component, the different components of the J component, and so on, and that's how they are actually recombining with each other to give you the different types of T cell receptors.

So this is one of the T cell receptors; this is another T cell receptor, and so on. Similarly, in the B cell receptor, B cells have different types of V, different types of J, and different types of D components. So, D, J rearranges DNA joints, and you will also see that V, D, J is giving you the different types of antibodies and different types of antigen-binding sites present in the antibody molecules. Now, another example is NHEJ, or non-homologous end joining. So NHEJ is a mechanism for repairing double-strand DNA breaks. Unlike homologous directed repairs, which require a homologous template, NHEJ directly ligates a broken end without the need for a template.

This makes NHEJ an efficient process that can operate in both dividing and non-dividing cells. The term non-homologous end joining was introduced by Maury and Haber in 1996. NHEJ is often guided by short homologous DNA sequences known as microhomologies. These sequences are typically found in a single standard overhang at the end of the double-strand break.

When the overhangs are compatible, NHEJ can accurately repair the breaks. However, if the overhangs are not compatible, imprecise repairs may occur, leading to the loss of the nucleotide. Inaccurate NHEJ repair can result in the loss of genetic material and may lead to translocation and telomerase fusion. These events are considered hallmarks of tumor cells, highlighting the importance of proper NHEJ function in maintaining genomic stability. NHEJ is a wide-step and exists in nearly all biological systems.

In mammalian cells, it is the predominant pathway for double-strand break repair. However, in budding yeasts like Saccharomyces cerevisiae, homologous recombination tends to dominate under common laboratory conditions. When NHEJ is inactive, double-stranded breaks may be repaired by an alternative, more error-prone pathway like microhomology-mediated end joining. In MMEJ, end resection reveals the short microhomologies on either side of the break, guiding the repair. Unlike classical NEGJ, MMEJ often results in the deletion of the DNA sequences between the microhomologies. So, what is the mechanism of NHEJ? So, you have double standard breaks, and then these double standard breaks are actually going to have the binding of the Ku7080.

So, the recognition of DNA ends by the Ku7080 heterodimers. So, they will go and bind to the DNA breaking sites, and then they will recruit the DNA PKCs to the nucleus, such as Artemis, to streamline the incompatible strands, and then the XRCC4 DNA ligates the

complete seal of the break, and then the DNA double-strand breaks are going to be repaired. Key players in the NHEJ and their associated mutation-related diseases. So you are going to have different types of diseases that are found if there is a mutation of the important components. You can have the Ku70, DNAPKCs, ArtemisCs, DNAPol, Ligase, XRCC4, and XLF. So if there is a mutation of XR70, it will actually be responsible for SID, lymphoma, and radiosensitivity.

If there are mutations in the ligases, then they are going to be for immunodeficiency, reduced growth and developmental issues, microcephaly, and malignancy. Then if there is a problem with XLF and other kinds of things, it is going to have embryonic lethality, syndrome, immunodeficiency, developmental delay, and microcephaly. Then let's move on to the next example, which is called transposons. So, transposons are also called "jumping genes," right? So, transposons are segments of DNA that can move from one locus to a relatively non-specific site in the genome.

Movement of a transposon can occur without duplicating the element. Accordingly, they are classified as class 1, or replicative DNA, and type 2, which is non-replicative DNA. DNA transposons carry inverted repeat sequences at their ends as recombination sites, flanking a recombinase protein called a transposase. which bring out the process of recombinations. Retrotransposons contain LDR and the genes for integrase and reverse transcriptase activity.

So there are different types of transposons. You can have the class 1 transposon, which is called a copy and paste, whereas the class 2 transposon is called a cut and paste. So, in class 1, you are actually going to have LTR transposons and retrotransposons. And the non-LTR retrotransposons. So when they are going to do a transcription, they will produce the messenger RNA, and then there will be a reverse transcriptase to produce double-stranded DNA, and then this double-stranded DNA will actually integrate into the target sites.

Similarly, for the non-LTR retrotransposons, there will be transcription; it will produce messenger RNA. Then, there will be target-primed reverse transcriptase, which will actually integrate into the target DNA. Similarly, for the class 2 cut and paste, you are going to have the DNA transposons, which will be excised from the other DNA, and then they will actually integrate into the target DNA. There are examples of transposons. You have the bacterial replicative transposon, which is called TN3 and F' mu.

Then you have non-replicative transposons, like IS elements and simple transposons. So IS1, IS2, IS50, IS10, TN1, TN7, TN501. Then we have a composite transposon, which is called TN5, TN9, or TN10. Then we have eukaryotic elements, such as HAT family

members. And so on, then we have a virus like transposons, such as the Ty element of yeast and the copia of Drosophila.

Then we have poly retrotransposons, which are the LINE and SINE in mammals or the Alu in humans. What is the outcome of the transpositions? So altered gene expression through insertion into the gene. For example, the ISI element can insert itself into a nearly 80-inch-long promoter sequence. IS elements can cause polar mutations and region-dependent recombinations. Site-specific combinations of the transposable elements in the same molecule can cause rearrangements such as deletions and inversions. And then the IS-1 element can cause deletions, while TN-3 and resolvase can cause the resolvase reaction toward the deletion and the mutation in general through the insertions.

Now, these are some of the approaches that have been very common or have been found when using the traditional methods of generating genetic recombinations. So, you can have homologous recombination, nonhomologous recombination, and so on. Now, when you are going to use these methods and trying to generate the genetic modifications, you are actually going to face a lot of limitations and you are going to have a lot of issues. What are these issues? These issues are that there will be a lack of precision, there will be low efficiency, it will be time-consuming, and there will be limited edition scopes and off-target effects.

These are very serious effects. You are going to lack precision, which means that if you use non-homologous recombination methods, they will be random. They can be on one side or they could be on another side. Even for homologous recombination, you also require very precise flanking sequences; only then can you edit the particular gene into the genome. But what happens is that even if all this depends on the sequence similarity, if the sequence similarity is unique, then it may actually be able to give you the.

Very precise removal of the mutated genes. But if there are any, you know, off-targets, right, if there is any kind of similarity between the sequences, then it may actually replace another gene or it may actually give you side effects. Similarly, all this depends on the integration of your externally supplied DNA into the system, and you can actually use the different types of DNA delivery methods to deliver the DNA to the site of action. That process is very inefficient. And on the other hand, once the DNA enters, it should actually go through the recombination process.

And during this journey, the DNA should not undergo any kind of DNA or other attacks. So that's why the efficiency of this recombination is going to be very, very low. Third, all of this requires the genetic cassettes to be prepared; then you are actually going to

transform that, and so it is actually very, very time-consuming. The fourth is because you are dependent on the flanking sequences. Or you are dependent on the cellular machinery, the scope is very, very limited; on the other hand, the first part is that the majority of the time when you are.

Transforming the cells with this externally supplied DNA, they may actually get integrated into the off-target sites. And because of that, it may not provide you with the desirable results. So these are some of the challenges that people are facing when they use traditional methods for genome editing. So what have we discussed? We have discussed the homologous recombinations. We have discussed the non-homologous methods.

Within homologous recombination, we have discussed the methods in both prokaryotic and eukaryotic systems. And within non-homologous recombination, we have taken an example of transpositions. We have discussed the VDJ recombinations and so on. So, with this, I would like to conclude my lecture. Thank you.