Molecular Biology

Prof. Vishal Trivedi

Department of Biosciences and Bioengineering

Indian Institute of Technology, Guwahati

Module - 11

Genome Editing

Lecture-48 Genome Editing (Part 2)

Hello everyone, this is Dr. Vishal Trivedi from Department of Bioscience and Bioengineering IIT Guwahati and what we were discussing? We were discussing about the genome editing in this particular module and so far what we have discussed in the previous lecture, we have discussed about the traditional approaches where we are using the homologous recombinations or the non-homologous recombination as an approach to modify the genome. Both of these approaches are having the many drawbacks and these drawbacks were you know corrected or will were improved very significantly in the modern approach. So, in today's lecture we are going to discuss about the modern approaches how you can be able to use these approaches to to you know to do the genome editing. So, as the name suggests the genome editing involves the alteration of the organism genetic code and I have you if you recall in the previous lecture, we have discussed very significantly that how these you know approaches are being used for curing the different types of diseases. Now, in the what was the major drawback of the previous approach was that in a homologous recombination or the homologous non-homologous recombination that they were having the off targets, they were expensive, they were time consuming and these things can be very much reduced with the modern approaches.

So, we have the you know couple of modern approaches what we can actually be able to use to do the genome editing and in the if you see how the genome editing is been evolved that in the beginning people were using the restriction enzyme for cutting the DNA a particular part of DNA and then using that for you know bringing the modification into the into the system. Then we have come up with the zinc finger nucleases and in the in the in the year of 1980s, the zinc finger nucleases were discovered and they were actually been engineered in such a way that they are actually going to be used for genome editing. Then we discuss about the TALENs. So, the TALENs are a structure similar to zinc finger nucleases, but they are been used for transcription activator like effector for DNA binding and they are also been used for DNA editing.

Then the modern approach is the usage of the CRISPR-Cas which I think you might have heard about. So, CRISPR-Cas is the latest approach what people are using for genome editing which involves a different types of protein and other kinds of factors. Then we have the base editing which is also been done with the help of the CRISPR and then we have the

primer editing approach which where you are actually using the Cas9 kind of protein and that is also been used for the genome editing. And then we also have the paste where you are actually going to have the different you know techniques to actually be able to use for the genome editing. So, these are the some of the edit approaches what people have used for genome editing.

What we are going to discuss today is we are going to discuss about the zinc finger nucleases, we are going to discuss about the TALENs and we are also going to discuss very briefly about the Cas-Cas9 method. And if you recall in the previous lecture also when I was discussing about the homologous recombination and the non-homologous recombination I said very clearly that these lectures are only to introduce you the topic. They are not going to be extensively been used for discussing the topic. Because there are very good genome editing MOOCs courses are available on to the NPTEL portal and you can actually be able to use or you can be able to follow them by getting the more detail about these approaches. So, in today's lecture we are discussing about we are going to discuss about the zinc finger nucleases, we are going to discuss about the TALEN and we are also going to discuss about the Cas-Cas.

And all these approaches if you see they are actually been having the you know I have been arranged in such a way that they are actually going to have the time and cost. So, time and cost is going to be more for the zinc finger nucleases and it is going to be less for the CRISPR-Cas. But as far as the feasibility is concerned the feasibility would be less, feasibility will be you know more for the CRISPR-Cas and it is going to be less for the zinc fingers. This means in any case the zinc finger the Cas phases the CRISPR-Cas approach is the best suited approach for genome editing. So, let us first discuss about the zinc finger nucleases, then we will discuss about the TALEN and then lastly we are also going to discuss very briefly about the CRISPR-Cas also.

So, zinc finger zinc finger proteins. So, zinc finger is a compact structural protein motif characterized by the binding of one or more zinc zinc ions contributing to the stabilization of its folded structure. The term zinc finger was first pointed based on the finger like appearances observed in the presumed structure of transcription factor 3a from the African frog Xenopus. However, it has been identified in a diverse area of protein structure within the eukaryotic cell. The discovery of zinc zinc protease is in the Xenopus transcription factor 3a essential for its function was reported in 1983 making marking the first document instance of the zinc requirement for a zinc regulation protein.

Subsequently, the similar findings were observed in the Cooper factor in the Drosophila and the zinc finger commons commonly manifest as a metal binding domain within the multi domain proteins. So, in a typical zinc finger proteins proposed a model for the TF3 proteins 30 kDa region featuring the 9 repeating region of 30 amino acid region the 25 of the 30 amino acid in the repeat fold around a zinc ion to form a small independent structure domain thus finger. And the 5 intervening amino acid provide the linkage

between the consecutive fingers. Each unit forms a zinc coordinate coordinated fingers stabilized by the conserved amino acid and a potential hydrophobic cluster. The rich the cluster the region is rich in basic and the polar residues are implicated in the nucleic acid binding.

So, zinc finger motifs or zinc finger nucleus is a is a small protein and this is a protein which has been designed where you are actually going to have the zinc finger motif in a you know 30 amino acid residue. And these are been repeated multiple times so that it can be actually have the binding of the zinc to these regions right. So, you see that in a in a particular one 30 nucleotide 30 amino acid residue you are actually going to have a zinc in the center and it is been coordinated by the many of the residue. Mostly the histidine what is been present on to the different helix and beta sheets which are actually going to be utilized for stabilizing the zinc what is present in the this particular motif. And this zinc particular motif is actually going to have the affinity for the DNA binding.

So, they are been actually been utilized or can be exploited for the genome editing applications. So, folding of the zinc finger relies on the tetrahedral coordination of the zinc ion with the two invariant pair of the cysteine and histidine forming a conserved cysteine-cysteine histidine-cysteine pattern. So, within this you are actually going to have a zinc in the center. Additionally each repeated includes the tyrosine or phenylalanine or phenylalanine 17 and elsines 23 large hydrophobic residues that may interact to create a stabilizing hydrophobic cluster into the compact finger module. The 30 amino acid repeat rich in basic and polar residues concentrated at significant number between the second cysteine and the first histidine emphasizing its role into the nucleic acid binding.

So, what you have here is that you are actually going to see that there is a zinc which is coordinated between the cysteine and tyrosine and histidine residues. In some cases you are actually not going to have the histidine you are also going to have the phenylalanine and the other hydrophobic residues like the leucine and this is kind of going to be an environment which is required for the zinc binding. The first zinc finger was identified in the 1980 for the transcription factor third A which gives rise to the discovery of new group of transcriptional activator protein with the 30 amino acid repeating region. This new class of protein was able to bind the specific sequence of the DNA. So, that is very very important that zinc finger motif or zinc finger is actually going to be utilized for recognizing the specific DNA sequence and that can be modified or that can be modulated in such a way that you can actually be able to use them for probing a particular DNA sequence.

The zinc finger structure is maintained by the zinc ion which coordinated cysteine and histidine in different combinations. Originally the zinc finger referred to a DNA binding motif in a xenophase, but now it includes various structural coordinated coordinating a zinc ion. Initially the classified by the number and order of the histidine residues recent classification focused on the fold groups determined by the overall shape of the folded domain. Common fold group like include the cysteine, cysteine like classical zinc finger and

the zinc ribbon. So, cysteine histidine to zinc finger motif consist of an alpha helix and anti parallel beta sheets the zinc ion is coordinated by the two histidine residues and the two cysteine residues.

That is what I have shown before also that you are going to have a zinc residue which is coordinated between the two histidine molecules and the two cysteine molecules. The cysteine histidine like fold group is by far the best characterized class of zinc finger and it is commonly common in mammalian transcription factor such domain adopt a simple beta beta alpha fold and have the amino acid sequence motif. So, in the amino acid sequence motif you are going to have the X 2, cysteine X 2, then cysteine X 12, histidine and the histidine. The specific amino acid interaction between the zinc finger and the DNA was elusive until 1991. A pivotal movement occur when the scientist successfully deciphered the crystal structure of complex involving a DNA oligonucleotide and the three finger zinc DNA binding domain of the mouse transcription factor ZIF 268.

In this structure the alpha helix play a primary role binding to the DNA major group through the precise hydrogen bonding interaction from the amino acid at a helical portion 1, 3 and 6 to 3 consecutive bases on the strand of the DNA. Despite this strand configuration there are notable variation in the sequences. Consequently the zinc finger motifs are considered excellent natural building block for the DNA for the design of the protein tailored to recognize a specific DNA sequence owing to their adaptability and versatility. So, that is a very very important portion that because of this kind of zinc finger coordination between the histidine and cysteine it is actually going to be very useful in you know exploiting to detect a specific DNA sequence. So, far there was no protein available which actually recognizes a specific DNA sequence.

So, and that was very very important that you actually can be able to target a specific DNA sequence utilizing this particular type of zinc finger. And once the zinc finger is been discovered the people have also started developing the zinc finger nuclease. So, the zinc finger nuclease are synthetic restriction enzyme created by combining a zinc finger DNA binding domain with the DNA cleavage domain. So, that is what it is actually happening. So, in the in a typical zinc finger nucleus what you have is a DNA binding domain which is from the zinc finger right where you are actually going to have the zinc coordinated into the histidine and cysteine and then you just put actually a DNA cleavage domain.

So, it is like a restriction enzyme which is actually going to chew a particular type of part of the DNA. The engineered zinc finger domain can be customized to recognize a specific target sequence allowing the zinc fingers to pinpoint the unique sequences within the integrate genome. Leveraging the inherent DNA repair mechanism in cells these agents are employed to accurately modify the genetic makeup of more complex organisms. Alongside CRISPR-SCAS9 and the TALEN ZFN stand out as a significant instrument in the relief of the DNA genome editing. So, what you are going to do is you are going to have the two tools actually you are going to have a DNA binding domain you are also going to have a DNA

cleavage domain right and utilizing them you can be able to take out a part of DNA right and once you take out this part of DNA it is actually going to be cured in due course of time and that is how it is actually going to give you the modification of that particular sequence.

So, in a DNA binding domain especially has 3 to 6 zinc finger repeats recognizing 9 to 18 base pair utilizing 3 finger array specific 9 base pair target site can be addressed. Similarly, you can actually be able to have the zinc finger nucleases which are actually going to be having a non specific globase domain typically type 2 restriction enzyme like the FOCI requiring the dimerization for the DNA cleavage. Standard zinc finger nucleases link this domain to the zinc finger domain and for the effective cleavage the two zinc finger nucleates must bind opposite DNA strand with a specific spacing. Protein engineering technique including the direct directed evolution and structure based design aim to improve the nucleus domains activity and facility in the zinc finger nucleus. Directed evolution produced a variant called as the SHARKEY while structure based design enhances cleavage specifically by modifying the dimerization interface.

So, what happen is that in a in a in a typical zinc finger nucleus what will happen is that it is actually going to utilize the 3 or 4 the zinc finger domain and they are actually going to be utilized a particular stretch of the nucleotide. And once that happens then the zinc finger cleavage domain will actually go and bind to the opposite side of the DNA and it is actually going to cleave and that is how you are going to have the two part right. And then these parts are actually going to be utilized for and will will participate into DNA repair and ultimately it is actually going to lead to the genome editing. Now, what is the application of the zinc finger nucleus? So, zinc finger nucleus as I said you know it is going to be utilized for the genome editing. So, you are what you are going to do is you are going to have the zinc finger nucleus you are also going to have a DNA cleavage domain and then you are going to utilize that for bringing the cleavage into the genome right.

And then this will go for the genome genome repairing right. So, it is actually going to activate the repair mechanisms. So, repairing could be by the non homologous recombinations or by the non homologous recombinations and once that happens it is actually going to bring the changes into the corresponding or the resulting genome. Now, let us move on to the next tool and next tool is the TALEN, Transcriptional Activator Like Effector Nucleus. So, TALEN and zinc finger are actually going to have the approach wise they are same actually.

In the in this case you are using the zinc finger domain as a weapon to identify the DNA where you are actually going to use the transcriptional factors. So, the TALEN as the name suggests transcription activator like effector plus nucleus right. So, transcription activator like effector nucleus are type of engineered restriction enzyme designed to target the specific DNA sequence for a cleavage. They are created by combining a TAL effector DNA binding domain with a DNA cleavage domain which act as a nucleus capable of cutting the DNA strands. By manipulating the transcription activator like effectors, scientists can

customize their binding to the virtually any desired DNA sequence.

When these engineered TALES are coupled with nucleus precise DNA cleavage occur at a predetermined locations. This engineered restriction enzyme can be introduced into the cell for applications such as genome editing or in-situ genome editing or the method known as the genome editing with a engineered nucleus. TALEN alongside zinc finger nucleus and KSPER9 stand out is the prominent tools in the field of genome editing. So, TAL or the Transcription Activator Like Effectors or TALES are the protein released by the certain beta and gamma proteobacteria which the being a prominent group particularly plant pathogenic xanthomonas bacteria utilize the TALES produced through the type 3 secretion system. These proteins have a capability to attack to the promoter sequences within the host protein stimulating the expression of plant gene that facilitate the bacterial infection.

Transcription Activator Like Effector protein contains a nuclear localization sequence or NLS, N-terminal localization sequence or the TS, Transcription Activator Domain AD and the central repeat domain or the repeat variant domains. So, the primary distinguishing feature of TAL effector is a central repeat domain consist of 1.5 to 3.5 repeat each typically contain the 34 received long. So, you have a N-terminal residue and then you are owing to have the nuclear localization sequences the N-terminal local translocation signals transcriptional activator domain and central repeat domain.

And in the central you are going to have a central repeat domain and this central repeat domain is unique to the TALES. The standard repeat sequence is this sequence and which hyper variably residue at the 12 and the 13 position known as the repeat variable dye residue or the RVD. So, this is the region this is the place where it is actually going to show you the variation. The identity of the two residue in consistently repeat correspond to the sequential DNA bases in the TALES effector target site. According to the crystal structure each repeat comprises two alpha helices and two short RVD containing loop.

In this loop the second residue of the RVD establishes the contact while the first residue stabilizes the RVD containing loop. TALE effector target site commonly include a thymine next to the 5 prime base target by the first repeat. Attributed to the contact between the thymine and the constructed tryptophan in the region N-terminals of the central repeat domain. So, this central repeat domain is going to have this particular sequence and within this the nucleotide number 13 or the residue number 13 and 12 and 13 is very very important because it is actually going to repeat and it is going to be a part of the RVD and that is actually going to decide the sequence identity or the sequence specificity. The experimentally validated code between the RVD sequence and the target DNA bases can be expressed as follows.

So, if you have the A, C or all those kind of thing then it is actually going to have the 5 methyl cytosine and so on. So, on the other hand it is also going to have the C-terminal domain which is a functional domain and is going to act as a activator or repressor or

nucleases or methylases and integrases. So, TALEN will have the 2 domains like first is the TAL DNA binding domain and the second is the DNA cleavage domain. This is exactly the same as what we have discussed for the zinc finger nucleases right. So, TAL effectors the DNA binding domain comprises a repeated 33 to 34 amino acid sequence with variable 12 and 13 base pair amino acid known as the repeat variable digestive or RVD.

This RVD variation correlates strongly with the specific nucleotide recognitions allowing the engineering of the customized DNA binding domain by the selecting the repeat segment with the desired RVDs. Minor alteration in the RVD and the incorporation of non-conventional RVD sequence can enhance the target. So, you can actually be able to altered the DNA of the TAL simply by modifying the RVD sequences or simply incorporating the particular type of RVD sequence. Then we have the non-specific DNA cleavage domain from the foci endonuclease is utilized to construct the hybrid nucleases demonstrated activity in each plant and animal cell. The foci domain operate as a dimer requiring the two construct with the distinct DNA binding domain for the target genome site with appropriate orientation and spacing.

The distance between the tail DNA binding domain and the foci cleavage domain as well as the spacing between the individual TAL domain binding site a crucial factor influencing the activity level. So, this is what exactly happened you have a TAL which is actually going to have the effector function and you are also going to have foci nucleases. So, what you have done is you have fused these two and after the fusion it is actually going to be TAL nucleus right which is actually going to have the DNA binding domain it is going to have the DNA cleavage domain. Now, what will happen is that this portion is actually going to utilize for recognizing a particular DNA sequence right. So, it is not only the DNA, but it also going to recognize a particular DNA sequence and how that happens that happens because of the RVD region where you are actually going to.

So, if you are changing the RVD region within this the central domain it is actually going to recognize a particular DNA sequence and then you also have the DNA cleavage domain. So, it is actually going to cut the DNA and then once it is been cut it is going to go through with the homologous or non homologous recombinations and ultimately it is actually going to give you the altered genome. Now, let us talk about the applications. So, applications of the TALENs are going to be the same as what we have discussed for the zinc finger nucleus and they have very very diversified applications in the various fields of the biotechnology. Then we have the third target or the third tool that is the CRISPR-Cas and CRISPR-Cas is the one of the most versatile and most economically very feasible tool which people can use.

So, CRISPR-Cas is a kind of a immune response. So, CRISPR-Cas is called as clustered regulatory interspaced short palindromic repeats and CRISPR-Cas systems are sophisticated adopted immune system employed by the prokaryotes. CRISPR is a DNA sequence found in the prokaryote like bacteria and archaea originating from the previous bacteriophage infection and served as a adoptive immune system allowing the cell to detect

and element the similar threat. It is present in about the 50 percent of the bacterial genome and nearly 90 percent of the archaea genome. CRISPR consists of the conserved repeat sequences interspersed with the short spacer sequence. CRISPR associated protein or Cas is a enzyme that uses the CRISPR sequences as a guide to recognize and open up of the specific strand of the DNA that are complementary to the CRISPR sequences.

So, CRISPR is a kind of a immune response which bacteria is using to fight against the viral infections. And it is been discovered by the two scientists which are actually been offered the Nobel Prize into the year of 2022, 2020. So, these two scientists have been offered the Nobel Prize in the year of 2020. And there is a long list of the history how the CRISPR-Cas system was evolved. So, it started from in the first time as a 1987 the people have when the people have observed this particular phenomena into the E.

coli. And then ultimately it is been end up in 2024, 2014 when the first CRISPR patent was granted to the feng jiang. So, that is how it is a complete systematic way in which the scientists have discovered the different aspects of the CRISPR-Cas and that is how the complete tool was evolved and this tool can be used for genome editing. So, CRISPR features the short conserved repeat sequences intersect with the similarly sized spacers ranging from the 23 to 47 base pair for repeat and 21 to 72 base pair for spacer. Bacterial genome features multiple CRISPR loci each with the diverse and hyper variable spacer sequence even among the closely related strain. This unique spacer sequence originates from the viral or plasmid DNA and serve as a recognition element.

The introduction of the new spacer allow the system to recognize and eliminate the matching viral or plasmid genome. The CRISPR loci include a conserved leader sequence and their activity depend on the adjacent CRISPR associated Cas genes which encode the essential Cas like protein like the nucleases helicases and polymerases working collaboratively for a CRISPR immune system. So, this is what is the combination you have the CRISPR gene Cas here you have a leader and promoters and then these are the CRISPR repeat spacer array and all these are being utilized for identifying the genome or the DNA sequence from the virus. So, that in the general mechanism you have the 3 events one is the adaptations you have the expression and the maturations and the third is the interference or the targeting. So, in the first event that is the adaptation the CRISPR Cas is actually in the adaptation stage the Cas 1 and Cas 2 complex acquires a proto spacer from the invading viral DNA and integrate it as a new spacer into the CRISPR array.

So, this is what exactly happened that when the virus is going to infect the bacterial cell it is actually going to inject its genome and that genome is going to be identified by the CRISPR Cas 1 and Cas 2 system and it is actually going to generate the proto spacer and that proto spacer is going to integrate into the into the into the CRISPR array and once that happens it is actually going to you know enter into the next phase and that next phase is the expression and the maturation phase. So, once it enter into this in the expression and maturation phase the CRISPR array undergoes the transcription and it is processed into the

mature CRISPR RNA each contains the transcription spacer and the part of basically repeat these CRISPR RNA form the nucleoprotein RNP complexes with the Cas proteins. And then we have the third phase and third phase is called an interfero phase. So, in the interfero phase the CRISPR RNA Cas RNP complex identify the invading target DNA through the complementary base pair cleaving the target sequence leading to a destruction of the invading viral genome providing the protection against the viral infection into the bacteria.

So, there are different types of CRISPR Cas system. So, you have the class 1 CRISPR Cas you have the class 2 CRISPR Cas and in the within the class 1 you have the type 1 type 2 type 3 and type 4 whereas, in the Cas 2 you are actually going to have the type 2 type 5 and type 6. And all these classification is very very complex right. And the class 1 system utilizes the multiple Cas protein while the class 2 protein involved a single protein for the interference. So, in the class 1 CRISPR Cas you are actually going to have the you know the you are going to have the multiple Cas proteins right. So, you are going to have the Cas 3 you are going to have the Cas 10 you are going to have the other kinds of CRISPR Cas proteins and these class 1 protein and class 1 CRISPR Cas system is going to be further subdivided into like type 1 type 3 and type 4 and 12 subtypes.

And these class 1 subtype system is found in the 90 percent of the CRISPR loci in bacteria and archaea and can target both the DNA and RNA. So, this is the type 1 where you are going to have the multiple Cas proteins like the Cas 6 and all that and it is actually going to you know target the CRISPR RNA and then it is actually going to enter into the targeting phase where the Cas 3 is actually going to target the particular DNA and it is going to induce the target DNA degradations. So, in a type 1 CRISPR the cascade complex and the Cas 3 nucleus play key roles. The process involves the cleavage of the pre CRISPR RNA by the Cas 6 generating the CRISPR RNA. This CRISPR RNA associated with the cascade identify the proto spacer in a target DNA.

Cas 8 a cascade subunit recognizes the proto spacer adjacent motif. CRISPR PAM function is crucial for a type 1 cascade Cas immune sequence. This functional PAM hinders the CRISPR RNA recognitions impeding the R loop formations. Functional CRISPR system activate the Cas 3 leading to its nick in the target single standard RNA single standard DNA and subsequent degradations. And then we have the type 3 in the type 3 the Cas 6 uses endonuclease endo ribonuclease mechanism to create the CRISPR RNA from the pre CRISPR RNA by cleaving it.

Unlike earlier CRISPR model this type introduces 8 nucleotide repeat sequence called the CRISPR RNA tags through this Cas 6 mediated cleavage. Even downstream to the spacer sequence the CRISPR tag into the 6 nucleotide increasing the size of the CRISPR complex. This is dealt in the Cas 10, CRISPR formation in 2, Cas 3 and Casper complex to B. Type 3 targets both DNA and RNA leading to the coarse transcriptional CRISPR RNA guided cleavage of the target DNA. The PAM domain of the CRISPR Cas 10 cleaves the DNA strands while CSM 3 and the CSR CMR 4 cleaves the RNA transcripts.

And important distinction is that the PAM is not is necessarily essential for the type 3 to initiate the immune mechanisms. Then we have the class 2 CRISPR Cas. So, in a class 2 CRISPR Cas characterized by the presence of a single effector molecule. There are 3 types of class 2 systems type 2, type 5 and type 6 and 9 subtypes. While the class 2 systems are more commonly known as Cas 9 is a class 2 system.

They only represent the 10 percent of the CRISPR loci and unlike class 1 they are found in the bacteria. Class 2 system can target both DNA and RNA depending upon the type. Type 2 the most common class 2 system find in type 2. The type 2 systems are characterized by the presence of Cas 9 as well as the ancillary protein like Cas 1 and Cas 2. Cas 9 the commonly used genome engineering endonuclease is a type 2 system.

Type 2 system requires the track RNA for the function. Type 5 system commonly know use the Cas 12 as their endonuclease of choice. Like Cas 9 Cas 12 target the DNA for the editing. Similar to type 2 system the type 5 also require the track RNA for the function. Type 6 is only the last 2 system that target the RNA for editing. So, class 13 is a type of is a type 6 endonuclease that enable the editing of the RNA.

So, then in a type 2 you are going to have the most prominent Cas protein that is a Cas 9 which is actually going to be a part of the type 2 Casper Cas system. And then these are the this is just a summary of the class 1 and class 2. So, in a type class 1 you have a type 1 type 3 and type 5 type 4. Whereas, in the class 2 you have the type 2 type 5 and type 6. In the class 1 you are going to have the signature Cas protein like the Cas 3 Cas 10 and CSF.

Whereas, in the type 2 class 2 you are going to have the class 9 class Cas 12 and Cas 13. So, Cas 9 is very common for genome editing and the Cas 13 is for the RNA. So, these 2 are being used very extensively for editing the genome of the organisms where or the DNA of the organisms whereas, this is being used for the editing of the RNA. So, in a CRISPR Cas 9 system, the CRISPR Cas 9 system particularly the type 2 A is widely studied Cas 9 or 160-kyridol-dal-10 DNA endonuclease is the sole protein needed for the interference. It works in a tandem in the single guide RNA, a fusion of CRISPR RNA and the TRAC RNA alongside the Cas 1, Cas 2 and CSN and Cas RNA 3 for DNA acquisition and processing of pre CRISPR RNA to the mature single SgRNA.

In 2012 this artificial Cas 9 system programmable for any DNA sequence revolutionize the DNA editing. Further research demonstrate the programming capability of Cas 9 from the S-thermophilous Casper system enhancing its application in the target genome editing. So, Cas 9 is actually going to be utilized very extensively into the CRISPR Cas 9 into the type 2 CRISPR system. Then we have the CRISPR Cas 9 system consist of 2 lobes. So, CRISPR Cas 9 system consist of 2 lobes in the Cas 9 proteins, the recognition and the nucleus lobe.

So, in the recognition lobe end you have the residue 62178, 718 and the nucleus lobe

which is actually going to be from 1 to 60 and 719 to 1368. The recognition lobe contain the Rec I, Rec 2 and the domain responsible for the nucleotide recognition. The arginine rich bridge helix act as a linker between the RuVc 1 and the Rec domain playing a crucial role in initiating the cleavage activity upon binding to the target DNA. The nucleus lobe contain the 2 endonuclear domain the HNH domain and the RuC domain. And the HNH domain cleaves the target DNA strand while the RuC domain cleaves the non target DNA strands.

The PAM interacting domain confers the PAM necessity and initiate binding to the DNA upon DNA binding the positively charged residue in the interface between the Rec and NUC lobes specifically at the bridge helix stabilizes the negatively charged SgRNA DNA hybrid. Additionally, positively charged residue in the linker range region between the RuC and NHC domain contribute to stabilizing the displaced non target DNA. So, this is the structure of the Cas 9 where you have the different types of domain you have HNH domain you have Rec 2, Rec 1, RuC 1 and PAM interacting domain and all of these are participating into the CRISPR Cas 9 mediated genome editing. So, what is a guide RNA? So, the these guide RNA take one of the two forms a synthetic trans activating CRISPR RNA or TRAC RNA plus a synthetic CRISPR RNA or CRISPR-CR RNA designed to cleave that gene target site of the interest. A synthetic or the expressed RNA guide DNA sorry RNA a single guide RNA a gRNA that consists of the both the CR RNA and TRAC RNA as a single construct.

So, you have the wild type Cas 9 and you have the modified Cas 9. So, modified Cas 9 is being utilized. So, in the modified Cas 9 you are actually having a D10 to A mutation rendered the RuC domain inactive causing this nickase to exclusively cleave the complementary or target RNA strand. On the other hand the H840A mutation in the HNH domain result in a nickase that cleaves the non complementary or non target DNA. Similarly, you have the dCas9. So, point mutation of D110A and the H840A can deactivate both RuC and HNH nucleus domain leading to a nucleus dead Cas 9 or the dCas9 molecule incapable of cleaving the target DNA.

However, the dCas9 molecule maintain its capacity to bind the target DNA through the guide RNA targeting sequences. So, this is the Casper Cas 9 of the class 12A. So, this is the class 12A system. So, you this is a comparison of the Casper Cas 9 and the Casper CPF 1. So, this is big in size this is complex and Cas 9 required the CR RNA and the TRAC RNA and Cas 9 contain the two nucleus domain HNH domain and RuC nucleus domain.

So, it is 5 prime NGG where N is the any nucleotide and G rich sequences. In this case it is a small in size it is simpler it is CPF 1 require only the CR RNA and CPF 1 contain only one nucleus domain or the RuC domain. And the recognition sequence is also different in this case you have the NGGG where N is any nucleotide whereas, in this case it is TTTN where N is any nucleotide whereas, T is a T rich spam. So, these are the some of the differences between the CR Cas 9 versus CR CPF 1. So, CRISPR is a very very you know robust tool for genome editing and it has the enormous applications every day there are people who are

utilizing or there are scientists who are utilizing this for genome editing applications. So, it is been employed in you know in genome editing of the plants or virus or the even for the human being also there are reports from the China that it is been utilized for genome editing of the some of the you know new born babies to make them resistance against a particular disease.

So, the CRISPR Cas is a very very robust tool for genome editing and we are going to discuss more about the application of the CRISPR Cas as well as the zinc finger nuclease and the TALENs in our subsequent lectures. So, with this I would like to conclude my lecture here you know subsequent lecture we are going to discuss some more aspects related to molecular biology. Thank you. Thank you.