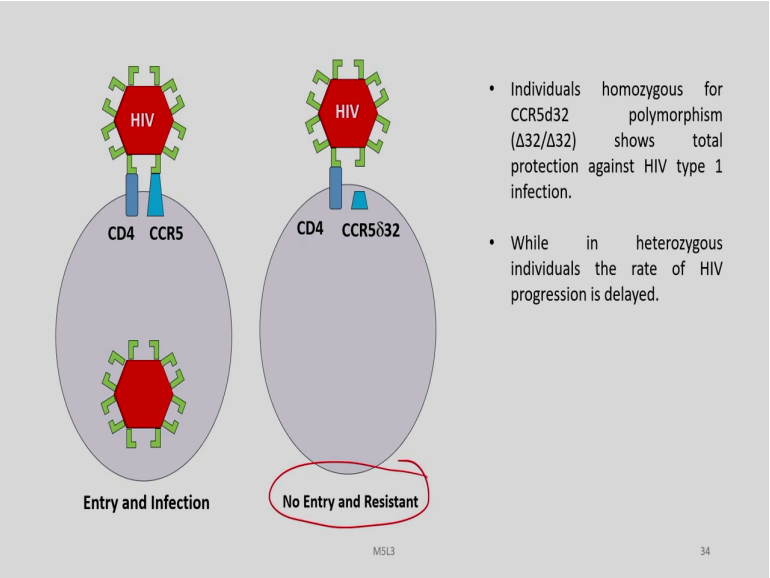
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 07
Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 technology
Lecture - 02
CRISPR/Cas9 in Genome Editing - Part B

Welcome back to my course on Genome Editing and Engineering. We were discussing about the CRISPR Cas9 system in Genome Editing in part A. In part B, we are going to discuss a little bit more of its specific applications.

(Refer Slide Time: 00:44)

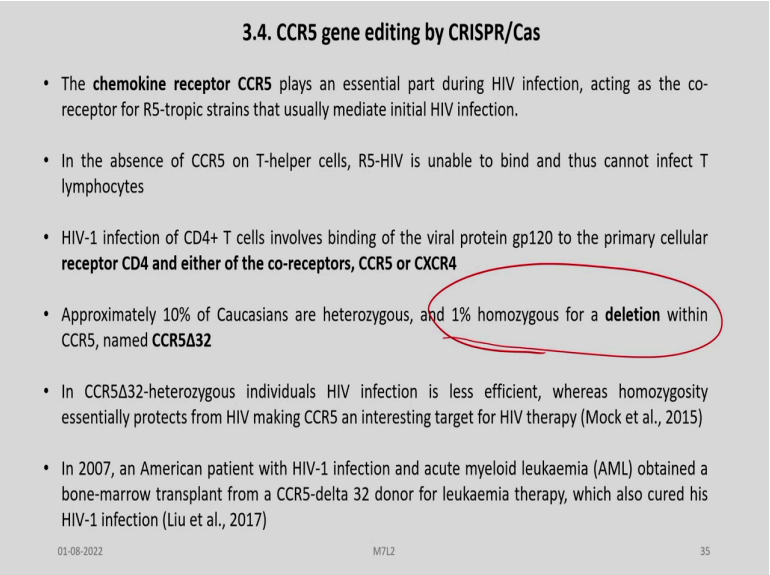


Here is a picture which may be familiar to you when we were discussing about the ZFN technology. So, this is basically the entry of HIV into a cell, which requires the presence of both CD4 and CCR5. And, you know about the CCR5 del 32 mutation, due to which there is a truncation in the protein and thereby the receptor becomes unavailable for binding of the HIV virus.

And, because of which it cannot enter the cell and such population of the cells or individuals are therefore, resistant to HIV infection. And particularly the individuals which are homozygous so, very high rate of resistance and those with heterozygous conditions can

delay the onset of HIV progression. So, this approach can be used for gene therapy of HIV infection.

(Refer Slide Time: 02:13)

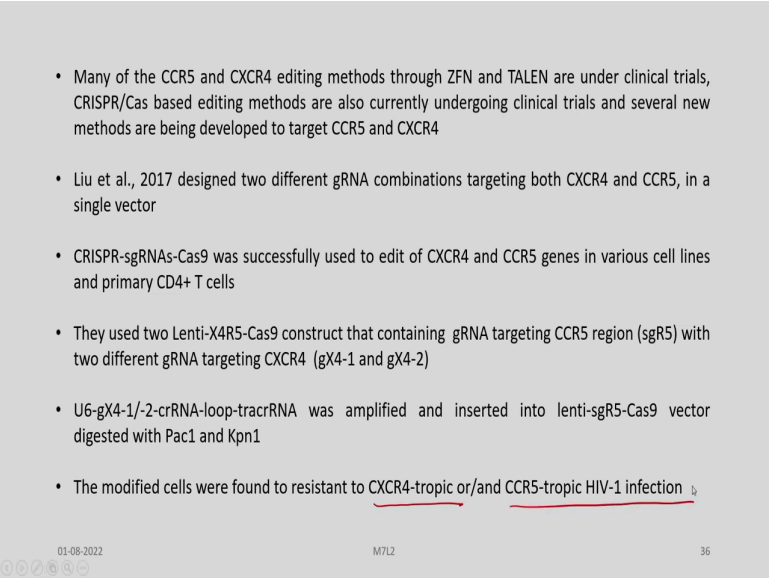


And, we have discussed this in the earlier two cases of ZFN and TALEN. We can also use CRISPR Cas9 in a similar way and all these statistics are known to you about that 1 % population with a homozygous condition.

And, then you also know about the American patient with HIV infection who was having acute myeloid leukemia, and he was given bone marrow transplant from a CCR5 delta 32 donor for this therapy and this also cured him of the HIV – 1 infection.

And these were some of the accidental things which help in the development of a gene therapy approach for HIV.

(Refer Slide Time: 02:56)

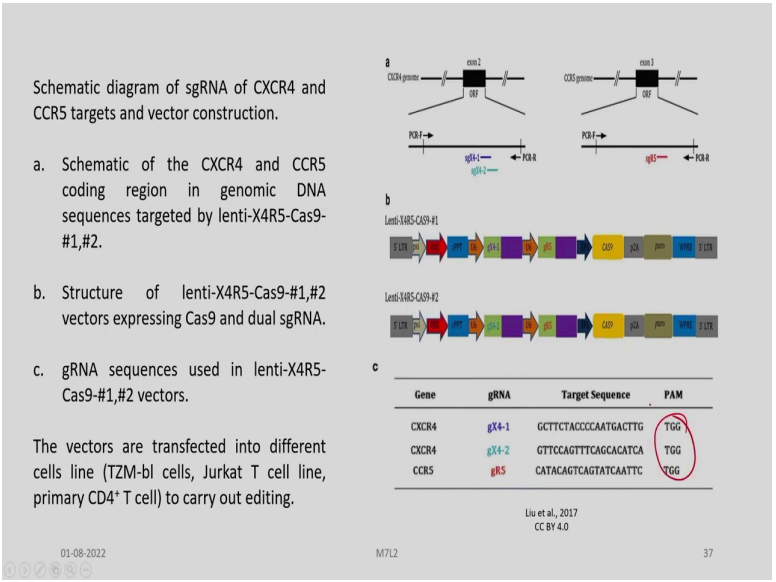


So, we now know that many of the CCR5 and CXCR4 editing methods through ZFN and TALEN are under clinical trials. In the case of CRISPR Cas9 Liu et al., in 2017 designed two different guide RNA combinations targeting both CXCR 4 and CCR5, in a single vector.

The CRISPR-sgRNA Cas9 was successfully used to edit CXCR4 and CCR5 genes in various cell lines and primary CD4 plus T cells and they used two Lenti-X4R5-Cas9 constructs that contain guide RNA targeting the CCR5 region with two different guide RNA targeting the CXCR4 gx41 and gx4-2.

U6-gX4-1/2-crRNA-loop-tracrRNA was amplified and inserted into the Lenti-sgR5Cas9 vector digested with Pac1 and Kpn1. The modified cells were found to be resistant to CXCR4-tropic or CCR5-tropic HIV – 1 infection.

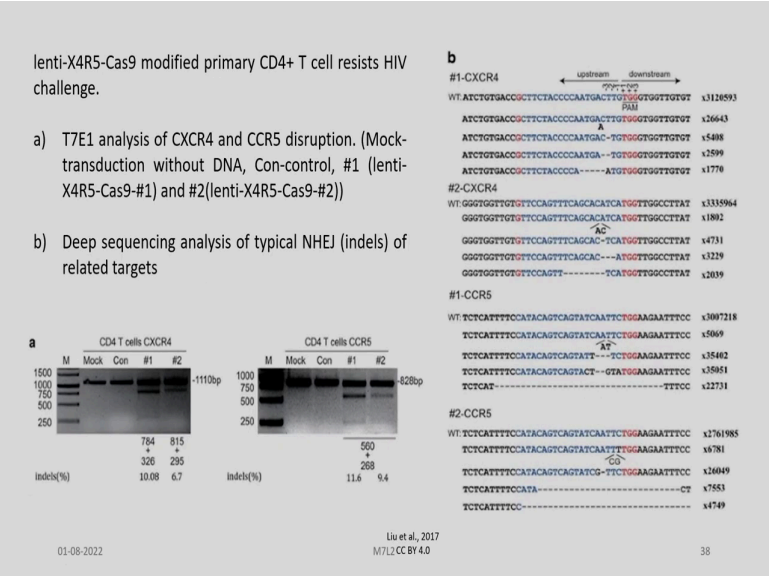
(Refer Slide Time: 04:30)



You can see here the schematic diagram of single guide RNA of CXCR4 and CCR5 targets and the vector construction. So, the figure (a) show the schematic of CXCR 4 and CCR5 coding region in the genomic DNA sequences which are targeted by lenti-X4R5-Cas9.

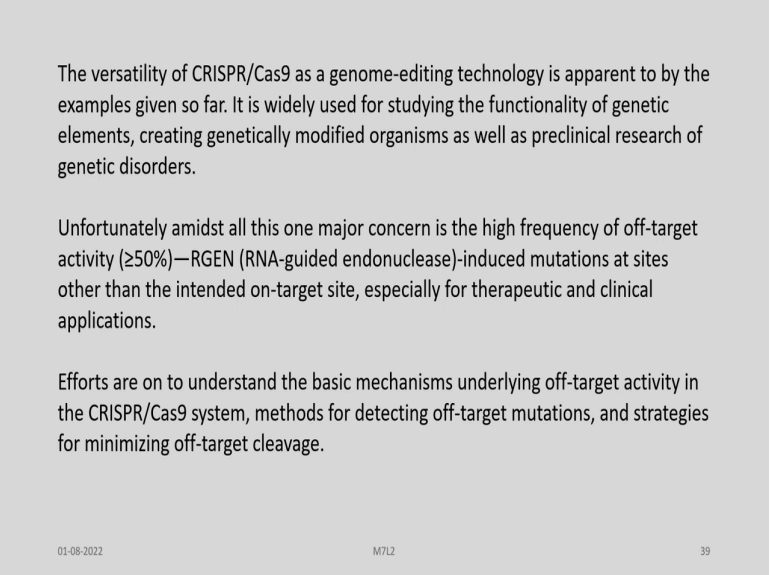
And, in (b), you can see the structure of the lenti-X4R5-Cas9 vectors expressing the Cas9 and the dual single guide RNA. In figure (c), we can see the guide RNA sequences which are used in the lenti-X4R5-Cas9 vectors. So, you can see here the CXCR4 and CCR5 genes and the corresponding guide RNAs which are being designed with the target sequences and the pump sequences here. The vectors are transfected into different cell lines to carry out the editing.

(Refer Slide Time: 05:50)



Lenti-X4R5-Cas9 modified primary CD4 plus T cells resists the HIV challenge or infection. In figure you can see the gel analysis of the various experiments. The T7E1 analysis of CXCR4 and CCR5 disruption is presented and in (b) the sequencing analysis of the typical non-homologous end joining indels of the related targets are presented.

(Refer Slide Time: 06:32)



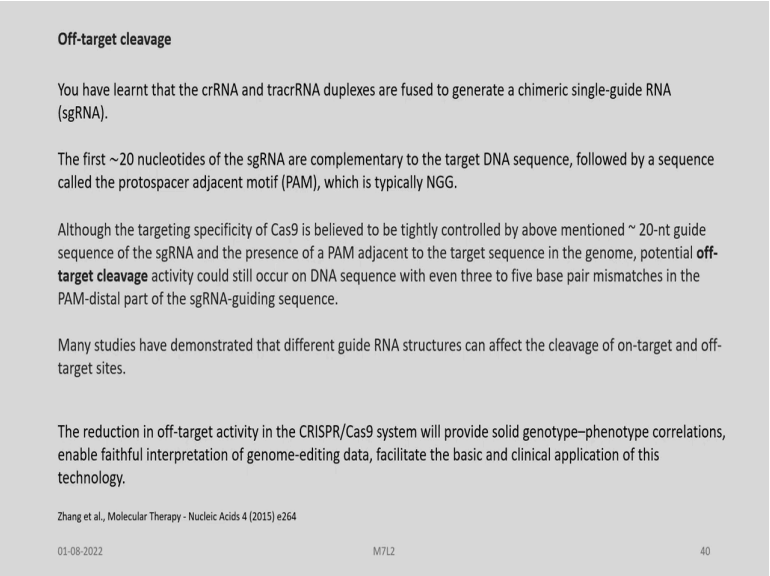
The versatility of the CRISPR Cas9 as a genome editing technology is apparent to you by the various examples given so far. It is widely used for studying the functionality of genetic

elements, creating genetically modified organisms as well as preclinical research of genetic disorders.

Unfortunately, amidst all these potential applications one major concern is the high frequency of off-target activity or RGEN – (RNA-guided endonuclease)-induced mutations, at sites other than the intended on-target site, especially in the case of therapeutic or clinical applications.

Efforts are on to understand the basic mechanisms underlying off-target activity in the CRISPR Cas9 system. Methods for detecting off-target mutations and strategies for minimizing off-target cleavage are being actively pursued or explored.

(Refer Slide Time: 07:38)

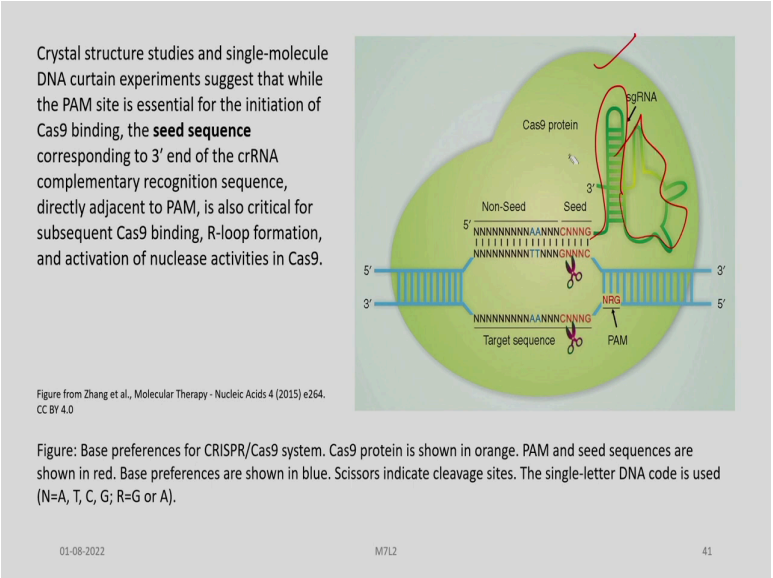


We now know that CRISPR RNA and tracrRNA duplex are fused to generate a chimeric single-guide RNA. And, the first 20 nucleotides of these single guide RNA are complementary to the target DNA sequence, followed by a sequence called the PAM motif which is typically NGG.

Although the targeting specificity of Cas9 is believed to be tightly controlled by the above mentioned approximately 20 nucleotide guide sequence of the single guide RNA and the presence of the PAM motif adjacent to the target sequence in the genome, potential off-target cleavage activity still occurs on DNA sequences with even 3 to 5 base pair mismatches in the PAM distal part of the sgRNA-guiding sequence.

Many studies have demonstrated that different guide RNA structures can affect the cleavage of on-target and off-target sites. The reduction of off-target activity in the CRISPR Cas9 system will provide solid genotype-phenotype correlations, and enable faithful interpretation of genome-editing data, facilitate the basic and clinical applications of CRISPR Cas9 technology.

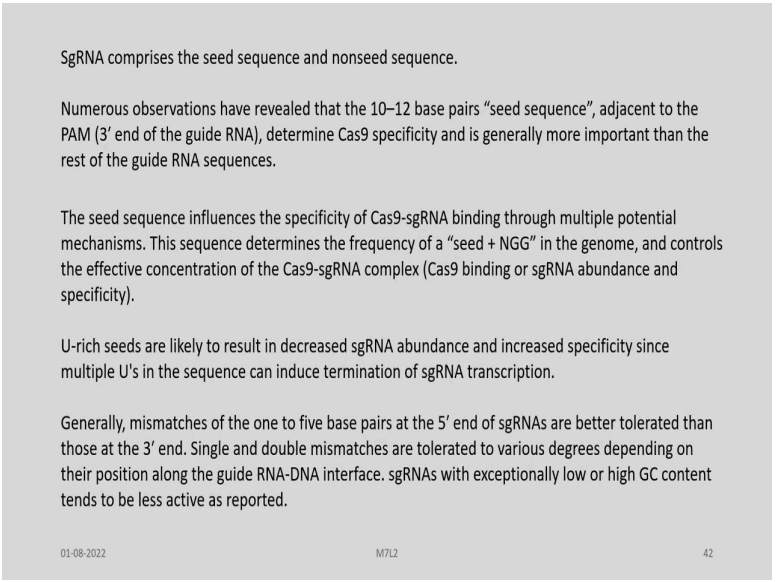
(Refer Slide Time: 09:03)



Let us have a look into the crystal structure of a single molecule DNA experiment and here is this CRISPR Cas9, you can see. And here is the single guide RNA which is binding to the target sequence and here we have the PAM sequence.

So, in these a crystal structure which was elucidated, it would suggest that while the PAM site is essential for the initiation of Cas9 binding, the seed sequence corresponding to 3' end of the CRISPR RNA complementary recognition sequence, directly adjacent to PAM, is also critical for subsequent Cas9 binding, R-loop formation, and activation of nucleus activities in Cas9.

(Refer Slide Time: 10:08)



What is this seed sequence? The single guide RNA comprises the seed sequence and a non seed sequence. Numerous observations have revealed that the 10 to 12 base pairs seed sequence adjacent to the PAM determine the Cas specificity and is generally more important than the rest of the guide RNA sequence.

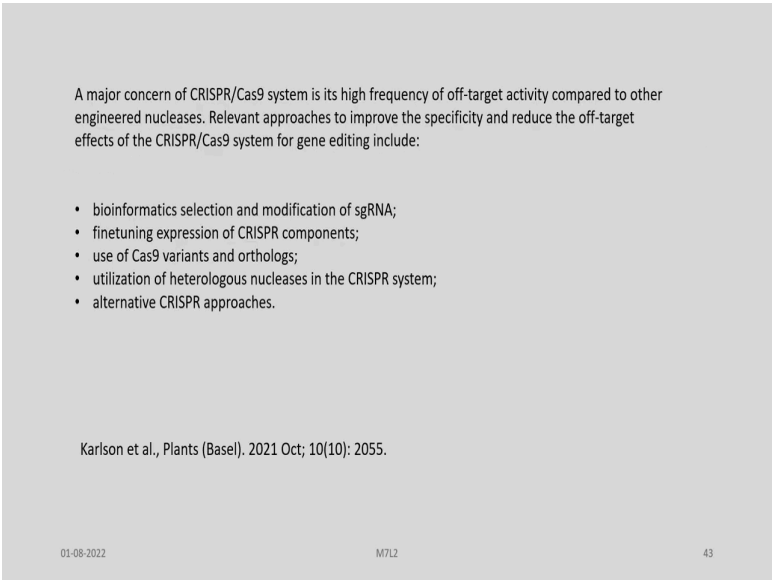
The seed sequence influences the specificity of Cas9 single guide RNA binding through multiple potential mechanisms. This sequence determines the frequency of a seed plus NGG in the genome and controls the effective concentration of the Cas9-sgRNA complex binding or sgRNA abundance and specificity.

What are these seed sequences and what are their importance? Single guide RNA comprises the seed sequence and nonseed sequence. The 10 to 12 base pair seed sequence adjacent to the PAM determine the Cas9 specificity and is generally more important than the rest of the guide RNA sequence.

The seed sequence influences the specificity of Cas9 single guide RNA binding through multiple potential mechanisms. It determines the frequency of a seed plus NGG in the genome and controls the effective concentration of the Cas9 single guide RNA complex. U-rich seeds are likely to result in decreased single guide RNA abundance and increased specificity since multiple U’s in the sequence can induce termination of sgRNA transcription.

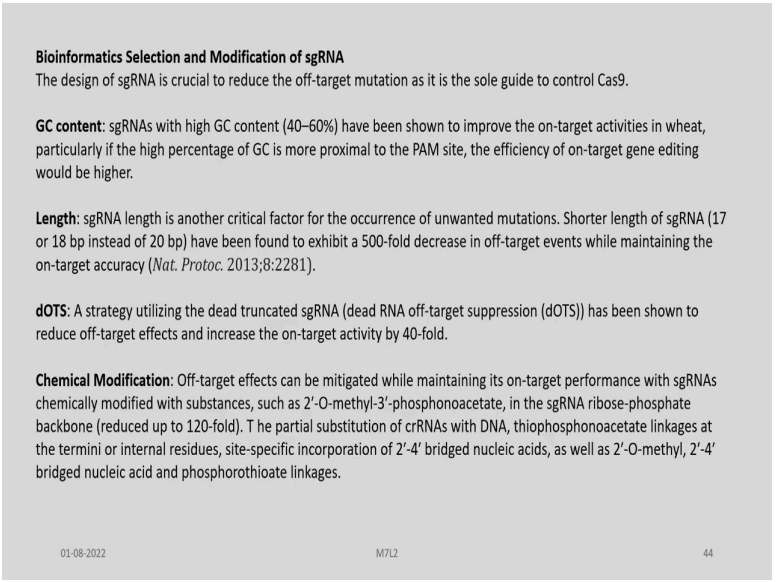
Generally, mismatches of one to five base pairs at the 5’ end of the single guide RNAs are better tolerated than those at the 3’ end. Single and double mismatches are tolerated to various degrees depending on their position along the guide RNA-DNA interface. sgRNAs with exceptionally low or high GC content tends to be less active as reported by many researchers.

(Refer Slide Time: 12:25)



A major concern of CRISPR Cas9 system is its high frequency of off-target activity compared to other engineered nucleases. Many relevant approaches to improve the specificity and reduce the off-target effects of the CRISPR Cas9 system for gene editing have been developed such as, bioinformatics selection and modification of single guide RNA; fine tuning expression of CRISPR components; use of Cas9 variants and othologs; utilization of heterologous nucleases in the CRISPR system and alternative CRISPR approaches.

(Refer Slide Time: 13:12)



We will discuss about these in brief one by one.

The first is the use of bioinformatics for selection and modification of single guide RNA. The design of single guide RNA is crucial to reduce the off-target mutation as it is the sole guide to control Cas9. Some of the important factors are GC content, length, dOTS and chemical modification.

sgRNA is with high GC content, 40 to 60, percent have been shown to improve the off on-target activities in wheat, particularly in the high percentage of GC is more proximal to the PAM site, the efficiency of on-target gene editing would be higher.

The length of single guide RNA is another critical factor for the occurrence of unwanted mutations. Shorter length of single guide RNA 17 to 18 base pair instead of 20 have been found to exhibit a 500-fold decrease in off-target events while maintaining the on-target accuracy.

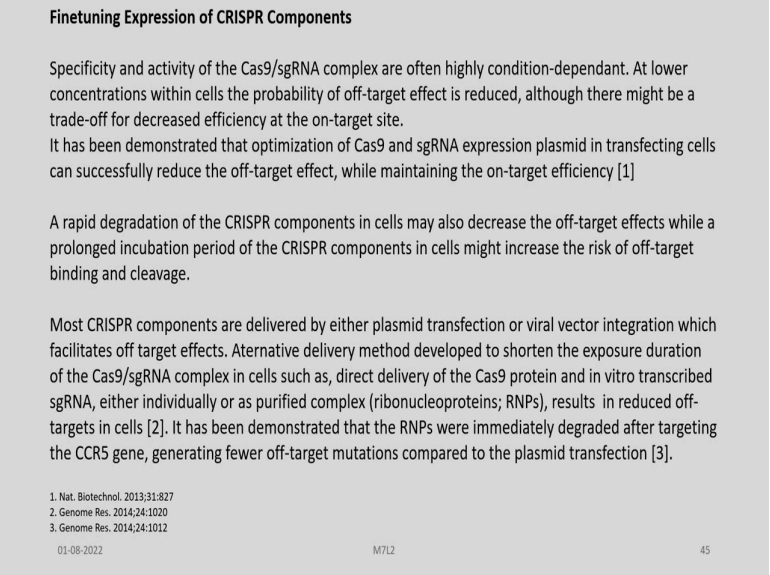
A strategy utilizing the dead truncated single guide RNA dOTS (dead RNA off-target suppression), has been shown to reduce off-target effects and increase the on-target activity by 40-fold.

Chemical modification is another strategy. Off-target effects can be mitigated while maintaining it is on-target performance with single guide RNAs chemically modified with

substances, such as 2-O-methyl-3-phosphonoacetate, in the single guide ribose-phosphate backbone and this reduces the off-targeting up to 120-fold.

The partial substitution of CRISPR RNA with DNA, thiophosphonoacetate linkages at the termini or internal residues, site-specific incorporation of 2' – 4' bridged nucleic acids, as well as 2' – O–methyl, 2' – 4' bridged nucleic acid, and phosphorothioate linkages are also suggested.

(Refer Slide Time: 15:24)



Another approach is the fine tuning of expression of CRISPR components.

Specificity and activity of the Cas9 single guide RNA complex are often highly condition dependent. At the lower concentrations within cells, the probability of off-target effect is reduced, although there might be a trade-off for decreased efficiency at the on-target site. It has been demonstrated that optimization of Cas9 and single guide RNA expression plasmid in transfecting cells can successfully reduce the off-target effect, while maintaining the on-target efficiency.

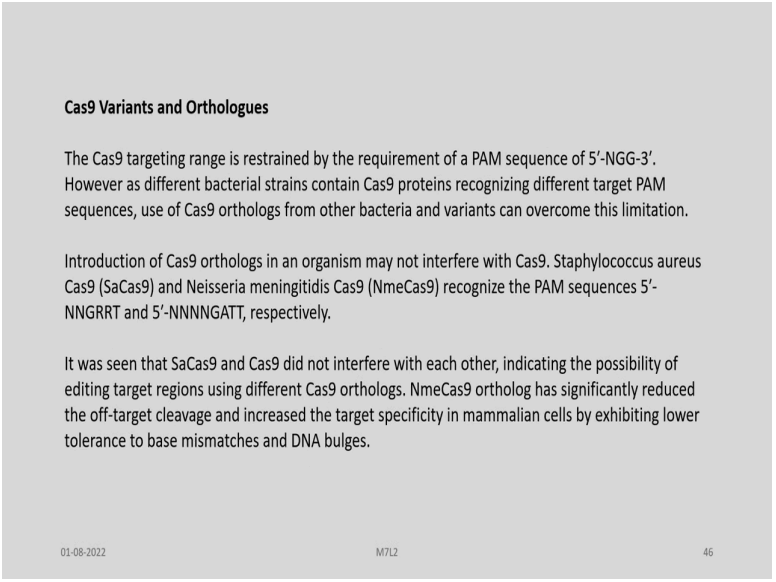
A rapid degradation of the CRISPR component in cells may also decrease the off-target effects while a prolonged incubation period of the CRISPR components in the cells might increase the risk of off-target binding and cleavage.

Most CRISPR components are delivered by either plasmid transection or viral vector integration which facilitates off-target effects. Alternative delivery methods developed to

shorten the exposure duration of the CRISPR Cas9 sgRNA complex in cells such as, direct delivery of the Cas9 protein and in vitro transcribed as the RNA, either individually or as purified complexes, results in reduced off-target in cells.

It has been demonstrated that the ribonucleoproteins were immediately degraded after targeting the CCR5 gene, generating fewer off-target mutations compared to the plasmid transfection.

(Refer Slide Time: 17:10)

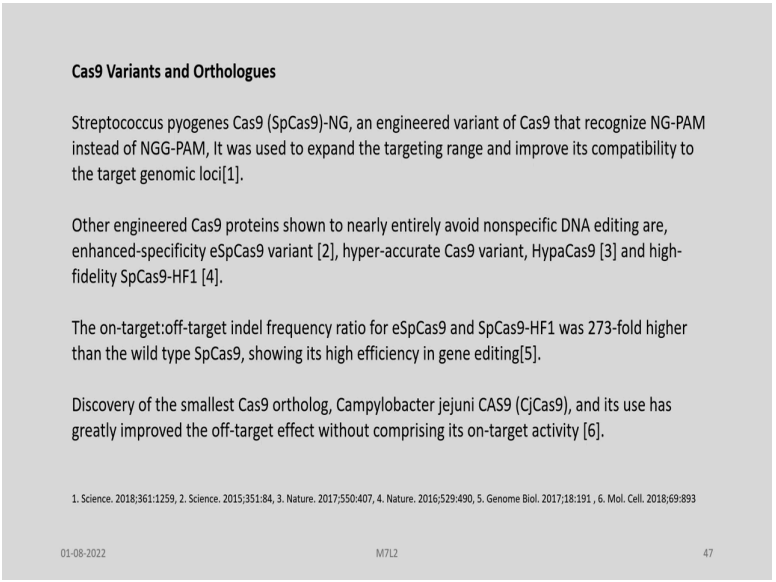


Cas9 variants and orthologues. The Cas9 targeting range is restrained by the requirement of a PAM sequence of 5' – NGG – 3' signature. However as different bacterial strains contain Cas9 proteins recognizing different target PAM sequences, use of Cas9 orthologs from other bacteria and variance can overcome this limitation.

Introduction of Cas9 orthologs in an organism may not interfere with Cas9. *Staphylococcus aureus* Cas9 and *Neisseria meningitidis* Cas9 recognize the PAM sequence 5'– NNGRRT and 5' – NNNNGATT, respectively.

It was seen that SaCas9 and Cas9 did not interfere with each other, indicating the possibility of editing target regions using different Cas9 orthologs. NmeCas9 ortholog has significantly reduced the off-target cleavage and increased the target specificity in mammalian cells by exhibiting lower tolerance to base mismatches and DNA bulges.

(Refer Slide Time: 18:27)

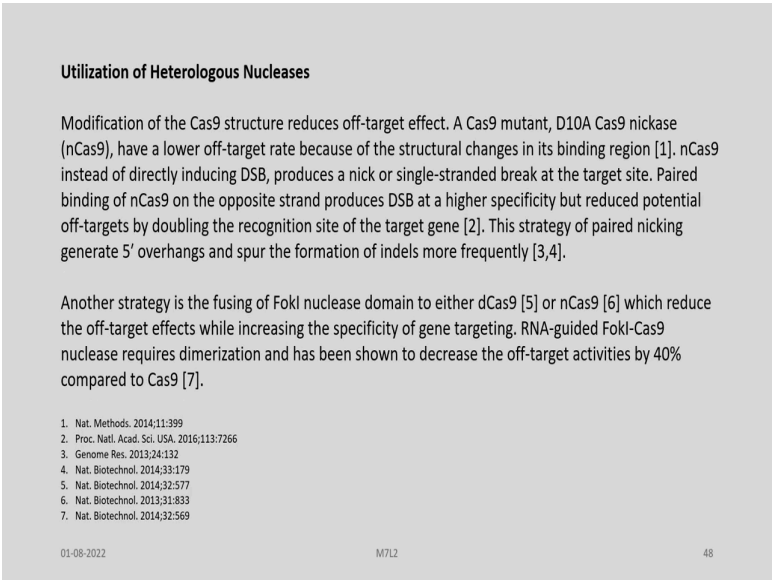


SpCas9-NG are engineered variant of Cas9 that recognize NG-PAM instead of NGG-PAM. It was used to expand the targeting range and improve its compatibility to the target genomic loci.

Other engineered Cas9 proteins shown to nearly entirely avoid nonspecific DNA editing are enhanced-specificity eSpCas9 variant, hyper-acute Cas9 variant or HypaCas9 and high-fidelity SpCas9 or HF1.

The on-target; off-target indel frequency ratio of eSpCas9 and SpCas9-HF1 was found to be 273-fold higher than the wild type SpCas9, showing its high efficiency in gene editing. The discovery of the smallest Cas9 ortholog, *Campylobacter jejuni* CAS9 CjCas9 and its subsequent use has greatly improved the off-target effect without compromising on its on-target activity.

(Refer Slide Time: 19:40)

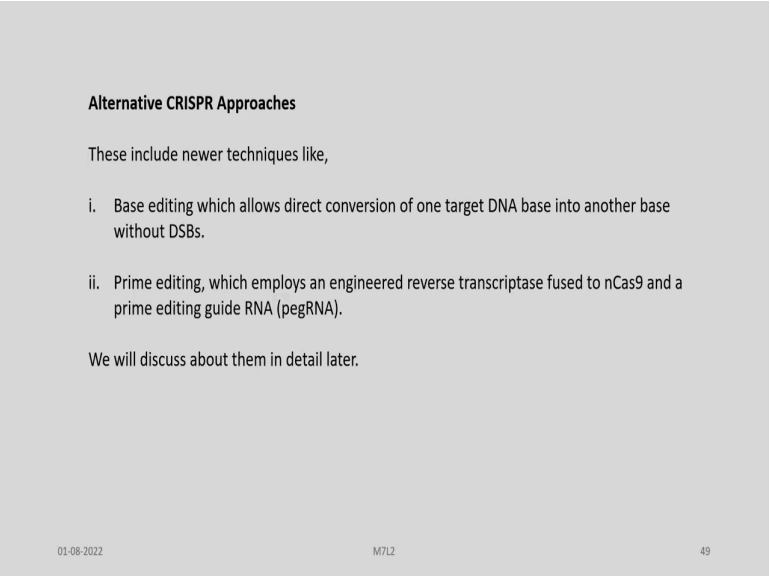


Another strategy is the utilization of heterologous nucleases. Modification of the Cas9 structure reduces off-target effect. A Cas9 mutant, D10A Cas9 nickase or nk nCas9, have a lower off-target rate because of the structural change in its binding region. nCas9 instead of directly inducing double strand breaks produces a nick or single-stranded break at the target site.

Paired binding of nCas9 on the opposite strand produces DSB at a higher specificity but reduced potential off-targets by doubling the recognition side of the target gene. The strategy of paired nicking generates 5’ overhangs and spur the formation of indels more frequently.

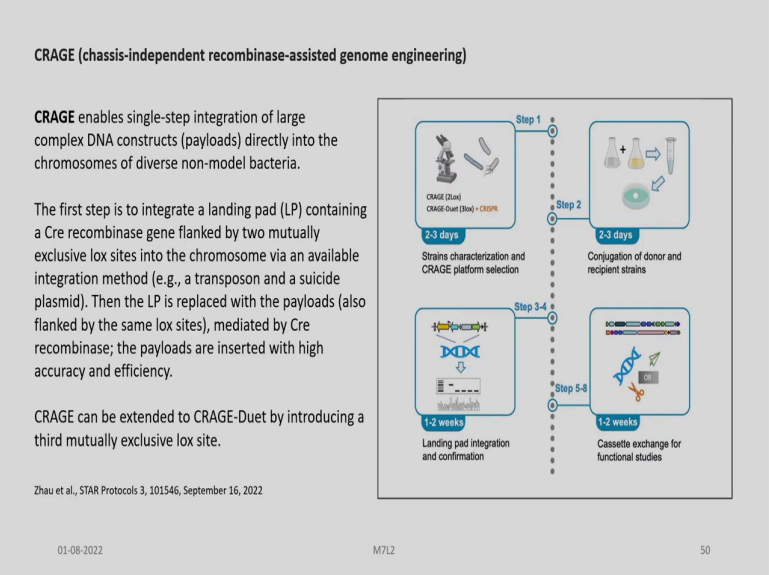
Another strategy is the fusing of FokI nuclease domain to either dCas9 or nCas9 which reduce the off-target effects while increasing the specificity of gene targeting. RNA-guided FokI Cas9 nuclease requires dimerization and has been shown to decrease the off-target activities by around 40 percent compared to Cas9 alone.

(Refer Slide Time: 21:00)



Many other strategies like alternative CRISPR approaches have also been suggested which include several techniques like base editing which allows direct conversion of one target DNA base into another base without the double strand breaks. Or prime editing which employs an engineered reverse transcriptase fused to nCas9 and a prime editing guide RNA or pegRNA. We will discuss about them later.

(Refer Slide Time: 21:30)



Let us discuss about another strategy called CRAGE, which is chassis-independent recombinase-assisted genome engineering. CRAGE enables single-step integration of large

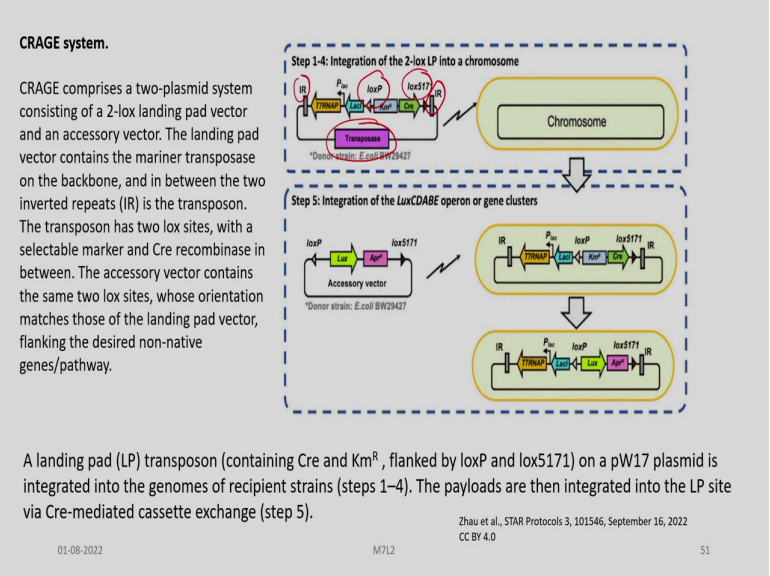
complex DNA constructs or payloads directly into the chromosomes of diverse non-model bacteria.

The first step is to integrate a landing pad containing a Cre recombinase gene flanked by two mutually exclusive lox sites into the chromosome via an available integration method. Then the LP (landing pad) is replaced with the payloads also flanked by the same lox sites mediated by Cre recombinase; the payloads are inserted with high accuracy and efficiency.

CRAGE can be extended to CRAGE-duet by introducing a third mutually exclusive lox site and there are various steps in this process and you can see the time required for carrying out step 1 requires around 2 to 3 days and here is the method of CRAGE duet plus CRISPR.

Step 2 requires around 2 to 3 days which is the conjugation of donor and recipient strains. Step 3 to 4 requires 1 to 2 weeks here the landing pad integration and confirmation is carried out and step 5 to 8 requires another 1 to 2 weeks where there is cassette exchange for functional studies.

(Refer Slide Time: 23:30)

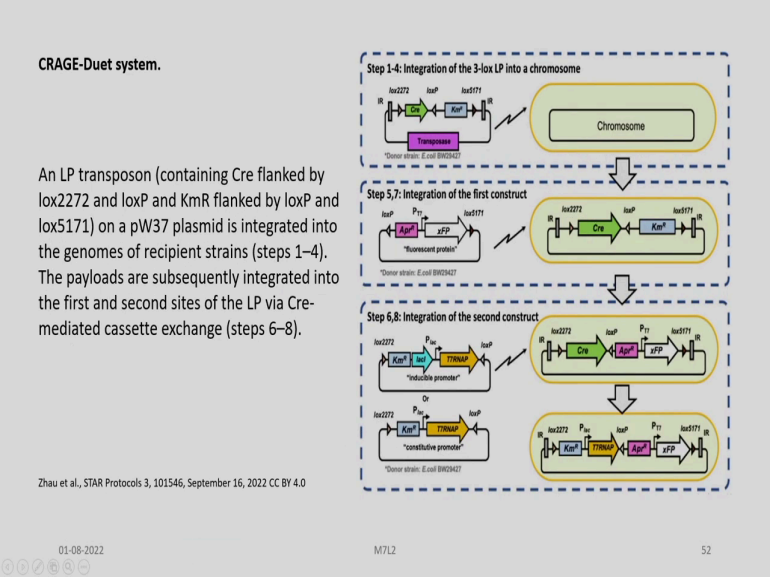


So, study the various steps of this process a little bit in detail. In step 4, you can see the integration of the 2-lox landing pad into a chromosome which is kind of a preparative stage. And, in the step 5 the integration of an operand or the gene cluster is taking place. These CRAGE system comprises a two-plasmid system consisting of a 2-lox landing pad vector and an accessory vector.

The landing pad vector contains the mariner transposase on the backbone, and in between the two inverted repeats is the transposon. The transposon has two lox sites, as you can see here with the selectable marker and Cre recombinase in between them. So, this is a marker. The accessory vector contains the same two lox sites, whose orientation matches those of the landing pad flanking the desired non-native genes and pathways.

Here you can see the landing pad transposon containing Cre recombinase and the selectable marker flanked by the loxP and lox5171 on a pW17 plasmid and this is integrated into the genomes of the recipient strains and the payloads are then integrated into the LP site via Cre-mediated cassette exchange.

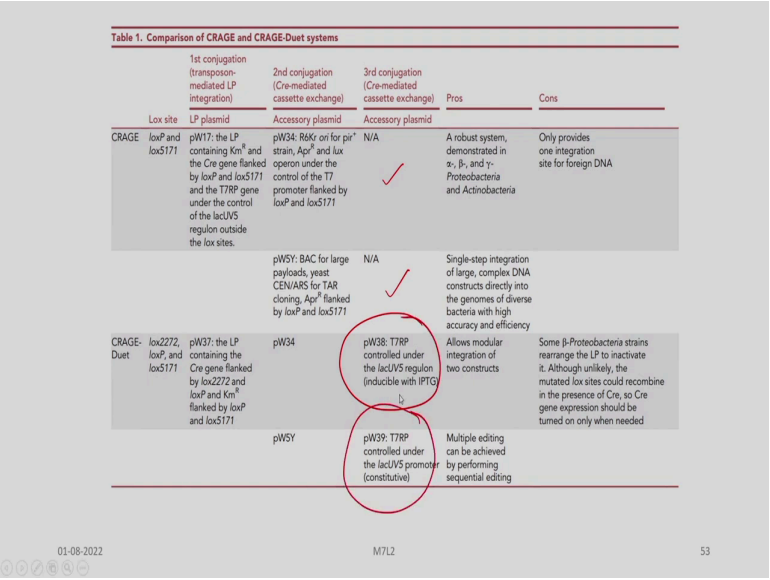
(Refer Slide Time: 25:22)



In CRAGE-duet system. An LP transposon containing Cre flanked by the lox sites and the selectable marker flanked by lox sites on a pW37 plasmid is integrated into the genomes of recipient strains in steps 1 to 4. The payloads are subsequently integrated into first and second sites of the LP via Cre recombination mediate recombinase mediated cassette exchange in step 6 to 8.

So, step 1 to 4 comprises of integration of the 3-lox LP into a chromosome. Step 5 to 7 contain integration of the first construct and step 6 to 8 involves the integration of the second construct.

(Refer Slide Time: 26:15)



This is a table who shows the comparison of CRAGE and CRAGE-duet system. So, you have the loxP sites here. As told earlier in CRAGE duet there is an additional third loxP site. And, regarding the 1st conjugation of the LP plasmid you have here the pW17: the LP containing Km^R and the Cre gene flanked by the lox sites and the T7RP gene under the control of the lacUV5 regulon outside the lox sites. While here you have pW37: the LP containing Cre gene flanked by lox 2272 and lox P and Km^R flanked by loxP and lox5171.

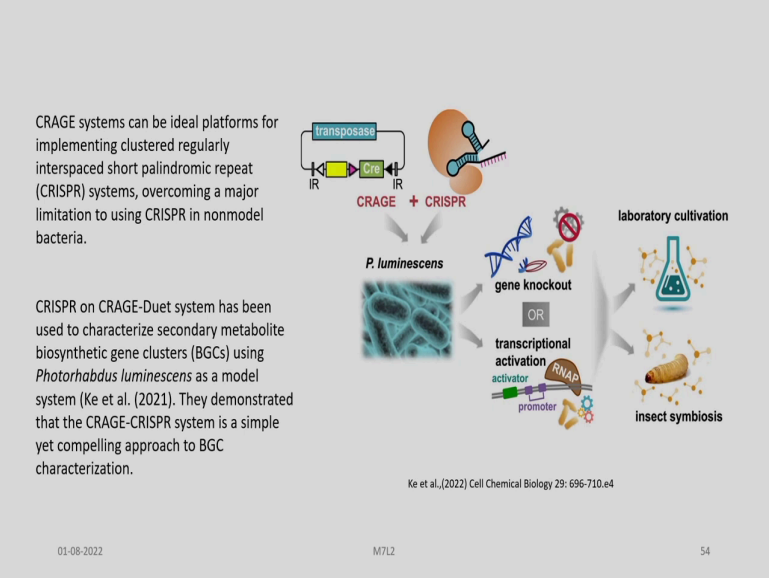
And the 2nd conjugation where the Cre-mediated cassette exchanges happen. The accessory plasmid here in this case of CRAGE is pW34 and pW5Y and here we can have pW34 and pW5Y as well. So, third conjugation it is not applicable in the case of CRAGE, and you can see here in CRAGE-duet you have a third conjugation as well.

And the advantages and disadvantages of these two systems: CRAGE is a robust system, demonstrated in alpha, beta, gamma, proteobacteria and actinobacteria, but the disadvantage is that it only provides one integration site for the foreign DNA and when we are using pW5Y BAC for larger payloads, yeast centromere ARS etcetera.

The single-step integration of large, complex DNA constructs directly into the genomes of diverse bacteria with high efficiency and accuracy is possible. In CRAGE-duet it allows modular integration of the two constructs, but some beta-proteobacteria strains rearrange the LP to inactivate it which is a disadvantage.

In the case of these pW5Y accessory plasmid multiple editing can be achieved by performing sequential editing. So, this in brief sums up the various features of CRAGE and CRAGE-duet system and their advantages and the disadvantages associated is in these two systems.

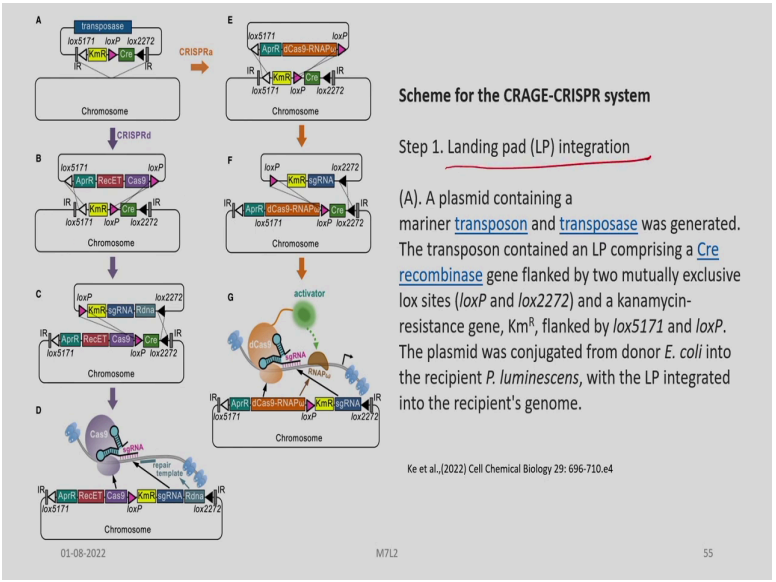
(Refer Slide Time: 29:13)



So, these CRAGE systems can be ideal platforms for implementing CRISPR based systems and it helps in overcoming a major limitation in non model bacteria.

The CRISPR on CRAGE-duet system has been used to characterize secondary metabolite biosynthetic gene clusters or BGCs as a model system. They demonstrated that the CRAGE-CRISPER system is a simple yet compelling approach to biosynthetic gene cluster characterization.

(Refer Slide Time: 29:57)



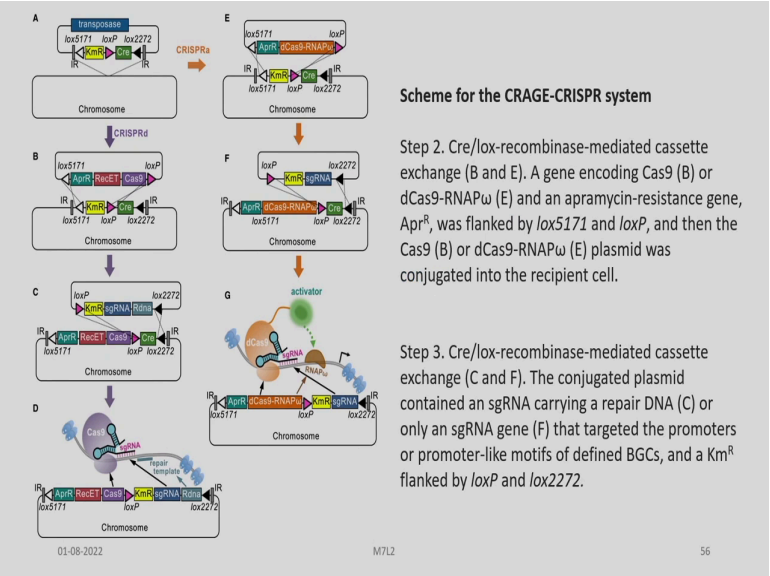
Here you can see the scheme for the CRAGE-CRISPER system. In step 1 as we all know the landing pad integration occurs.

In the figure you can see a plasmid containing a mariner transposon and transposase was generated. The transposon containing an LP comprising a Cre recombination gene flanked by two mutually exclusive lox sites and a kanamycin-resistance genes, flanked by the other two lox sites are shown. The plasmid was conjugated from donor *E. coli* into the recipient with the landing pad integrated into the recipient's genome.

So, here you can see the scheme for the CRAGE-CRISPR system. In step 1 as we have already discussed the landing pad integration takes place. And you can see the location of the various components like the mariner transposon and the transposase and the transposon contained an LP comprising a Cre recombination gene flanked by two mutually exclusive lox sites and the kanamycin resistance gene.

The plasmid was conjugated from donor *E. coli* into the recipient *P. luminescence* with the LP integrated into the recipient's genome.

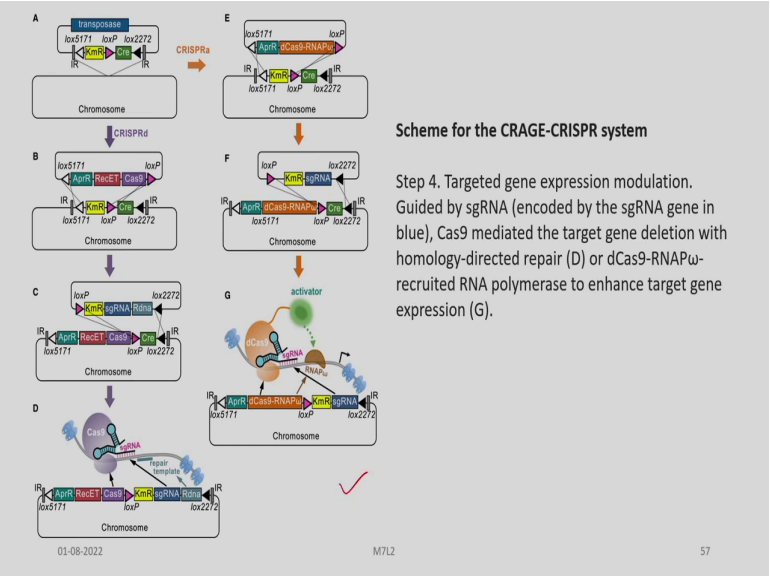
(Refer Slide Time: 31:27)



In step 2, the Cre lox recombination-mediated cassette exchange occurs as shown in B and E. A gene encoding Cas9 or dCas9-RNAP omega and an apramycin-resistance genes, Ap^R, was flanked by lox5171 and loxP and then the Cas9 b or dCas9-RNAPw plasmid was conjugated into the recipient cell.

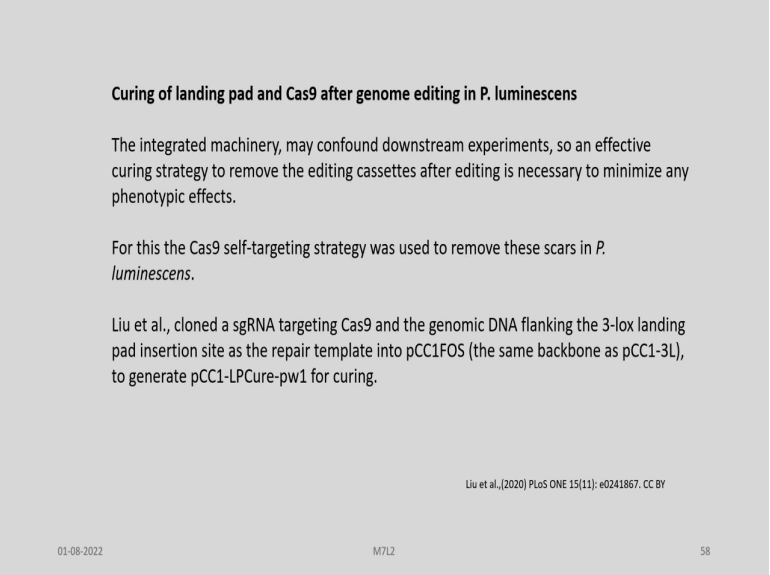
In a step 3, Cre lox-recombinase-mediated cassette exchange occurs (C and F). The conjugated plasmid contained a single guide RNA carrying a repair DNA or only a single guide RNA, as shown in F, that targeted the promoters or promoter-like motifs of defined BGCs, and a Km^R flanked by loxP and lox 2272.

(Refer Slide Time: 32:27)



In step 4, targeted gene expression modulation is depicted. Guided by these single guide RNA, Cas9 mediated the target gene deletion with homology-directed repair (D) or dCas9-RNAP omega recruited RNA polymerase to enhance target gene expression, as shown in figure G.

(Refer Slide Time: 33:01)

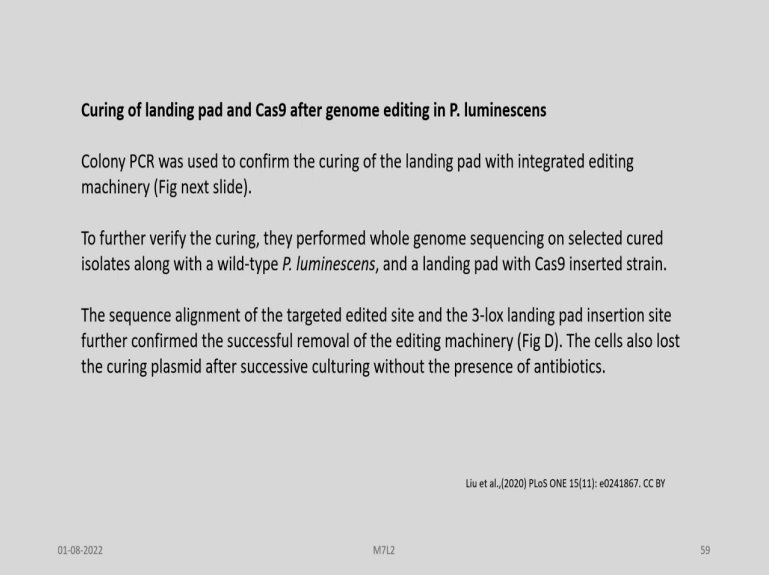


Curing of landing pad and Cas9 after genome editing in *P. luminescens*. The integrated machinery may confound downstream experiments, so, an effective curing strategy to remove the editing cassettes after editing is necessary to minimize any phenotypic effects.

For these the Cas9 self-targeting strategy was used to remove these scars in *P. luminescens.*

Liu et al., cloned a single guide RNA targeting Cas9 and the genomic DNA flanking the 3-lox landing pad insertion site as a repair template into pCC1FOS (the same backbone as pCC1-3L), to generate pCC1-LP cure pw1 for curing.

(Refer Slide Time: 33:47)

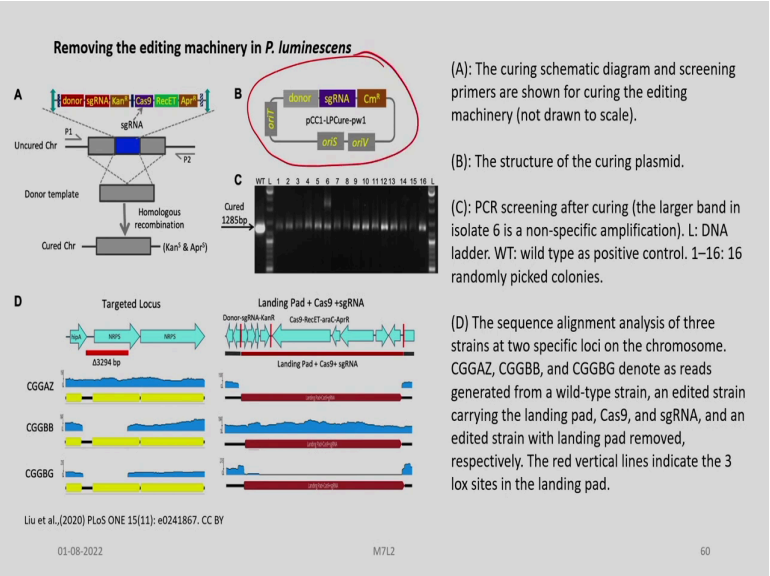


Colony PCR was used to confirm the curing of the landing pad with integrated editing machinery.

To further verify the curing, they performed whole genome sequencing on selected cured isolates along with a wild-type P. *luminescens*, and a landing pad with Cas9 inserted strain.

The sequence alignment of the targeted edited site and the 3-lox landing pad insertion further confirmed the successful removal of the editing machinery. The cells also lost the curing plasmid after successive culturing without the presence of antibiotics.

(Refer Slide Time: 34:20)



So, here you can see the removing of the editing machinery in *P. luminescens*. In step A, the curing schematic diagram and screening primers are shown for curing the editing machinery. The structure of the curing plasmid is shown in B, comprising of donor sgRNA and other components.

In figure C, we can see the PCR screening after curing and in figure D, we can see the sequence alignment analysis of the three strains at two specific loci on the chromosome. CGGAZ, CGGBB, and CGGBG denotes as the reads. Here these are generated from a wild-type strain, an edited strain carrying the landing pad, Cas9, and sgRNA, and an edited strain with landing pad being removed, respectively.

The red vertical lines indicate the 3 lox sites in the landing pad.

(Refer Slide Time: 35:31)

By integrating all the components into the chromosome, overall, in this particular experiment Liu et al., have demonstrated that genome editing in bacterial cells without a need for a replicating plasmid system is possible and the integrated landing pad and Cas9 can be removed to achieve scarless editing of bacterial genome. With this we come to the end of this lecture.

Thank you.