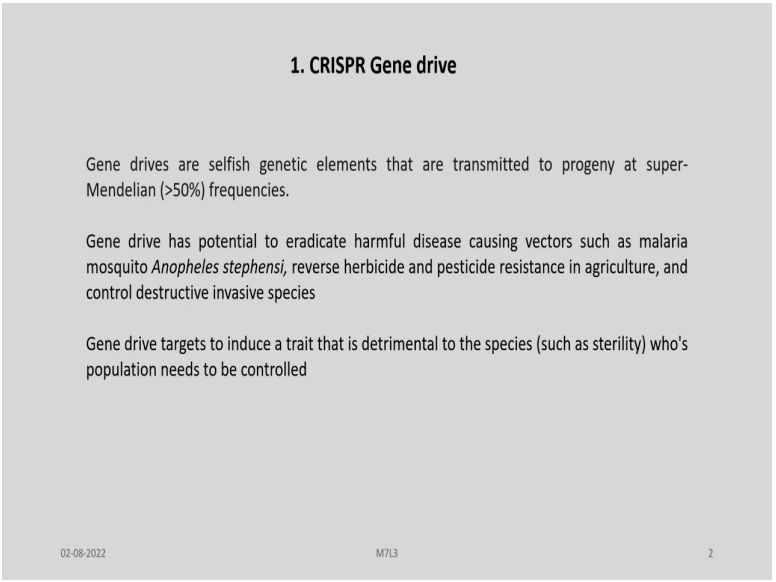
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 07
Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 technology
Lecture - 03
Applications of CRISPR/Cas9 Part-A

Welcome to my course on Genome Editing and Engineering. We are discussing module 7 which is on CRISPR Cas9 Technology. In lecture number 3, we are going to discuss about the various Applications of this CRISPR Cas9 Technology. To begin with, we are going to discuss about the CRISPR Gene Drive.

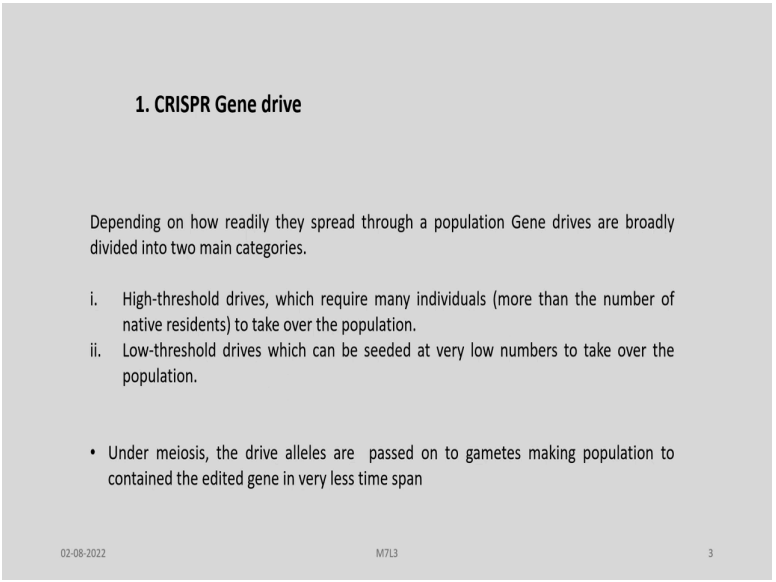
(Refer Slide Time: 00:54)



So, what are gene drives? Gene drives are selfish genetic elements that are transmitted to progeny at super Mendelian frequencies, which are more than 50 percent when counted on the progenies. We will discuss about the gene drives with illustrative examples in the next slide. But gene drives have various applications like the potential to eradicate harmful disease causing vectors such as malaria mosquito *Anopheles* and reverse herbicide and pesticide resistance in agriculture and control destructive invasive species, particularly weeds.

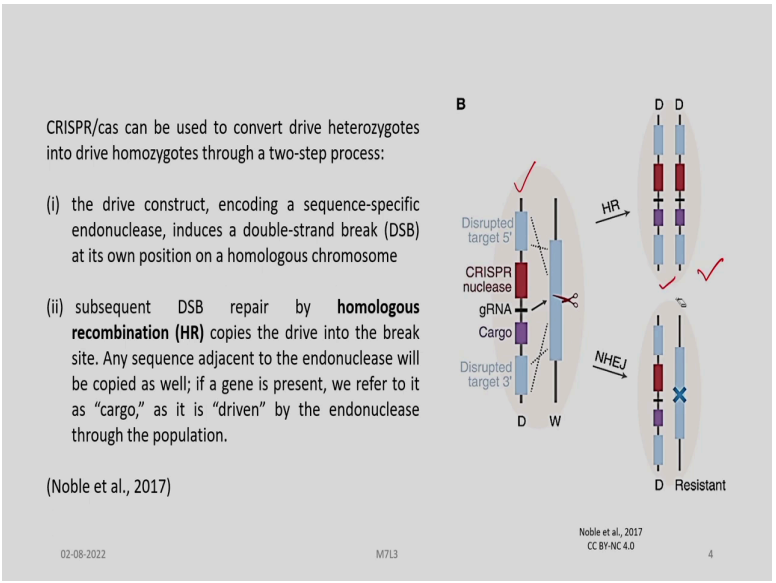
Gene drive targets to induce a trait that is detrimental to the species such as sterility whose population needs to be controlled. For example, in a malaria vector control program if we are able to induce sterility into the mosquitoes, we can very easily control their population.

(Refer Slide Time: 01:50)



CRISPR gene drive depending on how readily the gene drives spread through a population they are broadly divided into two main categories. Number 1 is the high-threshold drives which require many individuals, more than the number of native residents to take over the population. Then, we have the low-threshold drives, which can be seeded at very low numbers to take over the population. Under meiosis, the drive alleles are passed on to gametes making population to contain the edited gene in a very less time span.

(Refer Slide Time: 02:28)

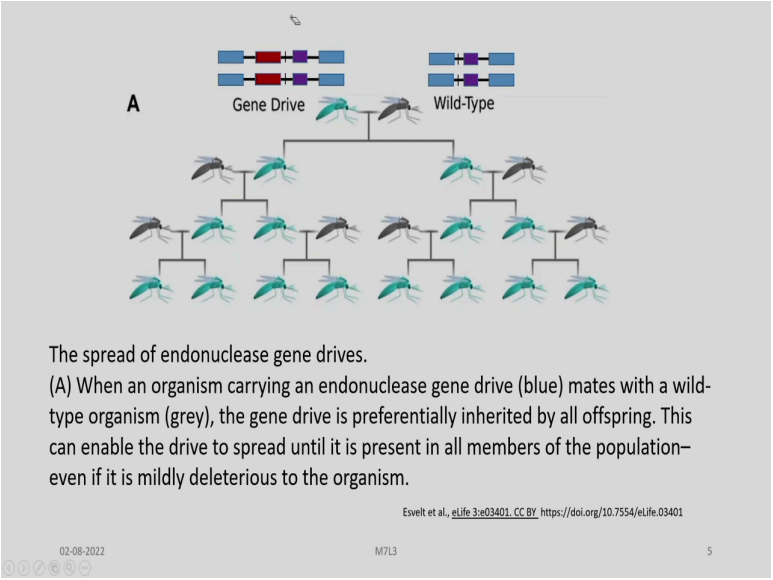


So, let us see exactly how a gene drive works and how CRISPR Cas9 is used for generating populations with gene drives. So, with CRISPR Cas9 we can convert drive heterozygotes into drive homozygotes through a two-step process. We have the drive construct here you can see which encodes a specific endonuclease which is specific to a particular sequence and these induces a double-strand break at its own position on a homologous a chromosome.

Now, it may undergo two pathways of repair either by homologous recombination, where the copies of these drive are copied into the break site. Any sequence adjacent to the endonuclease will also be copied as well; if a gene is present, we refer to it as a "cargo", as it is "driven" by the endonuclease through the population. So, here you can see the homologous recombination due to which the copy becomes available in both the chromosomes, but in NHEJ that does not happen.

So, this population will remain in its wild-type state, but of course, in a heterozygote state, but in this case homozygosity is attained.

(Refer Slide Time: 04:18)

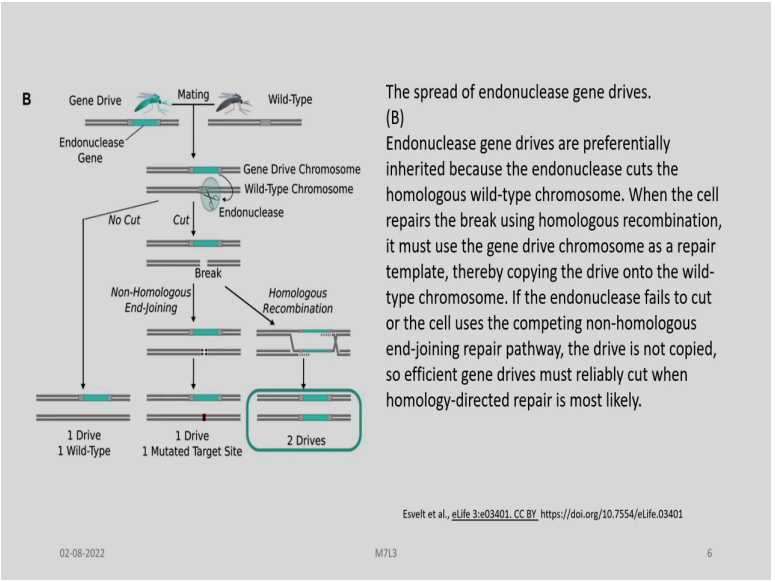


So, for example, we have a population here with a gene drive and we have a population here with wild-type. And, when we meet these population having the gene drive with a population having the wild-type as per Mendelian ratio we expect the 50 percent inheritance or distribution in both the population as you can see over here. But, as the generation passes you can see at the end it does not any longer get inherited as per the Mendelian ratio, but the entire population becomes homozygous for the gene drive.

So, that is why we tell that gene drives are selfish DNA and this is the super Mendelian ratio that we were referring to in the earlier slides. So, how the spread of the endonuclease genes in a gene drive happen? When an organism carrying an endonuclease, for example, here shown by the blue mates with a wild-type organism, the gene drive is preferentially inherited by all offspring. This can enable the drive to spread until it is present in all members of the population – even if it is mildly deleterious to the organism.

And, this skewed ratio is due to the CRISPR Cas9 which helps in creating double-strand breaks and also participates in the mobilization of the gene to the other chromosome.

(Refer Slide Time: 06:04)

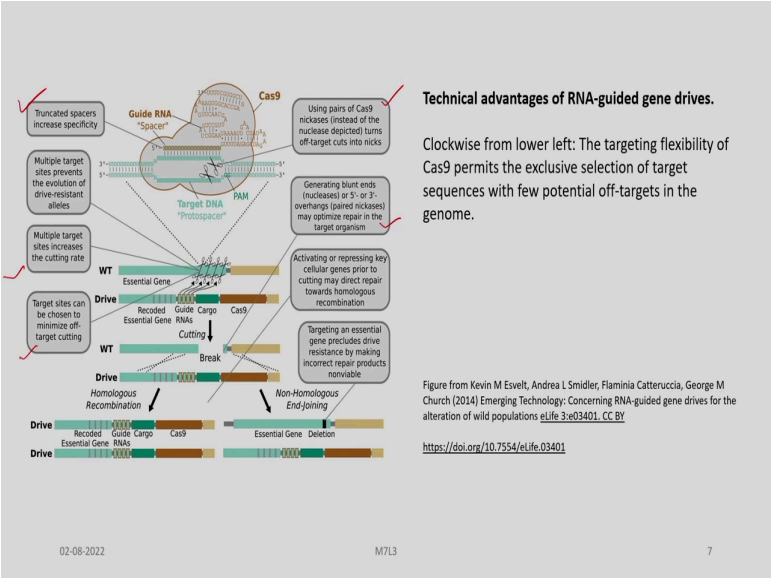


You can see here we have a gene drive population with a endonuclease gene over here and here is the wild-type population. So, the gene drive chromosome you can see here and the wild-type chromosome see here and the endonuclease X on the wild-type chromosome partner, it initiates cuts. And as already discussed, it can either undergo homologous recombination or non-homologous end joining.

In the case of homologous recombination, we have the drives now present in both the copies of the chromosome and if there is no cut induced, we will have only one drive and one wild-type. And in the case of non-homologous end joining we have one drive and one mutated target site as shown over here. The endonuclease gene drives are preferentially inherited because the endonuclease cuts the homologous wild-type chromosome.

When the cell repairs the break during homologous recombination, it must use the gene drive chromosome as a repair template, thereby copying the drive on to the wild-type chromosome. If the endonuclease fails to cut or the cell uses the competing non-homologous end joining repair pathway, the drive is not copied, so, efficient gene drives must reliably cut when homology-directed repair is most likely to happen.

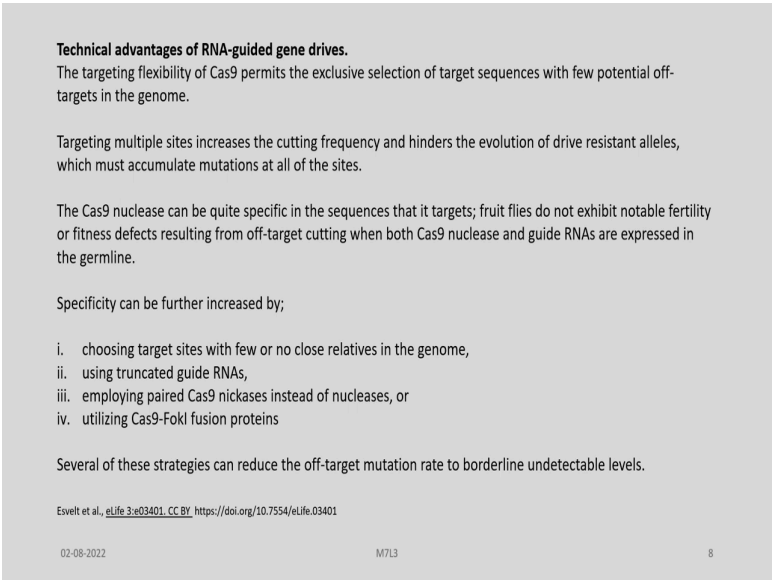
(Refer Slide Time: 07:35)



We have certain technical advantages of RNA guided gene drives. So, you can see here the target sites can be chosen to minimize off-target cutting. We can have multiple target sites which increases the cutting rate, then the multiple target sites prevents the evolution of drive resistance alleles. Then we have truncated spacers which increases the specificity and using pairs of casting decays instead of the nuclease as shown here, turns off-target cuts into nicks.

Then, generating blunt ends nuclease or 5' or 3' overhangs, may optimize repair in the target organism. Activating or repressing key cellular genes prior to cutting may direct repair towards the homologous recombination. Then targeting an essential gene precludes drive resistance by making incorrect repair products non viable. So, with so many advantages offered by RNA guide gene drives many experiments has been conducted with potential applications.

(Refer Slide Time: 08:54)



The targeting flexibility of Cas9 permits the exclusive selection of a target sequence with few potential off-targets in the genome. Targeting multiple sites increases the cutting frequency and hinders the evolution of drive resistant alleles which must accumulate mutations at all of the sites. The Cas9 nuclease can be quite specific in the sequence that it targets; fruit flies do not exhibit noticeable fertility of or fitness defects resulting from off-target cutting when both Cas9 nuclease and guide RNAs are expressed in the germ line.

So, how can we increase specificity further? Number 1 – by choosing target sites with few or no close relatives in the genome; secondly, using truncated guide RNAs; thirdly employing paired Cas9 nickases instead of nucleases or fourthly utilizing Cas 9 Fok1 fusion proteins. Several of these strategies can reduce the off-target mutation rate to undetectable levels.

(Refer Slide Time: 10:05)

Technical advantages of RNA-guided gene drives.

The frequency at which the drive is correctly copied might be increased by using Cas9 as a transcriptional regulator to activate HR genes and repress NHEJ genes.

By choosing target sites within an essential gene, any non-homologous end-joining event that deletes all of the target sites will cause lethality rather than creating a drive-resistant allele, further increasing the evolutionary robustness of the RNA-guided gene drive.

Other options include using distinct promoters and guide RNAs to avoid repetitiveness and increase stability or employing newly characterized, engineered, or evolved Cas9 variants with improved properties.

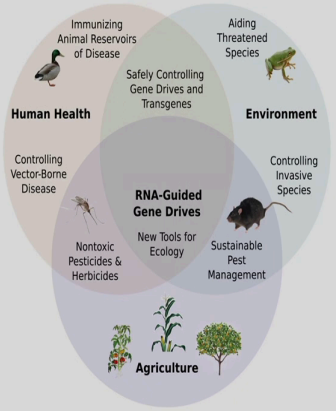
Esvelt et al., eLife 3:e03401. CC BY https://doi.org/10.7554/eLife.03401

The frequency at which the drive is correctly copied might be increased by using Cas9 as a transcriptional regulator to activate HR genes and repress NHEJ genes. By choosing target sites within an essential gene, any non-homologous end joining event that deletes all of the target sites will cause lethality rather than creating a drive-resistance allele, further increasing the evolutionary robustness of the RNA-guided gene drive.

Other options include using distinct promoters and guide RNAs to avoid repetitiveness and increase stability or employing newly characterized, engineered, or evolved Cas9 variants with improved properties.

(Refer Slide Time: 10:50)

Potential applications of RNA-guided gene drives.



Disease vectors such as malarial mosquitoes might be engineered to resist pathogen acquisition or eliminated with a suppression drive.
Wild populations that serve as reservoirs for human viruses could be immunized using Cas9, RNAi machinery, or elite controller antibodies carried by a gene drive.
Reversal and immunization drives could help ensure that all transgenes are safe and controlled. Drives might quickly spread protective genes through threatened or soon-to-be-threatened species such as amphibians facing the expansion of chytrid fungus.
Invasive species might be locally controlled or eradicated without directly affecting others. Sensitizing drives could improve the sustainability and safety of pesticides and herbicides.
Gene drives could test ecological hypotheses concerning gene flow, sex ratios, speciation, and evolution.

Figure from Esvelt et al., eLife 3:e03401. CC BY
https://doi.org/10.7554/eLife.03401

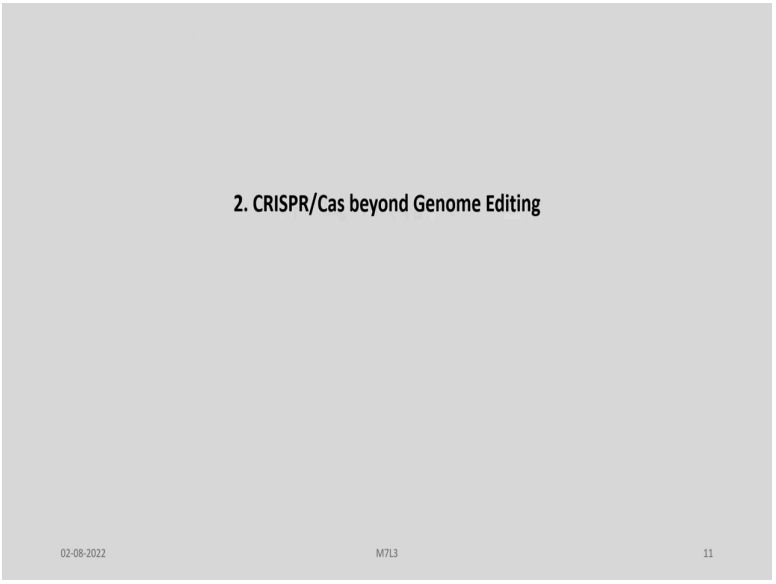
So, there are various areas in which the RNA-guided gene drives can be used. For example, in agriculture, in human health and in environmental areas. So, in agriculture we can use these drives for sustainable pest management or non-toxic pesticides and herbicides. And, in case of human health we can go for controlling the vector-borne diseases. As we have already told about say malaria control by controlling the vectors.

Then immunizing animal reservoirs of diseases, particularly those who causes zoonotic diseases. Then in the case of environment, we can use these for aiding threatened species and also for controlling the invasive species particularly invasive weeds. So, the potential applications of RNA guided gene drive is really very vast and this is being seen as a very important technology for the sustainable economy.

Disease vectors such as malarial mosquitoes might be engineered to resist pathogen acquisition or eliminated with a suppression drive. And, wild populations that serve as reservoirs for human viruses could be immunized using Cas9, RNAi machinery. Again, the reversal in immunization drives could help ensure that all transgenes are safe and controlled. Drives might quickly spread protective genes threaten or soon-to-be threatened species such as amphibians facing the expansion of say for example, chytrid fungus.

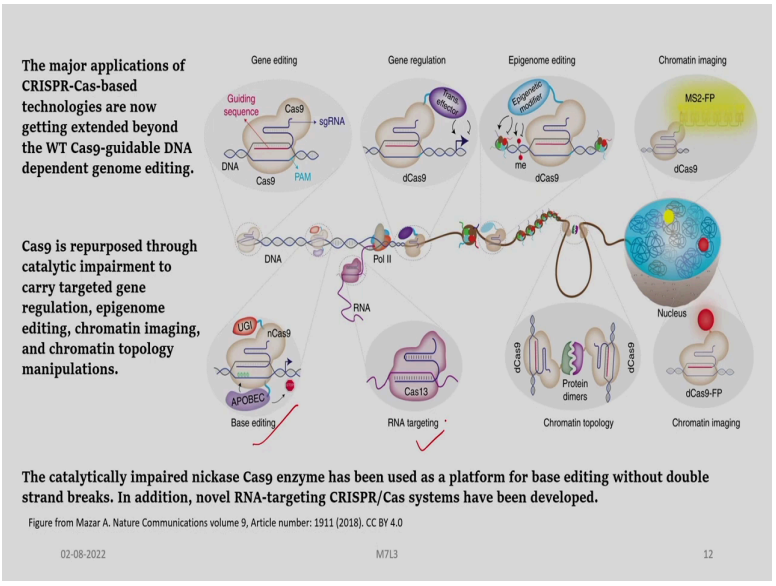
Invasive species might be locally controlled or eradicated without directly affecting others. Sensitizing drives could improve the sustainability and safety of pesticides and herbicides. And, gene drives could test ecological hypothesis concerning gene flow, sex ratios, speciation and evolution.

(Refer Slide Time: 13:02)



So, we know that CRISPR Cas9 is a genome editing workhorse with which we can do so many different kind of things. Now, the technology is emerging beyond genome editing. So, what we can do with CRISPR Cas9 or CRISPR Cas system beyond genome editing?

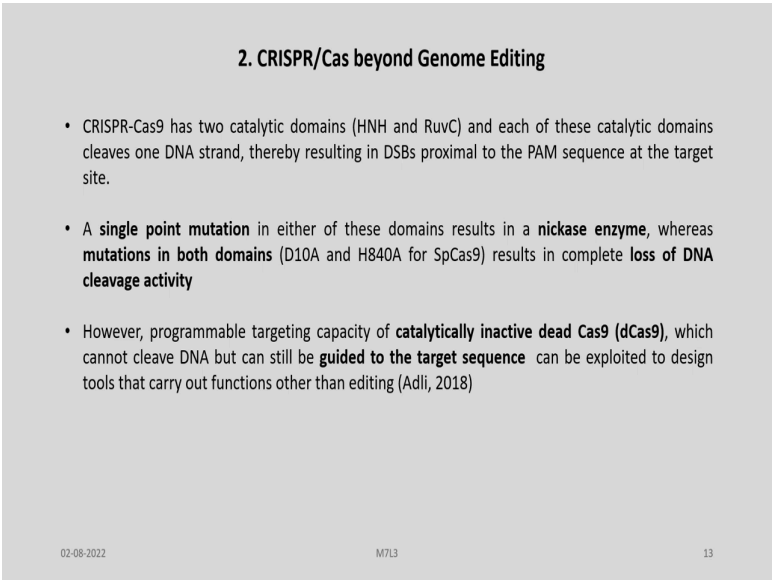
(Refer Slide Time: 13:27)



So, we know that major application of this technology is in genome editing which are now getting extended beyond these genome editing platform. Cas9 is getting repurposed through catalytic impairment to carry targeted gene regulation, epigenome editing, chromatin imaging and chromatin topology manipulations as you can see over here and also in case of base editing as well. The catalytically impaired nickase Cas9 enzyme has been used as a platform

for base editing without double-strand breaks. In addition, novel RNA-targeting CRISPR Cas9 systems has also been developed.

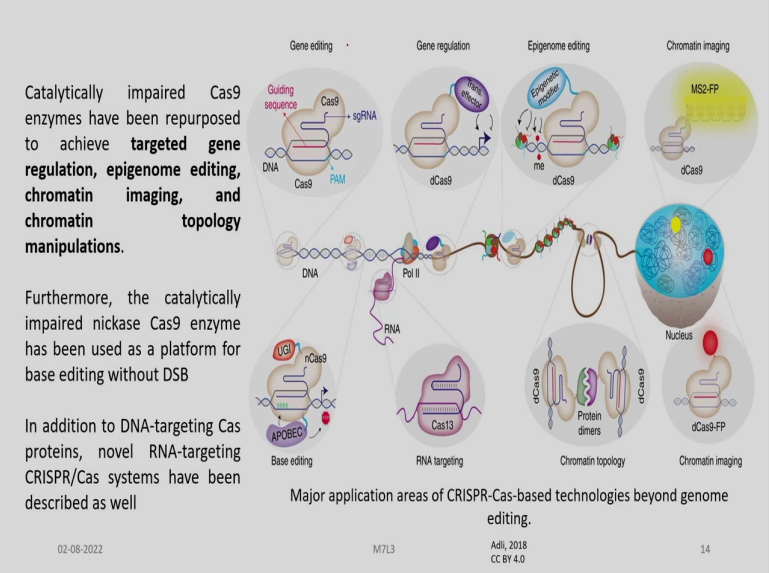
(Refer Slide Time: 14:23)



We know that CRISPR Cas9 has two catalytic domains, the HNH domain and the RuvC domain and each of these catalytic domains cleaves one DNA strand, thereby resulting in double strand breaks proximal to the PAM sequence at the target site. A single point mutation in either of these domains results in a nickase enzyme, whereas mutations in both domains D10A and H840A for SpCas9 results in complete loss of DNA cleavage activity.

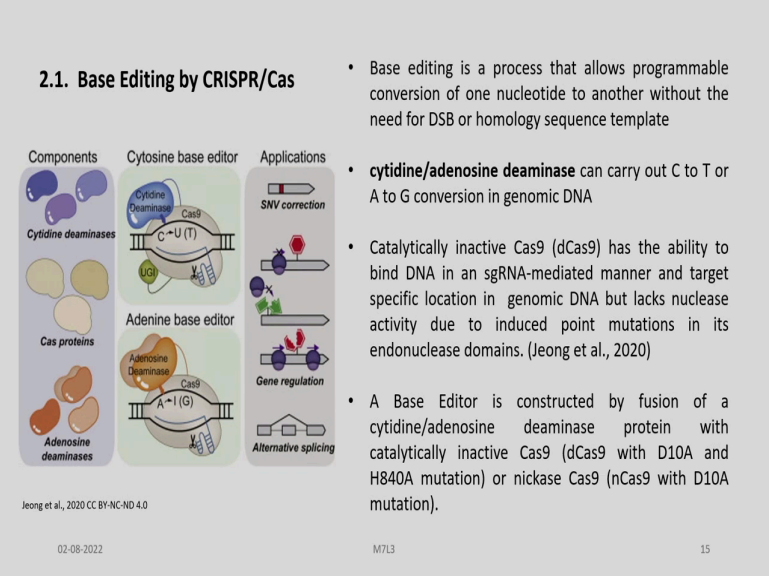
However, programmable targeting capacity of catalytically inactive dead Cas9 or which we call as dCas9, which cannot cleave DNA but can still be guided to the target sequence and can be exploited to design tools that carry out functions other than editing.

(Refer Slide Time: 15:21)



And, some of these we have discussed in the earlier slide. So, here you can see the various applications which we can do with the dCas9 enzyme.

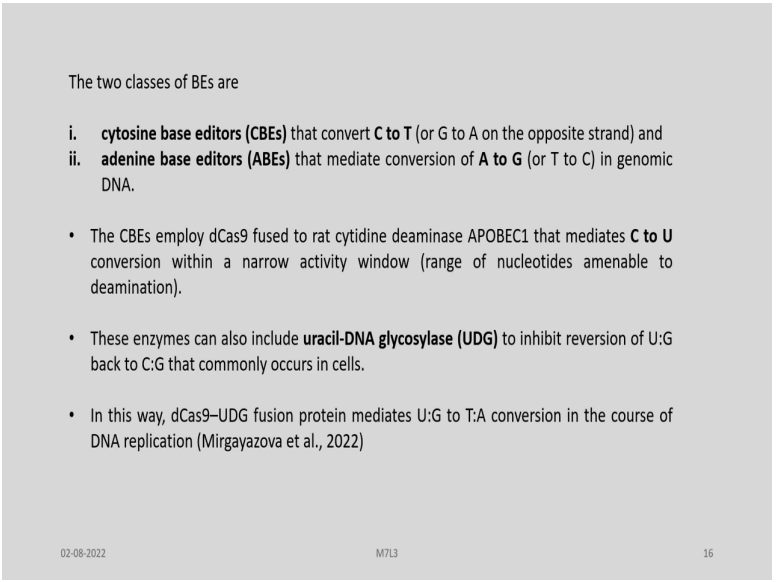
(Refer Slide Time: 15:34)



Let us start with base editing by CRISPR Cas9. So, here you have various components like cytidine deaminases, then Cas proteins and adenosine deaminases. And, then you have here the cytosine based editor and adenine based editor in the lower part and the corresponding applications like SNV correction, then gene regulation and alternative splicing. What is base editing? It is a process that allows programmable conversion of one nuclear nucleotide to another without the need for DSB or homologous sequence template.

Cytidine or adenosine deaminase can carry out C to T or A to G conversion in genomic DNA. Catalytically inactive Cas9 or dead Cas9 has the ability to bind DNA in an sgRNA-mediated reaction and target specific locations in genomic DNA but lacks nuclease activity due to induced point mutations in its endonuclease domains. A base editor is constructed by fusion of a cytidine or adenosine deaminase protein with catalytically inactive Cas9 or nickase Cas9.

(Refer Slide Time: 16:58)

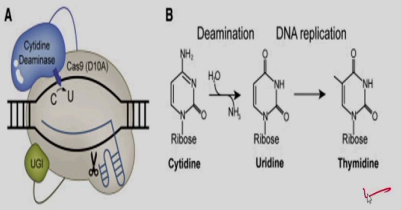


So, we have two classes of base editors, one is the cytosine base editor that converts C to T and the adenosine base editors ABEs which mediate the conversion of A to G in genomic DNA. The CBEs employ dCas9 fused to rat cytidine deaminase APOBEC1 that mediate C to U conversion within narrow activity window. These enzymes can also include uracil-DNA glycosylase (UDG) to inhibit reversion of U:G back to C:G that commonly occurs in cells. In this way, the dCas9 UDC fusion protein mediates U:G to T:A conversion in the course of DNA replication.

(Refer Slide Time: 17:49)

A. **Schematic diagram of a cytosine base editor**. A Cas9-sgRNA complex forms an R-loop at the target site in the DNA. The linked cytidine deaminase converts the exposed cytidine into a uridine. An additional linked protein, **UGI (uracil-DNA glycosylase inhibitor)**, protects uracil from uracil-DNA glycosylase (UDG).

B. After deamination, the resulting uridine is read as thymidine by DNA polymerase.



Jeong et al., 2020
CC BY-NC-ND 4.0

02-08-2022 M7L3 17

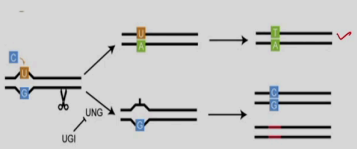
We can see here the schematic diagram of a cytosine based editor. A Cas9 sgRNA complex forms an R-loop at the target site in the DNA. The linked cytidine deaminase converts the exposed cytidine into a uridine. An additional linked protein UGI protects uracil from uracil-DNA glycosylase. After deamination, the resulting uridine is read as thymidine by the DNA polymerase.

(Refer Slide Time: 18:30)

C. The artificial uracil deoxynucleotide created by the CBE system is **repaired by two major pathways**.

1. In the upper pathway, DNA polymerase reads uracil as thymine, pairing the uracil with adenine. Ultimately, a T:A pair is made.
2. In the bottom pathway, the UDG protein removes uracil from the DNA. The resulting deoxyribonucleotide, which lacks a base, is repaired to cytosine, which is complementary to the guanine on the opposite strand.

- Alternatively, the uracil is excised, which induces the formation of insertions, deletions, or substitutions. The UGI protein inhibits the bottom pathway.



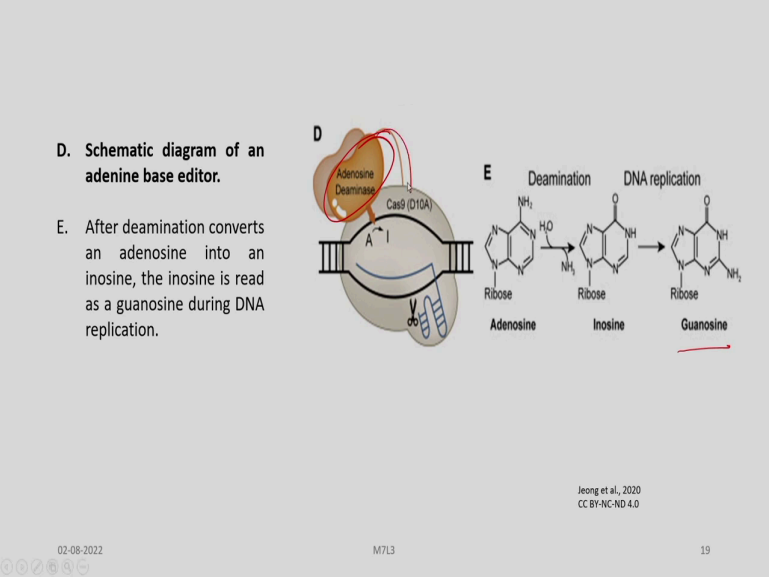
Jeong et al., 2020
CC BY-NC-ND 4.0

02-08-2022 M7L3 18

The artificial uracil deoxynucleotide created by the CBE system is repaired by two major pathways. In the upper pathway, DNA polymerase reach uracil as thymine, ultimately a T:A pair is made. In the bottom pathway, the UDG protein removes uracil from the DNA. The

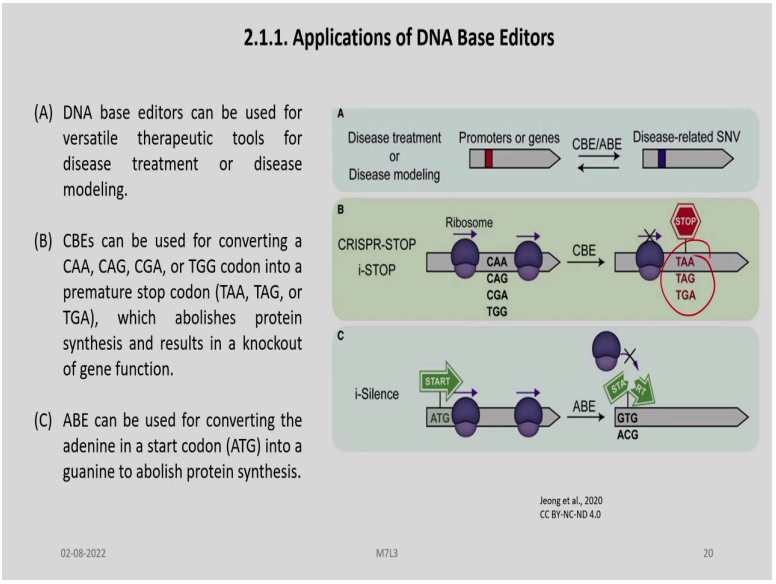
resulting deoxyribonucleotide which lacks a base, is repaired to cytosine, which is complementary to the guanine on the opposite strand. Alternatively, the uracil is excised, which induces the formation of insertions, deletions or substitutions. The UGI protein inhibits the bottom pathway.

(Refer Slide Time: 19:21)



This is a schematic diagram of an adenine based editor where there is adenosine undergoing deamination to form inosine which finally, gets conversion converted to guanosine. So, let us see that after this deamination, the inosine is read as a guanosine during this DNA replication process and you have this adenosine deaminase here connected to the Cas9 through a loop.

(Refer Slide Time: 19:57)



What are the various applications of DNA base editors? You can see here in figure (a), DNA based editors can be used for versatile therapeutic tools for disease treatment or disease modelling. So, here a disease related SNV is generated as a result of this base editing. In step B, you can see CBEs can be used for converting a CAA or CAG or CGA or TGG codon into a premature stop codon which abolishes protein synthesis and results in a knockout of gene function.

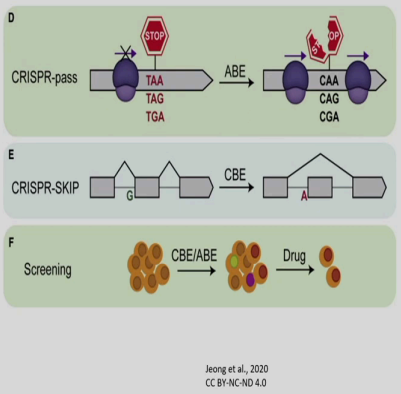
In step C, you can see that adenosine base editor can be used for converting the adenine in a start codon into a guanine to abolish protein synthesis. So, either we do here a knockout by creating a stop codon or we abolish the start codon and thereby you know abolished protein synthesis.

(Refer Slide Time: 21:31)

(D) ABE can be used for converting a premature stop codon into a codon for an amino acid. Although the resulting codon is not always the same as the unmutated original, whole protein synthesis is no longer abolished.

(E) CBEs can be used for converting Gs within splicing acceptor sites into As by editing Cs in the complementary strand of the target site.

(F) BEs can be used for generating various substitutions for screening experiments.



Jeong et al., 2020
CC BY-NC-ND 4.0

02-08-2022 M7L3 21

ABE can be used for converting a premature stop codon into a codon for an amino acid. Although the resulting codon is not always the same as the unmutated original, whole protein synthesis is no longer abolished. CBEs can be used for converting Gs within splicing acceptor sites into As by editing Cs in the complementary strand of the target. And, in figure F you can see BEs can be used for generating various substitutions for screening experiments.

(Refer Slide Time: 22:31)

2.2. Prime Editing by CRISPR/Cas

Prime editing has been developed to generate precise base-edits beyond the four transition mutations possible in Base editing

It allows precise and efficient base-to-base conversion for all 12 possible variations, including insertions, deletions, and combinations thereof without the requirement for DSB or DNA template

Prime editor consists of three main components:

I. prime editing extended guide RNA (**pegRNA**) that functions as both sgRNA and donor template for the desired alteration
II. fusion protein consisting of **Cas9 nickase** and optimized **M-MLV (Moloney murine leukemia virus) reverse transcriptase (RT)**
III. sgRNA that mediates cleavage of non-edited DNA strand by Cas9 nickase

(Kantor et al., 2020 & Mirgayazova et al., 2020)

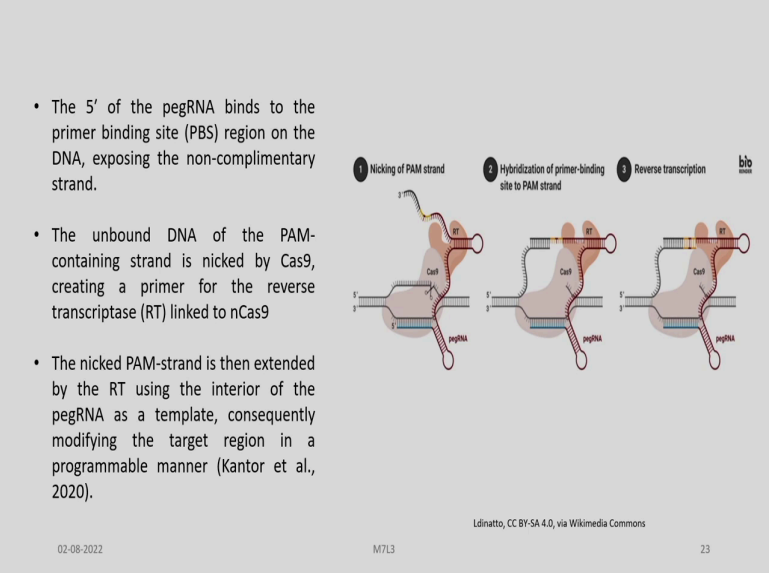
02-08-2022 M7L3 22

Let us now discuss about prime editing by CRISPR Cas9. Prime editing has been developed to generate precise base-edits beyond the four transition mutations possible in base editing. It allows precise and efficient base to base conversion for all 12 possible variations including

insertions, deletions, and combinations thereof without the requirement for DSB or DNA template.

Prime editor consists of three main components: 1 – prime editing extended guide RNA or pegRNA, that functions as both single guide RNA and donor template for the desired alterations. Then we have a fusion protein consisting of Cas9 nickase and optimized moloney murine leukemia virus for M-MLV reverse transcriptase; single guide RNA that mediates cleavage of non-edited DNA strain by Cas9 nickase.

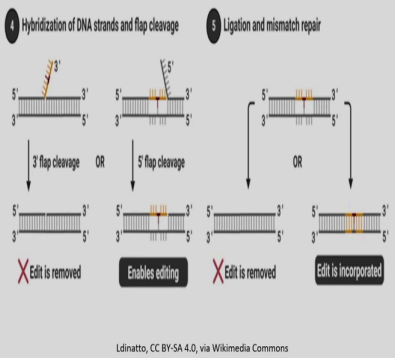
(Refer Slide Time: 23:13)



So, the 5’ of the pegRNA binds to the primer binding site region of on the DNA, exposing the non-complementary strand. The unbound DNA of the PAM-containing strand is nicked by Cas9 creating a primer for the reverse transcriptase linked to the nCas9. The nicked PAM-strand is then extended by the reverse transcriptase using the interior of the pegRNA as a template, consequently modifying the target region in a programmable manner as seen in this picture.

(Refer Slide Time: 23:52)

- The result of this step is two redundant PAM DNA flaps: **the edited 3′ flap** that was reverse transcribed from the pegRNA and **the original, unedited 5′ flap**.
- 5′ flaps are preferentially **degraded by cellular endonucleases** that are ubiquitous during lagging-strand DNA synthesis.
- The resulting heteroduplex containing the unedited strand and edited 3′ flap is resolved and stably integrated into the host genome via cellular replication and repair process (Kantor et al., 2020).



Ldinatto, CC BY-SA 4.0, via Wikimedia Commons

The result of this step is two redundant PAM DNA flaps: the edited 3 prime flap as you can see in the figure that was reverse transcribed from the pegRNA and the original, unedited 5' flap. 5' flaps are preferentially degraded by cellular endonucleases that are ubiquitous during lagging-strand DNA synthesis. The resulting heteroduplex containing the unedited strand and edited 3' flap is resolved and stabely integrated into the host genome via cellular replication and repair process.

(Refer Slide Time: 25:20)

2.3. CRISPR/Cas9 System for Gene Regulation

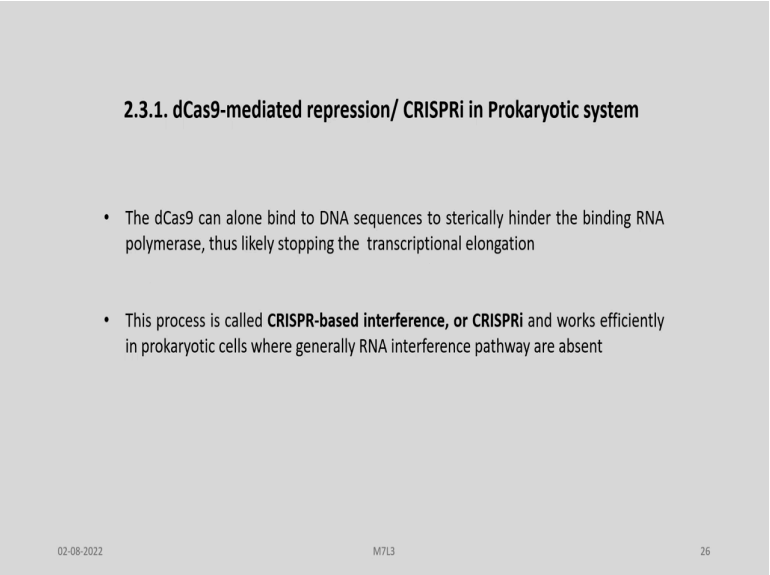
- **dCas9** could be fused with a variety of other enzymes or transcription factors to mediate site-specific regulation of gene expression
- It includes modulation of downstream gene expression by means of **transcriptional activation** (CRISPR activation, or CRISPRa), e.g., VP64, P65, Rta, or **transcriptional repression** (CRISPR interference, or CRISPRi), for example, KRAB
- The regulatory proteins can be either fused to **dCas9 or to RNA-binding proteins (RBPs)** recruited to the target site through interaction with **a hybrid RNA scaffold** coupling **sgRNA** and **RNA hairpins** (e.g., MS2, PP7) (Mirgayazova et al., 2020 & Hsu et al., 2014).

Another area in which the CRISPR Cas9 modified system is used is for gene regulation. So, dead Cas9 could be fused with a variety of enzymes or transcription factors to mediate site

specific regulation of gene expression which includes modulation of downstream gene expression by means of transcriptional activation which we call as CRISPR activation or CRISPRa. Examples are VP64, P65, Rta or transcriptional repression (CRISPR I or CRISPR interference), for example, like K R A B – KRAB.

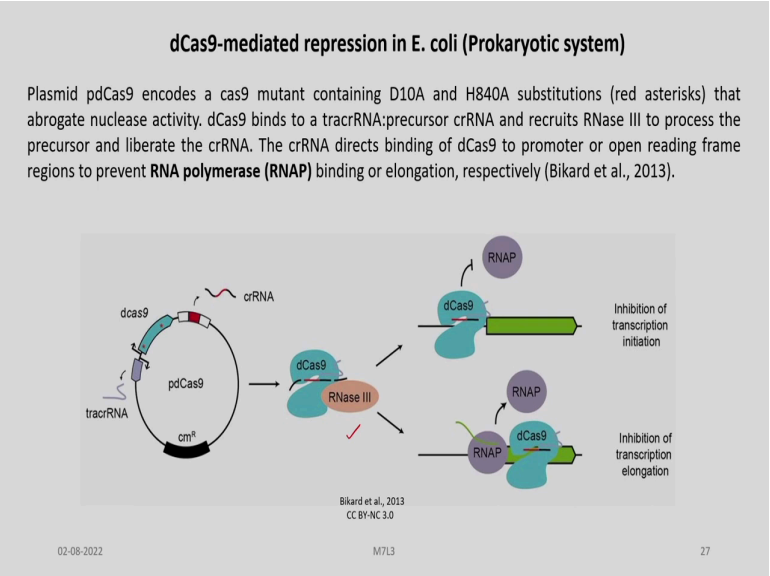
And, the regulatory proteins can be either fused to dCas9 or to a RNA-binding protein recruited to the target site through interaction with a hybrid RNA scaffold coupling the single guide RNA and the RNA hairpins.

(Refer Slide Time: 26:25)



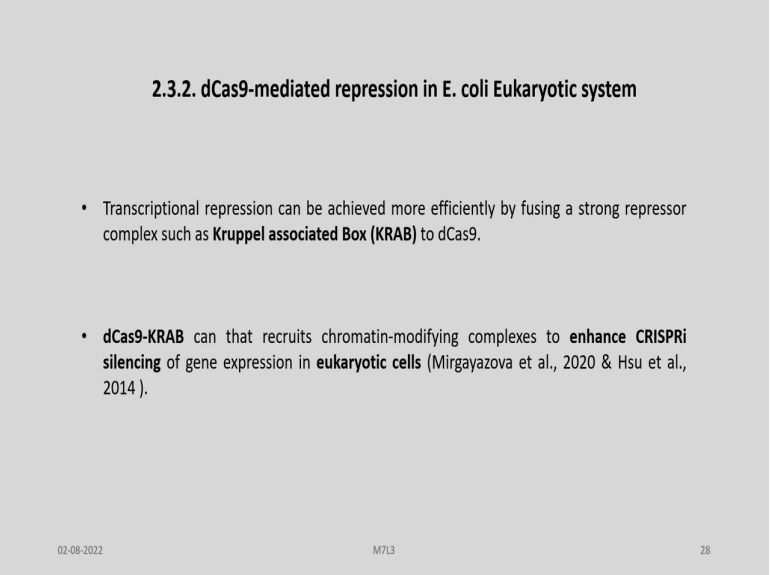
Let us discuss about the dCas9 mediated repression or CRISPRi in prokaryotic system. The dCas9 alone bind to DNA sequences to sterically hinder the binding RNA polymerase, thus likely stopping the transcriptional elongation. We call this process as CRISPR-based interference, or CRISPRi and works efficiently in prokaryotic cells where generally RNA difference pathways are absent.

(Refer Slide Time: 26:57)



So, plasmid pdCas9 encodes a Cas9 mutant containing D10A and H840A substitutions as shown with the red asterisks that abrogate nuclease activity. The dCas9 binds to a tracrRNA precursor crRNA and recruits RNase III to process the precursor and liberate the CRISPR RNA. The CRISPR RNA directs binding of dCas9 to promoter or open reading frame to prevent RNA polymerase binding or elongation respectively.

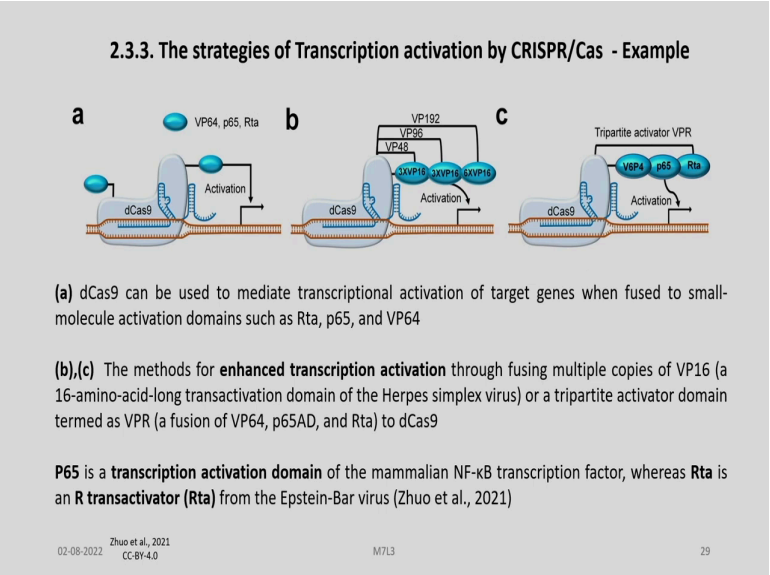
(Refer Slide Time: 27:45)



dCas9 mediated repression in *E.coli* Eukaryotic system. Here transcriptional repression can be achieved more efficiently by fusing a strong repressor complex such as KRAB or Kruppel

Associated Box to dCas9. dCas9-KRAB can recruit chromatin-modifying complexes to enhance CRISPRi silencing of gene expression in eukaryotic cells.

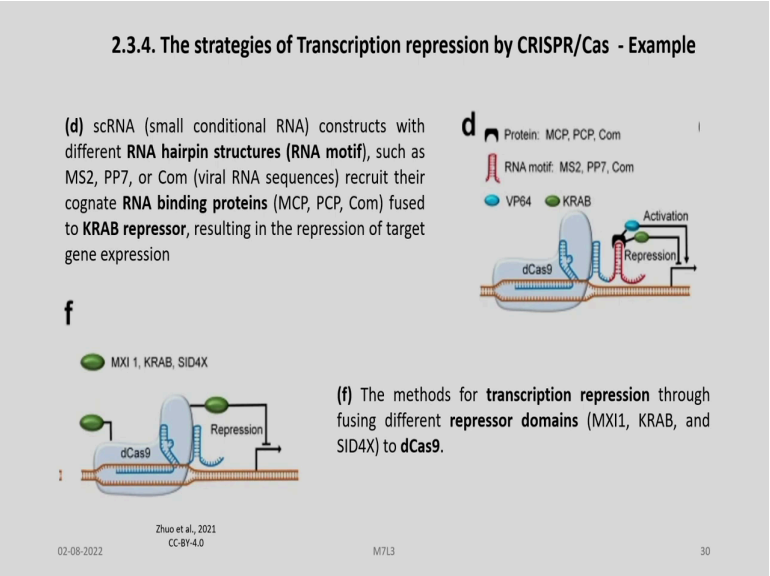
(Refer Slide Time: 28:13)



You can see in this figure the strategies of transcription activation of CRISPR Cas example. In (a), you can see d Cas9 is used to mediate the transcriptional activation of target genes when fused to small-molecule activation domains such as p65, Rta and VP64. In figure c, you can see the depiction of the method for enhanced transcription activation through fusing multiple copies of VP16 or a tripartite activated domain termed as VPR to dCas9.

P65 is a transcription activation domain of the mammalian NF kappa B transcription factor whereas, Rta is an R transactivator from the Epstein-Bar virus.

(Refer Slide Time: 29:09)



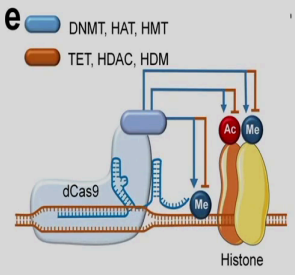
In figure d you can see small conditional RNA constructs with different RNA hairpin structures or RNA motifs such as MS2, PP7 or Com recruit their cognate RNA binding proteins fused to KRAB repressor, resulting in the repression of target gene expression. In figure (f), you can see the methods for transcription repression through fusing different repression domains like MXI1, KRAB and SID4X to dCas 9.

So, we can see that using dCas9 and fusing various transcription initiators or repression factors or any other modifying factors, we can develop so many different applications of CRISPR Cas9 beyond its simple editing applications.

(Refer Slide Time: 30:30)

2.4. Epigenetic editing by CRISPR/Cas

- Epigenetic chemical modifications on DNA or histone proteins in mammalian cells influence chromatin organization and gene expression
- **Fusion of dCas9 with epigenetic enzymatic domains** can act as an excellent tool to study the relationship between epigenetic modification and gene expression
- Epigenetic-modifying enzymes, e.g. histone demethylase (**HDM**)/methyltransferase (**HMT**), histone acetyltransferase (**HAT**)/deacetylase (**HDAC**), and DNA-demethylating enzymes (**TET**)/methyltransferase (**DNMT**), are fused with dCas9 to control diverse epigenetic states of targeted endogenous genes (Zhuo et al., 2021)



Zhuo et al., 2021
CC-BY-4.0

Thank you, for listening to this Part – A. We will be continuing our discussion on the applications of CRISPR in the second part as well.