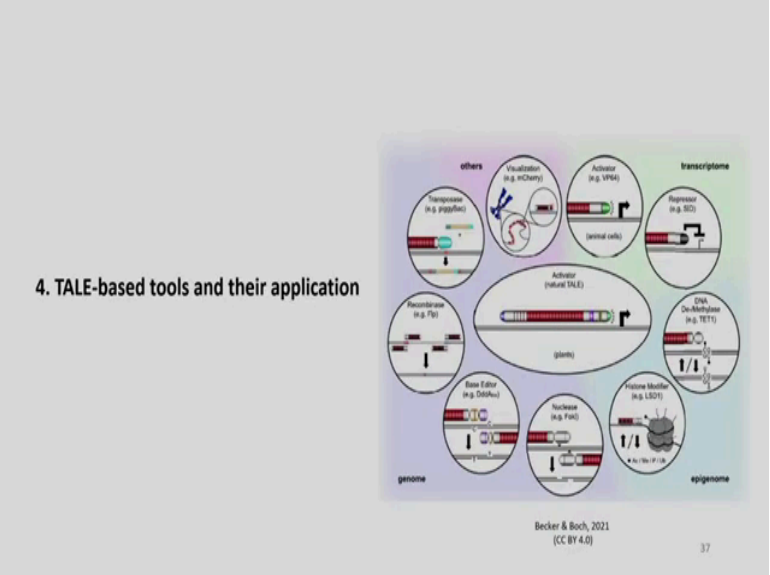
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

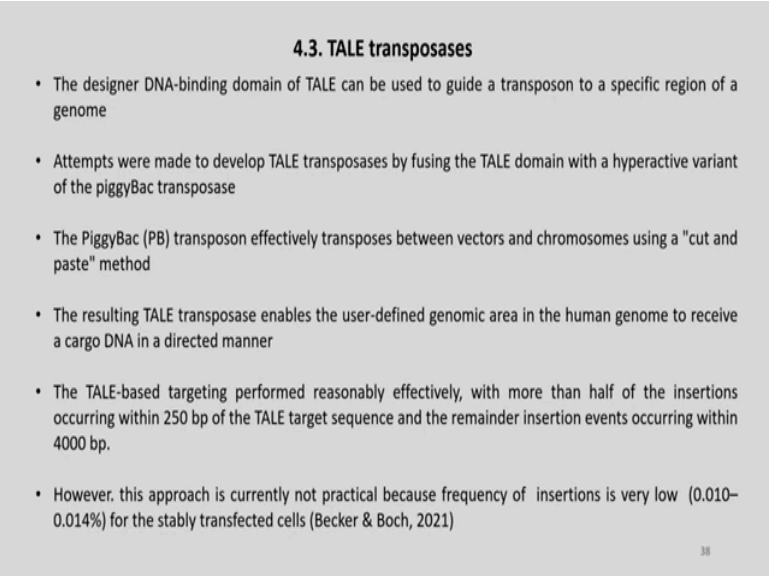
Module - 06
Transcription activator-like effector nuclease (TALEN) Technology
Lecture - 03
Application of TALEN - Part B

(Refer Slide Time: 00:36)



Welcome back to my lecture on Applications of TALEN. So, part B. So, we have discussed about some of the promises and potentials of applications of TALEN and where it can be used apart from cleaving DNA molecules for example, as transposes and then for activation and repression and DNA methylation and so, on and so, forth.

(Refer Slide Time: 01:04)

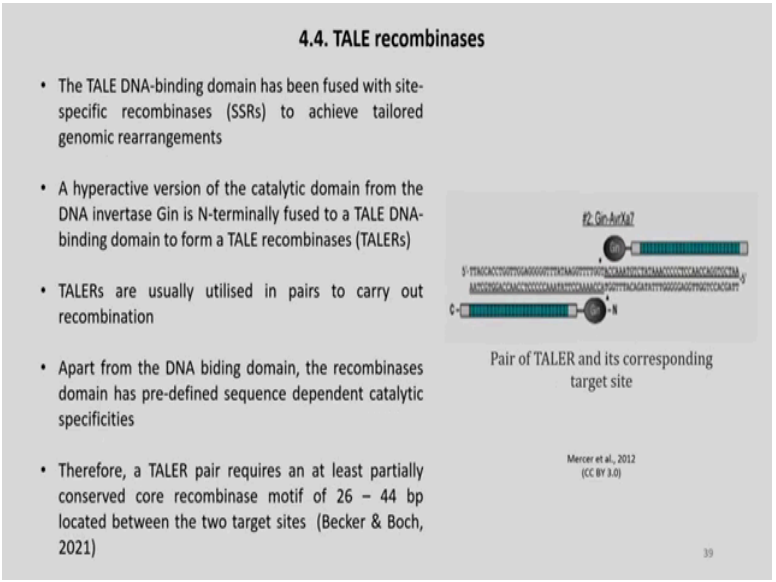


So, let us continue our discussion with TALE transposases. So, the designer DNA-binding domain of TALE can be used to guide a transposon to a specific region of a genome. Attempts were made to develop TALE transposes by fusing the TALE domain with a hyperactive variant of the piggyBac transposes. The piggyBac PB transposon effectively transposes between vectors and chromosomes using a cut-and-paste method.

The resulting TALE transposes enables the user-defined genomic area in the human genome to achieve a cargo DNA in a directed manner. The TALE-based targeting performed reasonably effectively with more than half of the insertions occurring within 250 base pairs of the TALE target sequence and the remainder insertion events occurring within 4000 base pairs.

However, this approach is currently not practical because the frequency of insertion is very low 0.01 percent or to 0.014 percent for the stably transfected cells, but the potential of the technology can be displayed through this work.

(Refer Slide Time: 02:31)

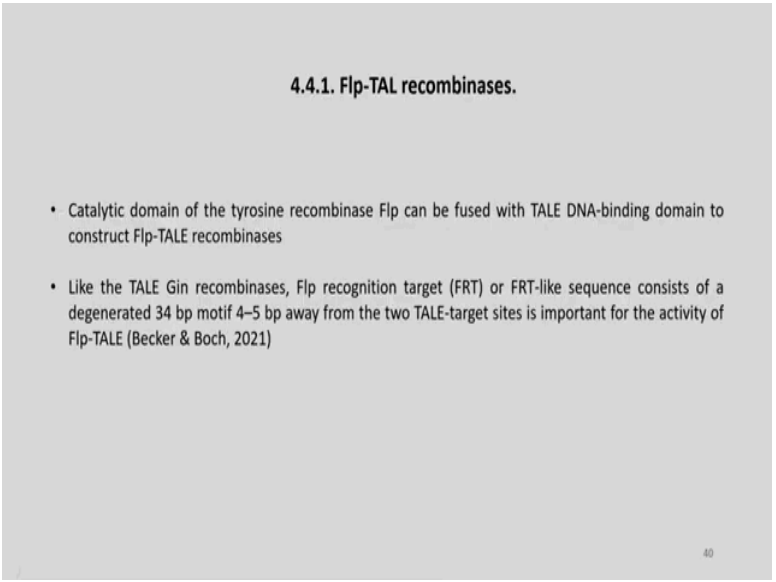


So, let us have a discussion on the TALE recombinases. The TALE DNA-binding domain has been fused with site specific recombinases (SSRs) to achieve tailored genomic rearrangements.

A hyperactive version of the catalytic domain from the DNA invertase Gin is N- terminally fused to a TALE DNA-binding domain to form TALE recombinases or TALERs. So, you can see here the Gin which is being fused to the N-terminal of the TALE and these TALERs are usually utilized in pairs to carry out recombination.

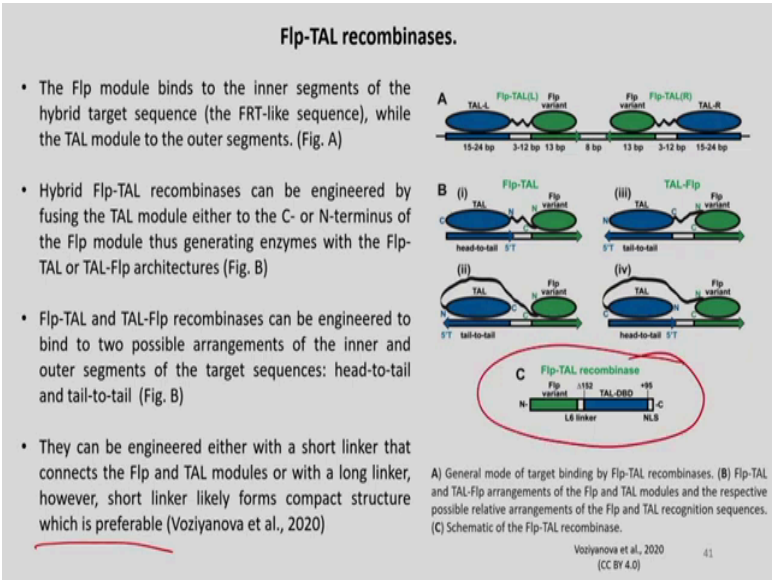
Apart from a DNA binding domain, the recombinase domain has predefined sequence dependent catalytic specificity. Therefore, a TALER pair requires at least partially conserved core recombination motif of around 26 to 44 base pair located between the two target sites Flp-TAL recombinases.

(Refer Slide Time: 03:55)



Catalytic domain of the tyrosine recombinases Flp can be fused with TALE DNA-binding domain to construct the Flp-TALE recombinases. Like the TALE Gin recombinases, Flp recognition target FRT or FRT like sequence consists of a degenerated 34 base pair motif 4 to 5 base pair away from the two TALE-target sites which is important for the activity of the Flp-TALE.

(Refer Slide Time: 04:30)



So, let us see the schematics of these Flp-TAL rrecombinases. In (A), you can see the general mode of target binding of Flp-TAL recombinases. In (B) you can see the Flp-TAL and

TAL-Flp arrangement of the Flp-TAL modules and respective possible relative arrangements of the Flp and TAL recognition sequences and (C) is the schematic of the Flp-TAL recombinases here.

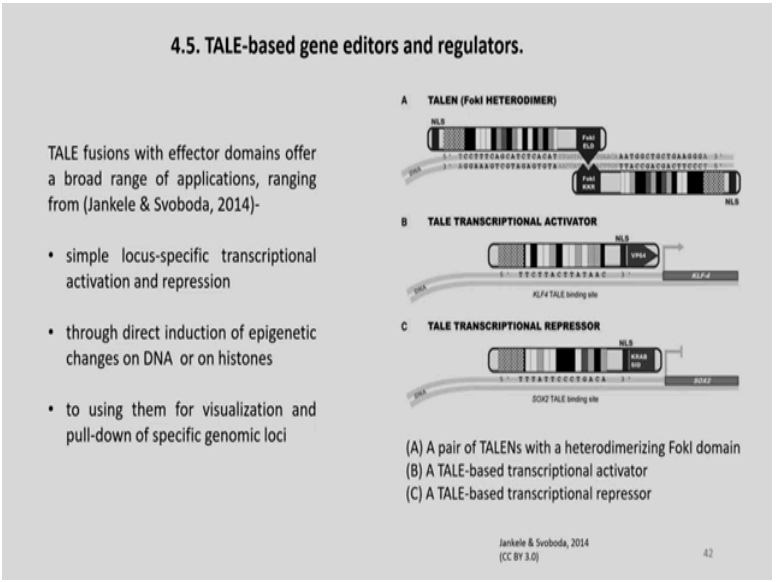
So, the Flp molecule binds to the inner segments of the hybrid target sequence called the FRT-like sequence while the TAL module to the outer segments. The hybrid Flp-TAL recombinases can be engineered by fusing the TAL module either to the C or N terminus of the Flp module, thus generating enzymes with the Flp-TAL or TAL-Flp architectures you can see in figure B.

So, this Flp-TAL you can see here the directionality and this is the TAL and the Flp is joined in this N terminus and here it is joined in the C terminus. So, you have Flp-TAL or TAL-Flp. This is a head-to-tail and this is a tail-to-tail and then there is another kind of an arrangement here you can see tail-to-tail C terminus to N and then this head-to-tail orientation in this figure number iv.

The Flp-TAL and TAL Flp recombinases can be engineered to bind to two possible arrangements of the inner and outer segments of the target sequences head-to-tail and tail-to-head as shown in figure B. They can be engineered either with a short linker that connects the Flp and TAL modules or with a long linker; however, short linker likely forms complex structure which is preferable.

So, here you see this is the short linker and these are the long linkers these are preferred over these because these linkers will add mass and volume and the complex structures are much more preferred. So, we already spoke about this Flp-TAL recombinase construct.

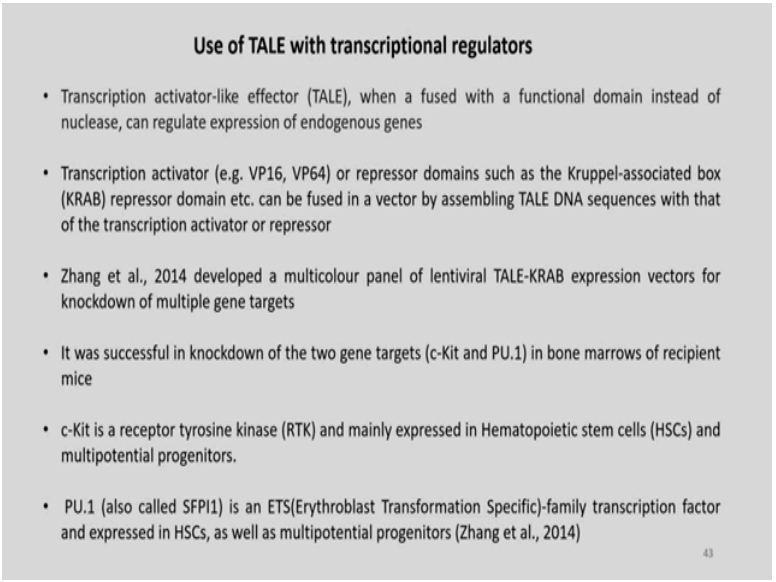
(Refer Slide Time: 07:24)



Let us have some discussion on the TALE-based gene editors and regulators. So, here you can see the TALEN-FokI heterodimer and there is a TALE transcriptional activator this is a TALE transcriptional repressor.

So, these FokI domains are different and that is why they are heterodimers and this is a TALE-based transcriptional activator and number C is the TALE based transcriptional repressor. So, TALE fusion with effective domains offer a broad range of applications like simple locus specific transcriptional activation and repression and through direct induction of epigenetic changes on DNA or on histones and using them for visualization and pull down of specific genomic loci.

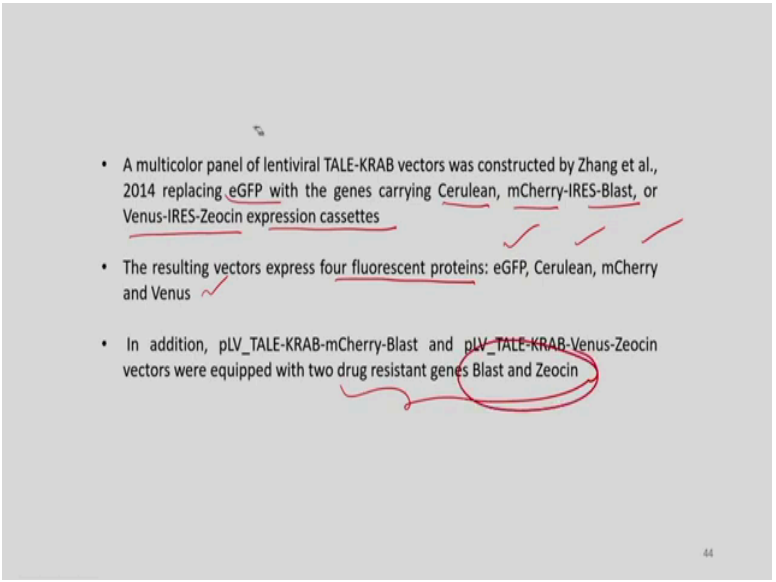
(Refer Slide Time: 08:31)



The application of TALE with transcriptional regulations. Transcription activator like effector when fused with a functional domain instead of nucleus, can regulate expression of endogenous genes. Transcription activators example VP16 VP64 or repressor domains such as the Kruppel-associated box KRAB repressor domain can be fused in a vector by assembling TALE DNA sequences with that of the transcription activator or repressor.

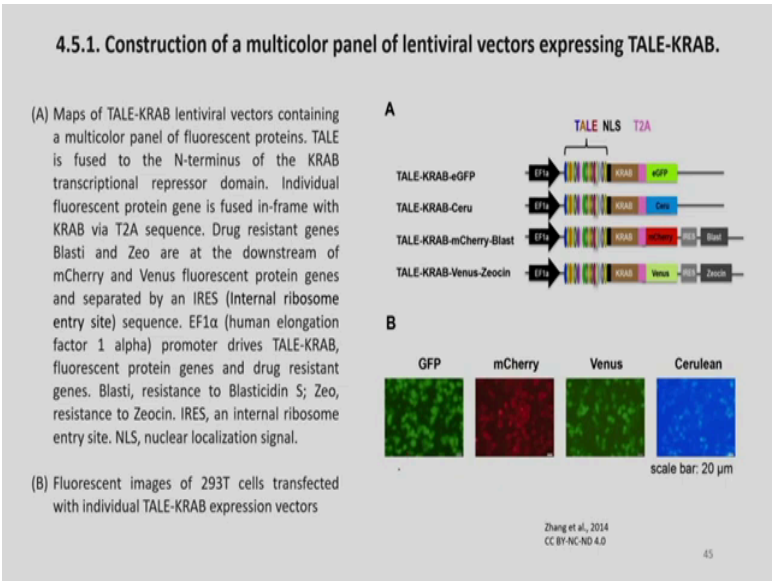
In 2014 Zhang et al developed a multicolour panel of lentiviral TALE-KRAB expression vectors for knockdown of multiple gene targets. It was successfully knockdown of the two gene targets c-Kit and PU.1 in bone marrows of recipient mice. c-kit it is a receptor tyrosine kinase and mainly expresses in Hematopoietic stem cells in multi potential progenitors. PU.1 also call SFPI1 is an ETS familiar transcription factor and expressed in hematopoietic stem cells as well as multi potential progenitors.

(Refer Slide Time: 10:05)



A multicolor panel of lentiviral TALE-KRAB vectors constructed by Zhang et al replacing eGFP with the genes carrying Cerulean, mCherry-IRES-Blast, or Venus-IRES-Zeocin expression cassettes was accomplished. The resulting vectors express four fluorescent proteins eGFP, Cerulean, mCherry and Venus. In addition pLV_TALE-KRAB mCherry blast and pLV_TALE-KRAB Venus zeocin vectors were equipped with two drug resistant genes Blast and Zeocin.

(Refer Slide Time: 10:47)

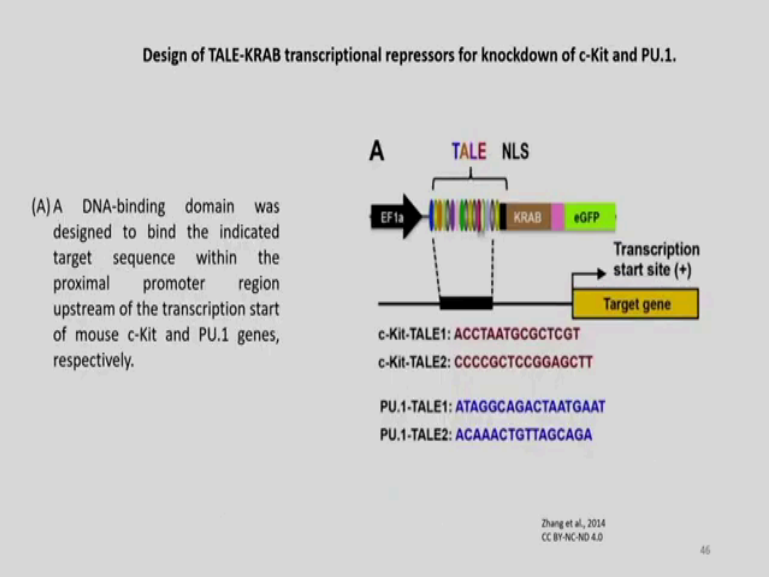


So, this is the construction of a multicolor panel of lentiviral vectors which we have discussed and these are the various fluorescence genes eGFP, Ceru, mCherry and Venus, and these last two mCherry and Venus constructs are also given these drug resistance genes blast and IRES and Zeomycin. And when transfected into cells you can see the expression of these various fluorescence proteins.

Let us look into the details of this experiments. So, in A you can see the maps of the TALE-KRAB lentiviral vectors containing a multi color panel of fluorescent proteins, TALE is fused to the N terminus of the KRAB transcriptional repressor domain here individual fluorescent protein gene is fused in frame with KRAB via the T2A sequence drug resistant genes.

Blast and Zeo are at the downstream of mCherry and Venus fluorescent protein genes and separated by an IRES internal ribosome entry site sequence EF1α human elongation factor one alpha promoter drives the TALE-KRAB fluorescent protein genes and drug resistant genes, Blasti resistance to Blasticidin S; Zeo resistance to Zeocin. IRES and internal ribosome entry site NLS is the nuclear localization signal and these are the images of the 293T cells transfected with individuals TALE-KRAB expression vectors.

(Refer Slide Time: 12:51)

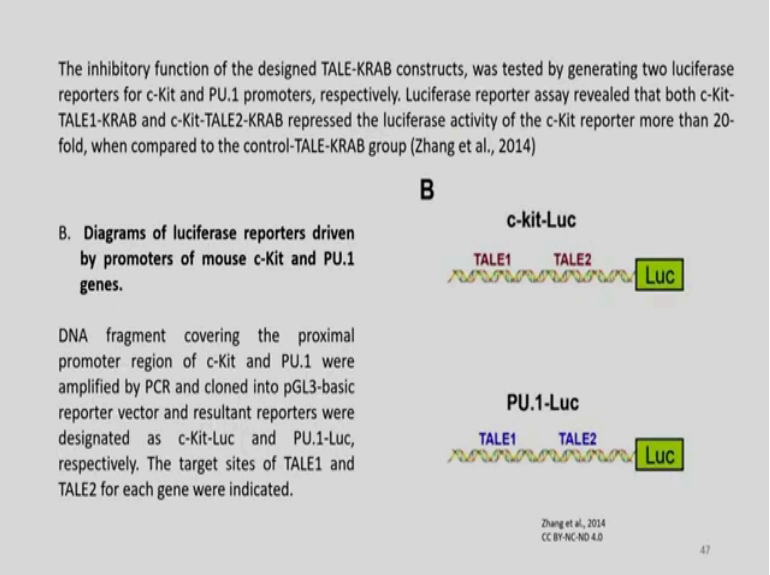


Design of TALE-KRAB transcriptional repressor for knockdown of c-Kit and PU.1 a DNA-binding domain was designed to bind the indicated target sequence. We need a

proximal promoter reason upstream of the transcriptional start of mouse c-Kit and PU.1 genes, respectively.

And you have these constructs already discussed earlier with KRAB and EF1 alpha and this is the target gene and this is the transcription start site.

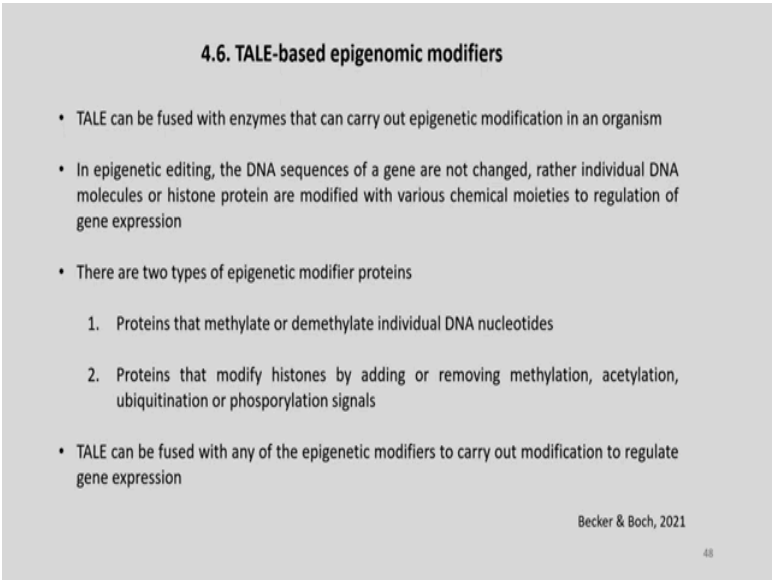
(Refer Slide Time: 13:33)



The inhibitory function of the design TALE-KRAB constructs was tested by generating two luciferase reporters for c-Kit and PU.1 promoters respectively. The luciferase reporter assay revealed that both c-Kit TALE1-KRAB and c-Kit TALE2 KRAB repress the luciferase activity of the c-Kit reporter more than 20-fold when compared to the control TALE-KRAB group.

So, here in this diagram you can see luciferase reporters driven by the promoters of mouse c-Kit and PU.1 genes on top and bottom respectively. The inhibitory function of the design TALE-KRAB constructs were tested by generating two luciferase reporters for c-Kit and PU.1 promoters respectively. Luciferase reporter assays revealed that both c-Kit TALE 1 KRAB and c-Kit TALE 2 KRAB repressed the luciferase activity of the c-Kit reporter more than 20-fold, when compared to the control TALE-KRAB group.

(Refer Slide Time: 15:05)

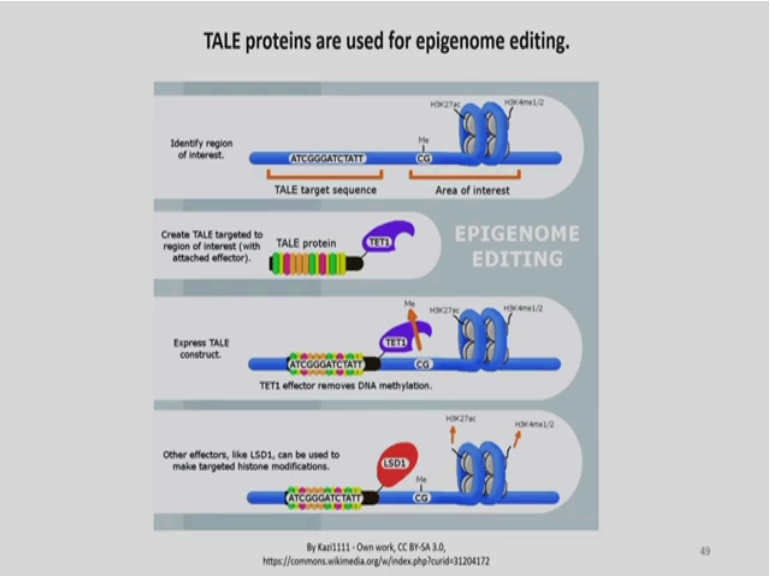


TALEs have also been used to make fusion proteins which can be used for epigenomic modification. So, fusion of TALE with enzymes which carry out epigenetic modification in an organism has been found to be quite promising. In epigenetic editing, the majority of the editing we have been discussing till now are genetic editing we are now using these TALE based approach for epigenomic editing here the DNA sequences of the gene are not changed.

So, there is actually no any mutation rather individual DNA molecules or histone proteins are modified with various chemical moieties to regulate gene expression. So, we have to know the difference between epigenetic editing and genetic editing. In genetic editing there is a change of the genetic sequence, but in epigenetic editing there is no any change of the genome genetic sequence, but some modification either in the base or in the histone protein takes place and this can be done with TALE based epigenomic modifiers.

Which are of two types mostly based on the modifier proteins, proteins that methylate or demethylate individual DNA nucleotides and proteins that modify histones by adding or removing methylation, acetylation, ubiquitination or phosphorylation signals. TALEs can be fused with any of these above epigenetic modifiers to carry out modification to regulate gene expression.

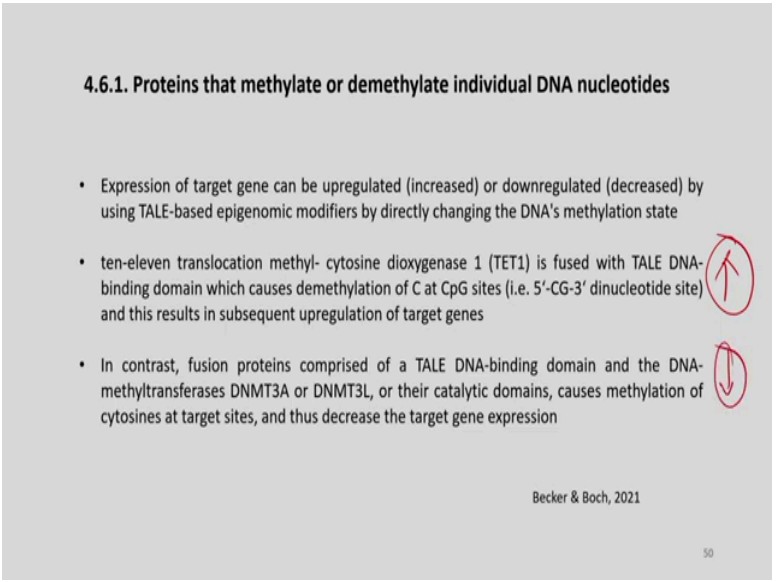
(Refer Slide Time: 17:13)



So, here you can see some schematics of TALE proteins used for epigenomic editing. So, first we need to identify the reason of interest. This is a TALE target sequence and this is the area of interest. So, the TALE protein will bind to the TALE target sequence and the modifier will act on the area of interest. So, you can see here some kind of modification being done on the DNA bases and here modifications are being done on the histone proteins around which the DNA wraps.

So, we can create TALE targeted to regions of interest with an attached effector this is a TALE protein and this is a TET1. So, we can express the TALE construct. So, this TET effector removes DNA methylation. So, the TALE binds to this DNA TALE target and these modifier TET 1 will remove these methyl group from here, ok. So, these is being used for demethylation. So, other effectors like LSD1 can be used to make targeted histone modifications here. So, H3K27-Ac or H3K4-Me 1/2. So, these are in brief the concepts of epigenome editing.

(Refer Slide Time: 19:16)

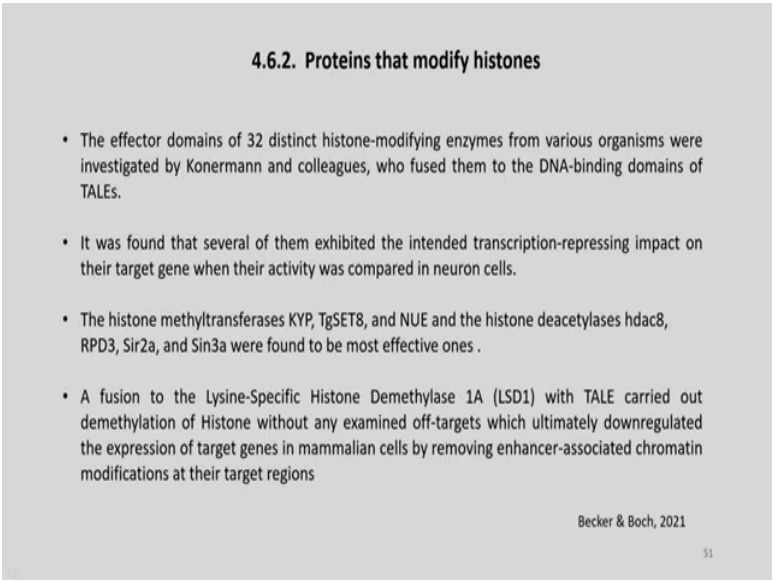


So, let us first discuss about the proteins that methylate or demethylate individual DNA nucleotides.

Expression of target genes can be upregulated or increased or downregulated (decreased) by using TALE-based epigenomic modifiers by directly changing the DNA's methylation state. Ten-eleven translocation methylcytosine dioxygenase 1 or TET 1 (ten-eleven translocation one) is fused with TALE DNA-binding domain which causes demethylation of C at CpG sites which we have shown in the figure earlier and this results in subsequent upregulation of target genes.

In contrast fusion proteins comprised of a TALE DNA-binding domain and the DNA-methyltransferases DNMT3A or DNMT3L or their catalytic domains causes methylation (the opposite of TET1) of cytosine at target sites and thus decrease the target gene expression. So, by TET1 we can upregulate gene expression and using DNMT3A and DNMT3L we can downregulate gene regulation; the first demethylates, the second one just does the reverse: it methylates.

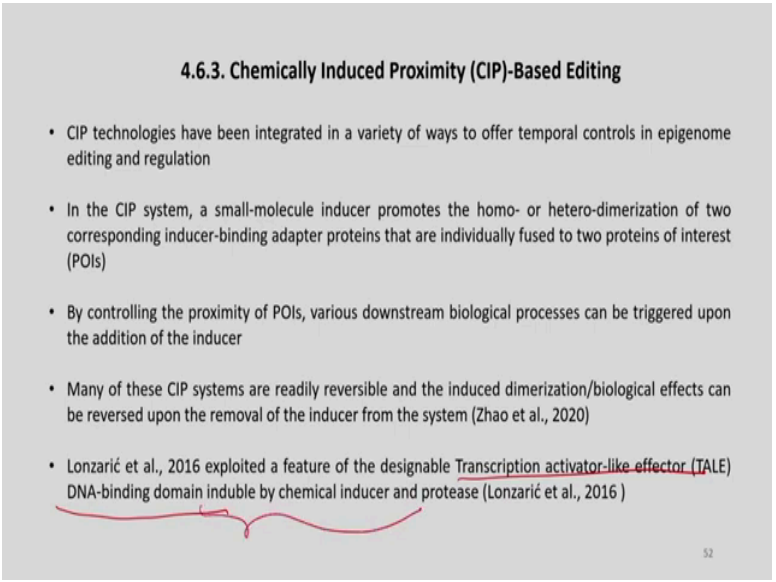
(Refer Slide Time: 20:54)



The second type are the proteins that modify histones. The affective domains of 32 distinct histone modifying enzymes from various organisms have been investigated by Konermann and his associates and fused them to the DNA-binding domain of TALEs. They found that several of them exhibited the intended transcription repressing impact on the target gene when their activity was compared in neuron cells. The histone methyl transferases KYP, TgSET8 and NUE and the histone deacetylases is hdac8, RPD3, Sir2a and Sin3a were found to be the most effective ones.

A fusion to the Lysine-Specific Histone Demethylase 1A (LSD1) with TALE carried out demethylation of Histone without any examined off-targets which ultimately downregulated the expression of target genes in mammalian cells by removing enhancer-associated chromatin modifications at their target region.

(Refer Slide Time: 22:12)



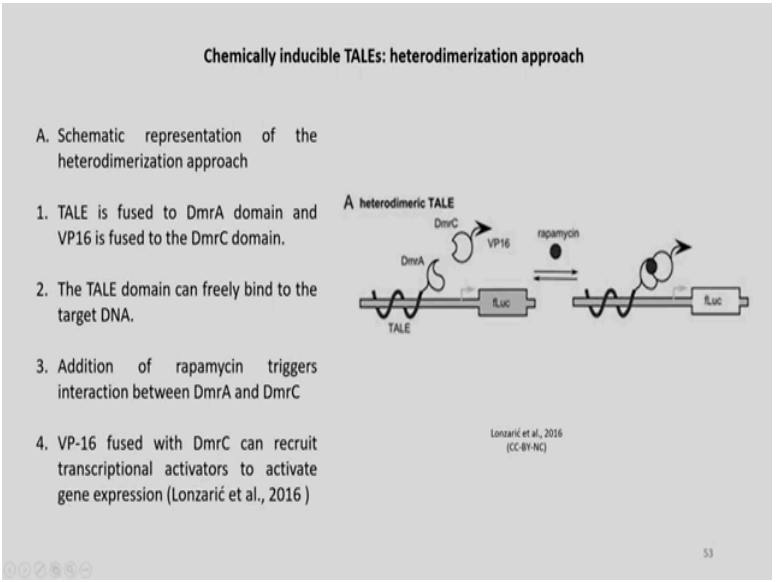
Let us now discuss about one interesting topic where we can use chemicals to do base editing and this is known as chemically induced proximity-based editing or CIP-based editing. CIP technologies have been integrated in a variety of ways to offer temporal controls in epigenomic editing and regulation.

In a CIP system a small molecule inducer promotes the homo or hetero dimerization of two corresponding inducer binding electro proteins that are individually fused to two proteins of interest by controlling the proximity of POIs proteins of interest various downstream biological processes can be triggered upon the addition of the inducer.

Many of these chemically induced proximity systems are readily reversible and the induced dimerization or biological effects can be reversed upon the removal of the inducer from the system. So, the system is operational in the presence of the inducer if you remove the inducer we can switch off the system giving us total control.

Lonzaric and his group exploited a feature of the designable Transcription activator like effector DNA-binding domain inducible by chemical inducer and a protease.

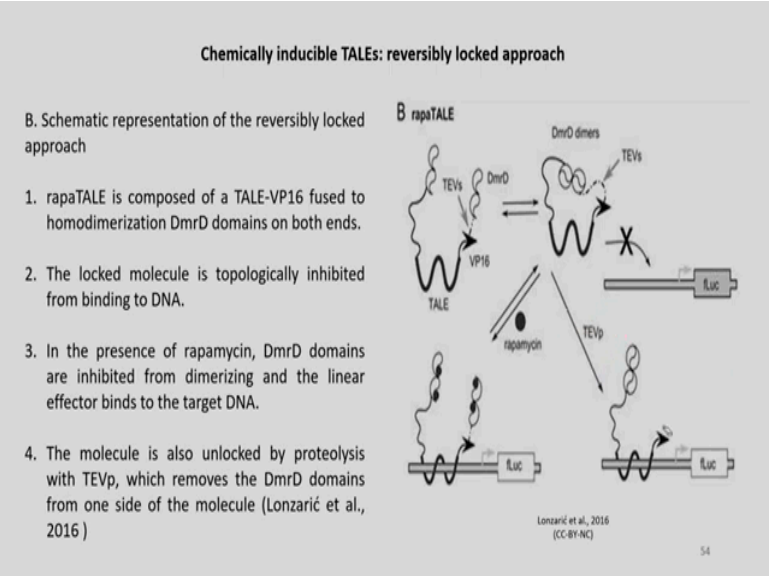
(Refer Slide Time: 23:53)



So, these are called as chemically inducible TALEs and here the heterodimerization approach is followed. So, there is a schematic representation of the heterodimerization approach. So, you can see here the TALE bound to DmrA and DmrC bound to VP16 in the presence of some molecule like rapamycin here they dimerize DmrA TALE DmrC VP16 will dimerize and some effect will be there as a result of this dimerization.

So, TALE is fused to DmrA domain and VP16 is fused to the DmrC domain. The TALE domain can freely bind to the target DNA and we know the mechanism of TALE binding to specific DNA targets. And here the addition of rapamycin triggers interaction between DmrA and DmrC leading to their dimerization. The VP16 fused with DmrC can recruit transcriptional activators to activate gene expression. So, these VP16 DmrC cannot itself bind to the DNA we are using TALE fused with DmrA which binds to DNA and which allows the DmrC bound to VP16 bound to DmrC to form active partners which can switch on the gene expression.

(Refer Slide Time: 25:56)

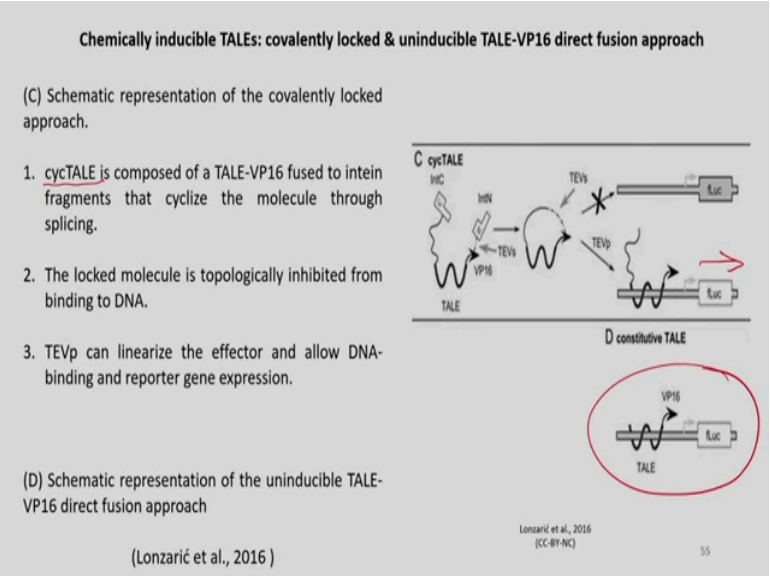


So, there is another approach the reversibly locked approach to do different kind of work and these are also chemically inducible TALEs and this is the schematic representation of the reversibly locked approach let us study the figure first. So, this is a rapaTALE which means there is a TALE domain and you can see this with the VP16 bound here and some other components and you see here DmrD and then under such conditions these rapamycin bind to these domains and it does not allow them to lock or bind to one another.

And here you can see some other kind of binding orientation let us now try to understand which will switch on the gene and which will switch off the gene. This locked rapaTALE is composed of TALE VP16 fused to homo dimerization domains on both end. DmrD is here and DmrD is also here. The locked molecule is topologically inhibited from binding to DNA, ok.

In the presence of rapamycin DmrD domains are inhibited from dimerizing and the linear effector binds to the target DNA. The molecule is also unlocked by proteolysis with TEVp which removes the DmrD domain from one side and now this molecule is able to bind to the DNA easily because there are no any restrictions anymore.

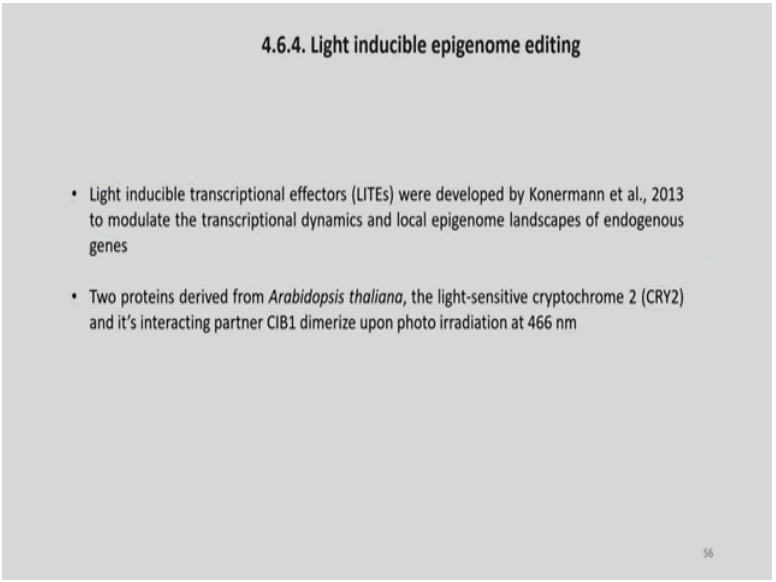
(Refer Slide Time: 28:31)



Let us look into covalently locked and un-inducible TALE VP16 direct fusion approach. The schematic representation of the covalently locked approach, the cycTALE is composed of TALE VP 6 fused to intein fragments that cyclize the molecule through splicing. So, there a splicing occurring here due to this there is a cyclization. The locked molecule is topologically inhibited from binding to DNA. TEVp can linearize this effector and allow DNA-binding and reporter gene expression and here it is unable to bind and there is no any gene expression.

So, in (D) you can see the schematic representation of the uninducible TALE-VP16 direct fusion approach.

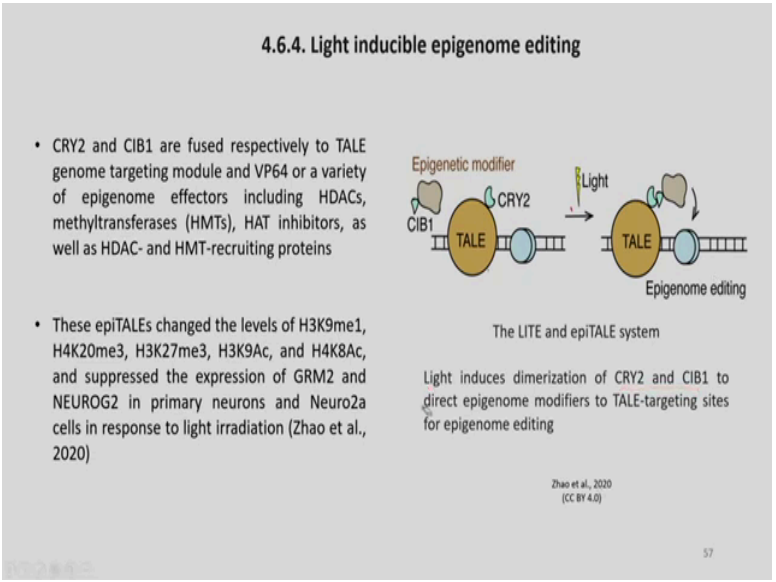
(Refer Slide Time: 29:53)



Apart from inducing TALEs with the help of chemicals we can also induce them with the help of light. So, these are the light-inducible epigenome editing approach. So, light-inducible transcriptional effectors LITEs were developed by Konermann et al in 2013 to modulate the transcriptional dynamics and local epigenome landscapes of endogenous genes.

Two proteins derive from Arabidopsis thaliana, the light-sensitive cryptochrome 2 (CRY2) and its interacting partners CIB1 dimerize upon photo irradiation at 466 nanometers. So, these CRY2 and CIB1 when exposed to light they will form dimers.

(Refer Slide Time: 30:55)

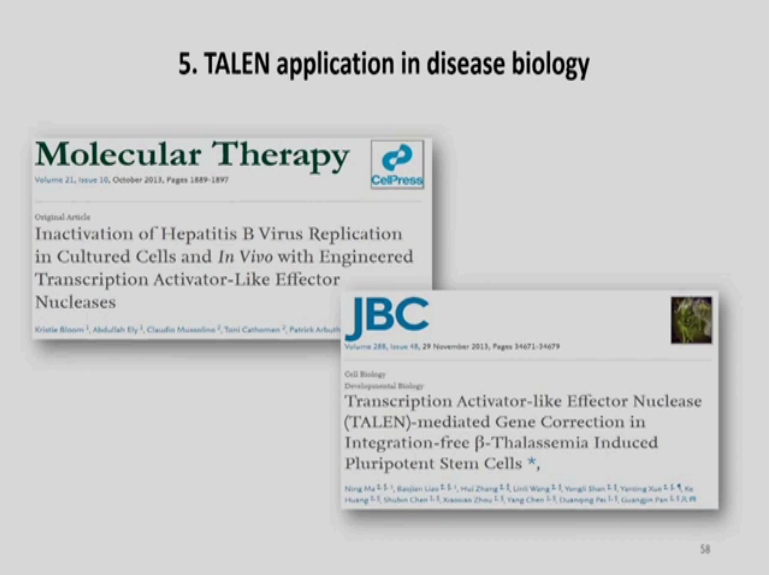


So, you can see here CIB1 and CRY2 in the presence of light, they are forming dimers and here we are using TALE to produce a TALE CRY2 fusion protein and this dimerization of CIB1 and CRY2 is exploited for epigenome editing.

CRY2 and CIB1 are fused respectively to TALE genome targeting module and VP64 or a variety of epigenome effectors including HDACs methyl transferase, HAT inhibitors as well as HDAC and HMT-recruiting proteins. These epiTALEs change the levels of H3K9me1, H4K20me3, H3K27me3 and several others and suppress the expression of GRM2 and NEUROG2 in primary neurons and Neuro2a cells in response to light radiation.

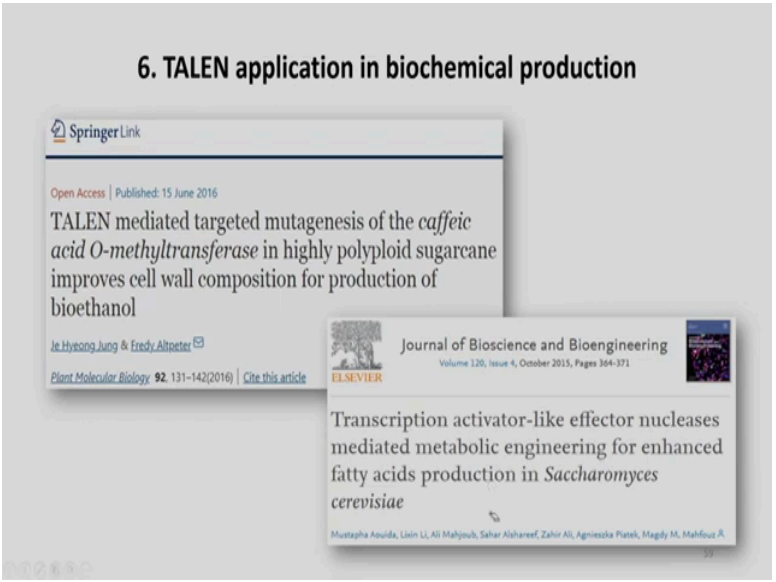
So, you can see in this figure that light induces the dimerization of CRY2 and CIB1 to direct epigenome modifiers to TALE-targeting sites for epigenome editing.

(Refer Slide Time: 32:35)



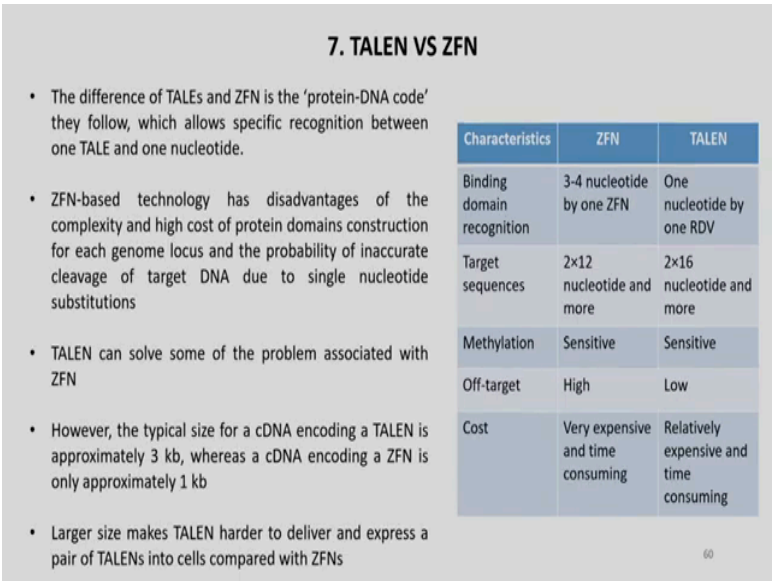
TALEN has been widely used in many applications in disease biology like inactivation of Hepatitis B virus replication in cultured cells and this has a potential to be translated into therapies. Then also TALEN mediated gene correction in integration free beta-thalassemia induced pluripotent stem cells.

(Refer Slide Time: 33:03)



And for the application in biochemical production example TALEN mediated targeted mutagenesis of caffeic acid O-methyl transferase in highly polyploid sugarcane improve cell-level composition for the production of bioethanol. So, in today's world where there is lot of demand for bioethanol TALEN is indeed a very very promising technology and others like transcription activator like effector nucleases used for metabolic engineering for enhanced fatty acid production in Saccharomyces cerevisiae.

(Refer Slide Time: 33:50)



So, this can have huge impact in the food industry as well as the fuel industry. So, before we end let us make a comparison between the two technology platforms we have discussed so, far. We discussed in both the cases their basic biology and then we discussed how they are constructed or synthesized and what are the bioinformatics tools or softwares available to carry out their design or their optimized operations.

So, TALEN and ZFN are mostly based on the FokI nuclease protein and which is fused to a DNA-binding protein domain which is zinc finger in the case of ZFN and which are TALEs in the case of TALEN. So, the difference of TALEs and ZFN is the protein DNA code they follow which allows specific recognition between one TALE and one nucleotide. ZFN based technology has disadvantages of the complexity and high cost of protein domain construction for each genome locus.

And the probability of inaccurate cleavage of target DNA due to single nucleotide substitutions TALEN can solve some of the problems associated with ZFN; however, typical size for a cDNA coding encoding a TALEN is approximately 3 kb whereas, a cDNA encoding ZFN is only approximately 1 kb. So, larger size makes TALEN harder to deliver and express a pair of TALENs into cells compared with ZFNs and we have seen how large constructs with both the ZFN pairs can be loaded into a single vector and ensure its efficiency and efficacy.

Now, let us discuss in brief a little bit about the two technology platforms which we have discussed so, far TALEN versus ZFN and we know they are mostly based on the nuclease called FokI and a binding domain which is TALE in the case of TALEN and zinc fingers in the case of ZFN and both are proteins. So, these are mostly DNA-binding proteins fused to a nuclease domain with which we can carry out genome editing.

Briefly the difference of TALEs in ZFN is the protein DNA code they follow which allows specific recognition between one TALE and one nucleotide. So, here in ZFN there are 3 to 4 nucleotides recognition by one zinc finger and here one nucleotide by one RVD. The ZFN-based technology has disadvantages of the complexity and high cost of protein domains construction for each genome locus and the probability of inaccurate cleavage of target DNA due to single nucleotide substitutions.

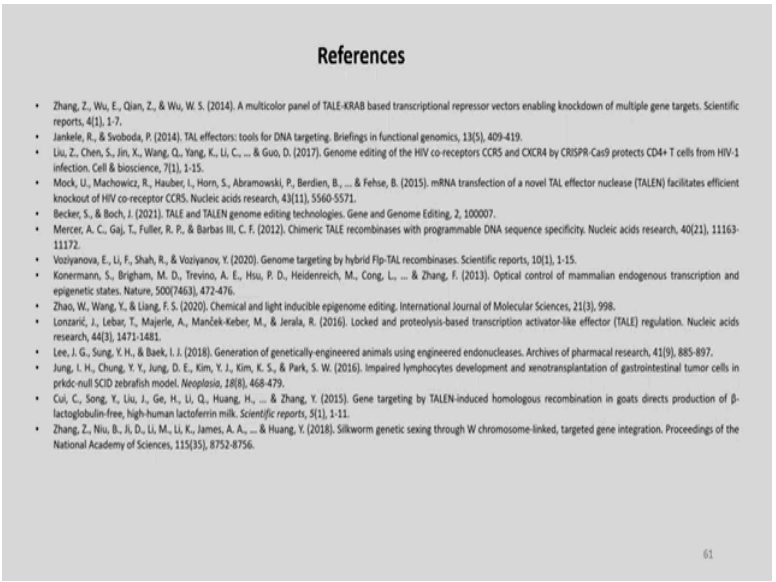
TALEN can solve some of the problems associated with ZFNs; however, the typical size for a cDNA and coding a TALEN is approximately 3 kb whereas, a cDNA encoding a ZFN is only

approximately 1 kb and we have discussed how both the left or forward or right or reverse ZFN can be loaded into a single vector and their co-expression can be ensured to increase the efficiency and efficacy of the process.

Larger size makes TALEN harder to deliver and express a pair of TALENs in to cells compared to ZFN. So, other things are as laid out in the field you can see that the target sequence 2×12 nucleotide and more this is little bit higher 2×16 both are sensitive methylations and then off target is quite high in ZFN which is a disadvantage.

And this is low in TALEN which is an advantage in spite of it being hard to deliver due to its bigger size and as already told the cost is expensive in case of ZFN and it is time consuming and difficult and TALEN is relatively inexpensive and less time consuming.

(Refer Slide Time: 38:51)



References

- Zhang, Z., Wu, E., Qian, Z., & Wu, W. S. (2014). A multicolor panel of TALE-KRAB based transcriptional repressor vectors enabling knockdown of multiple gene targets. Scientific reports, 4(1), 1-7.
- Jankele, R., & Svoboda, P. (2014). TAL effectors: tools for DNA targeting. Briefings in functional genomics, 13(5), 409-419.
- Liu, Z., Chen, S., Jin, X., Wang, Q., Yang, K., Li, C., ... & Guo, D. (2017). Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4+ T cells from HIV-1 infection. Cell & bioscience, 7(1), 1-15.
- Mock, U., Machowicz, R., Hauber, I., Horn, S., Abramowski, P., Berdien, B., ... & Fehse, B. (2015). mRNA transfection of a novel TAL effector nuclease (TALEN) facilitates efficient knockout of HIV co-receptor CCR5. Nucleic acids research, 43(11), 5560-5571.
- Becker, S., & Boch, J. (2021). TALE and TALEN genome editing technologies. Gene and Genome Editing, 2, 100007.
- Mercer, A. C., Gaj, T., Fuller, R. P., & Barbas III, C. F. (2012). Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic acids research, 40(21), 11163-11172.
- Voziyanova, E., Li, F., Shah, R., & Voziyanov, Y. (2020). Genome targeting by hybrid Flp-TAL recombinases. Scientific reports, 10(1), 1-15.
- Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., ... & Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. Nature, 500(7463), 472-476.
- Zhao, W., Wang, Y., & Liang, F. S. (2020). Chemical and light inducible epigenome editing. International Journal of Molecular Sciences, 21(3), 998.
- Lonzarić, J., Lebar, T., Majerle, A., Manček-Keber, M., & Jerala, R. (2016). Locked and proteolysis-based transcription activator-like effector (TALE) regulation. Nucleic acids research, 44(3), 1471-1481.
- Lee, J. G., Sung, Y. H., & Baek, I. J. (2018). Generation of genetically-engineered animals using engineered endonucleases. Archives of pharmacal research, 41(9), 885-897.
- Jung, I. H., Chung, Y. Y., Jung, D. E., Kim, Y. J., Kim, K. S., & Park, S. W. (2016). Impaired lymphocytes development and xenotransplantation of gastrointestinal tumor cells in prkdc-null SCID zebrafish model. *Neoplasia*, *18*(8), 468-479.
- Cui, C., Song, Y., Liu, J., Ge, H., Li, Q., Huang, H., ... & Zhang, Y. (2015). Gene targeting by TALEN-induced homologous recombination in goats directs production of β-lactoglobulin-free, high-human lactoferrin milk. *Scientific reports*, *5*(1), 1-11.
- Zhang, Z., Niu, B., Ji, D., Li, M., Li, K., James, A. A., ... & Huang, Y. (2018). Silkworm genetic sexing through W chromosome-linked, targeted gene integration. Proceedings of the National Academy of Sciences, 115(35), 8752-8756.

61

So, some of the references with which these lecture has been prepared.

Thank you for your patient hearing.