Review

Genome editing to epigenome editing: towards unravelling the enigmas in developmental biology

Suresh Kumar*

Division of Biochemistry, ICAR-Indian Agricultural Research Institute, New Delhi-110012, India.

ABSTRACT

Development of an organism depends on differential expression/regulation of genes in the genome to produce diverse cell types during the developmental processes. A typical animal/plant genome contains a certain/fixed number of genes, all of these are not expressed all the times. Depending on the differential expression of the genes under varying environmental conditions, the morphology/ physiology of the organism may vary. Epigenetic variations (DNA and histone modifications, and variation in small-RNA biogenesis) play important roles in the regulation of gene expression during the developmental process and environmental stresses. Genome editing helps unravelling the function(s) of a gene, and editing/correcting the gene of interest. Once we identify the epigenetic mark(s) associated with the trait of interest, we can use epigenome editing tools and techniques for manipulation of gene expression. Epigenome editing uses a fusion protein comprising a specific DNA recognition domain that recruits the attached enzymatic domain to the defined genomic site. Discovery of dCas9 provides a valuable tool for epigenome editing. However, gene expression is regulated depending on the dynamic and reversibly modifiable biological and chemical information in the epigenome/epitranscriptome. Recruitment of dCas9 fused with histone-acetyltransferase or Tet1 DNA-demethylase can be used to activate enhancers and thus gene expression. Even reversible recruitment of endogenous chromatin complexes to a genomic locus is possible in almost any cell type. Deploying heterochromatin complex (e.g. Hp1/Suv39h1) and chromatin-remodelling complex (e.g. BAF), there are possibilities of repression or activation of genes through epigenome. The recent developments with respect to base- and prime-editing promise to add precision in epigenome and epitranscriptome editing, respectively. This review provides newer insights into fine-tuning of gene expression through genome/epigenome editing, and thus might help unravelling some of the enigmas of developmental biology.

INTRODUCTION

Human population is predicted to cross 9 billion by 2050 that would require to increase the food production by 70% [1]. However, one of the challenges would be to produce more nutritive food from the continuously shrinking arable land and freshwater. Another important challenge would be to produce the food, feed and fodder in a safe and sustainable manner [2, 3]. Therefore, the conventional approaches might not be adequate to meet the predicted requirements, particularly in terms of the quality. Hence, there are needs to utilize the modern tools and techniques to further enhance productivity of crops with the limited natural resources. However, the biosafety and

---

*Email ids: sureshkumar3_in@yahoo.co.uk;
sureshkumar@iari.res.in

biosecurity issues of genetically modified organisms (GMOs) would have to be kept in mind [4, 5]. Genome editing promises to provide unprecedented opportunities not only for the manipulation of biological systems for a better understanding of the regulatory mechanisms but also for efficient manipulation of gene/genome having lesser safety issues, if any. Epigenomics and epimarks (epigenetic markers) associated with the traits of interest would enable functional integration of epigenetics and their usage towards improving crop plants.

Most of the plant scientists aim at developing better varieties with broader adaptation under the changing global climate. Since the beginning, researchers have been interested to decipher the mechanisms that enable plants adapt to diverse environmental conditions. Improving understanding of genomes and their functional characterization have paved the way for targeted genome and epigenome editing towards crop improvement. However, site-specific manipulation in the genome has been elusive in the past. But scientific efforts have made tremendous advances in targeted genetic manipulation and editing the genome.

1. Genome editing: a gene revolution technology

Targeted editing of gene, even at base level, has revolutionized the genetic manipulation of organisms. Earlier, this has been possible due to the availability of the nucleases with programmable, site-specific DNA-binding domains such as like zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALENs). However, with the advent of Clustered regularly interspaced short palindromic repeats/CRISPR-associated-9 nuclease (CRISPR/Cas9) gene/genome editing has become simpler and more efficient. CRISPR/Cas9, a bacterial prokaryotic adaptive immune system originally identified in *Streptococcus pyogenes*, makes double-strand based on sequence specificity provided by two interacting RNA moieties: (i) CRISPR RNAs (crRNA), and (ii) trans-activating RNAs (tracrRNA). It was subsequently found that even a single chimeric RNA molecule can serve the function of recruiting the Cas9 nuclease, which made its usage easier in genome editing. Thus, any sequence (~20 nucleotides long) in the gene/genome can be a target for editing if it possesses two essentially required sequences (i) a unique sequence in the genome, and (ii) a Proto-spacer Adjacent Motif (PAM) immediately downstream of the target sequence. PAM sequence is essential for identification of the target sequence. The Cas9 nuclease and the guide-RNA (gRNA) form a riboprotein complex and become active to bind to the target DNA. However, the spacer sequence of gRNA remains free to base pair with the target DNA. The extent of interaction/complementarity between the spacer and target DNA determines whether Cas9 will make double-strand break (DSB) or not. The Cas9 nuclease cleaves opposite strands of the target DNA 3-4 nucleotides upstream of the PAM sequence. Subsequently, the DSB is repaired by one of the two repair mechanisms (i) an error-prone Non-Homologous End Joining (NHEJ) pathway, and (ii) a less efficient, having high-fidelity, Homology Directed Repair (HDR) pathway (Figure 1).

2. Recent advances in genome editing technology

Discovery of a catalytically-dead (dCas9) mutant of *Streptococcus pyogenes* SpCas9 provides a commendable tool for manipulation of regulating gene expression [6]. However, more efficient recruitment of multiple effectors has become possible with novel tagging approaches. Recruitment strategy has also been combined with chemically inducible approach to achieve temporal control on transcription process by the recruitment of reversible synthetic modulators. Braun *et al.* [7] used FIRE–Cas9 for rapid and reversible recruitment of endogenous chromatin complex to a genomic locus. The exogenous activators/repressors require quite a long time to turn on and off the gene expression; use of synthetic ligands becomes active within minutes of recruitment enabling us to determine the linkage between epigenetic regulators.

Nucleases like Cas9 are being searched for a more effective and precise cleavage of different types of nucleic acids like single stranded DNA and RNA. A CRISPR-based endonuclease from Prevotella and Francisella 1 bacteria (Cpf1) is an alternative platform for genome editing. The CRISPR–Cpf1 system enhances genome-editing efficiency and its speed. Cpf1 is also known as Cas12a which is
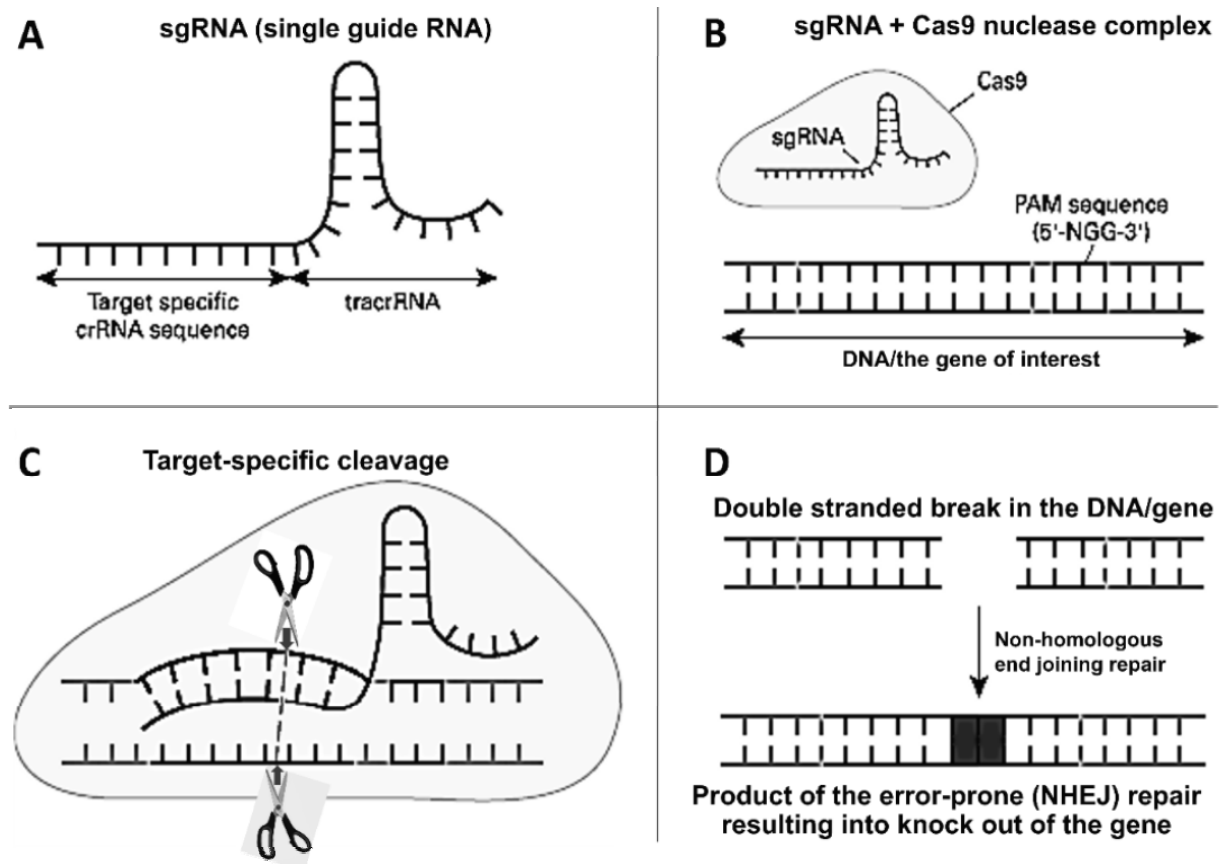
Figure 1. CRISPR/Cas9-mediated gene editing. (**A**) Two components (i) CRISPR RNAs (crRNA), and (ii) trans-activating RNAs (tracrRNA) of a chimeric single-guide RNA (sgRNA), (**B**) The sgRNA-Cas9 riboprotein complex bound at the target DNA with the PAM sequence, (**C**) Cleavage of target DNA by double-strand break at the target DNA, (**D**) Repairing of the breaks using an error-prone non-homologous end joining (NHEJ) mechanism leading to InDel, causing mutation/inactivation of the gene.

more effective and precise than the Cas9. Another newly discovered Cas14 nuclease has smaller size (half the size of Cas9) and targets single-stranded DNA (ssDNA); therefore, it can be utilized for detecting ssDNA viruses of clinical, ecological, and economic importance. RNA editing for programmable A to I replacement (REPAIR), a transcriptome-editing technology (targeting and altering RNA bases), offers an opportunity to edit mRNA [8]. PspCas13b nuclease is being used in REPAIR for RNA knockdown and editing with broad applicability for basic and applied research. Transcript editing might allow answering some of the basic questions about alternative splicing, and translation. This might confer temporary and reversible genetic edits, rather than the permanent genome edits in case of the current editing at gene level. This might also minimize the ethical issues that arise due to the presently practiced genome editing. However, the RNA base editing would require to be administered repeatedly to be functional.

Further advancements in gene editing technology are being made with respect to base editing. Base editing is a CRISPR/Cas-based approach which can convert one nucleotide base to another with precision in DNA or RNA without causing DSB. In fact, all the four possible transition mutations, $C \rightarrow T$, $G \rightarrow A$, $A \rightarrow G$, and $T \rightarrow C$, can be incorporated at the desired site in the gene/genome with the help of CRISPR/Cas base editors (BEs). While the cytosine base editor (CBE) can mount C–G to T–A conversion, the adenine base editor (ABE) can mount A–T to G–C alteration.

In case of RNA, editing of Adenine to Inosine is also possible with the help of RNA base editor (RBE). Herein, a catalytically dead Cas9 (dCas9) developed by mutation in both the domains (RuvC and HNH) of Cas9 is used which lacks DNA-cleavage activity but retains RNA-directed DNA-binding activity. While CBE can be used to edit/incorporate stop codon in the gene, thus causing premature termination of translation, ABE can be used to correct/remove stop codon from a defective allele containing an early stop-codon.

Prime editing is a newer approach of genome editing technique that uses an engineered prime editing guide RNA (pegRNA) which specifies the target site and contains the desired edit using a prime editor protein. The primer editor protein consists of a Cas9 nickase attached with a reverse transcriptase. The Cas9 nickase is guided to the target DNA by the pegRNA. Once the nick is formed by Cas9, reverse transcriptase uses pegRNA as a template to accomplish the desired edit by directly polymerizing DNA onto the nicked target DNA strand. The original DNA strand is replaced by the edited DNA strand, creating a heteroduplex which contains one edited and one unedited strand (Figure 2). Subsequently, the editor guides in copying the edited strand onto the unedited strand and completing the process. Thus, gene/genome editing technology has a very promising future in the areas of research and therapeutics.

3. Epigenomics: a new dimension in functional genomics

The genetic information needed for synthesis and assembly of RNAs and proteins in a living organism is encoded by the cellular genome. But, instructions regarding access to this information in a temporal and spatial manner are encrypted in the epigenome, which grants selective access to the information in the gene. Since plants are sessile in nature, they face several environmental stresses throughout their life. Although plants possess innate capability to tolerate adverse climatic conditions, yet they require further improvement in their efficiency to produce more under unfavourable climatic conditions. Until recently, it has been thought that isolation of a gene associated with a trait of interest would be sufficient enough to transfer the trait from an organism to a plant, and to create the expected phenotype. However, evidence suggests that nucleotide sequence of the gene provides only part of the genetic information; the surrounding environment like chromatin conformation contributes much in expression of the trait. Since the epigenetic states of chromatin are variable, transfer of a trait from one organism to another would not only require the transfer of the gene(s) but the appropriate chromatin/epigenetic states must also be made available so that the trait can express successfully under suitable epigenetic environment. Therefore, it is essential to study the epigenetic states of the donor, and to ensure that proper re-establishment of the epigenetic state takes place in the recipient for appropriate expression of the trait [9].

While a sum total of all the genes in an organism is known as genome, epigenome refers to the sum total of all the epigenetic changes in DNA (without any alteration in the underlying nucleotide sequence) and/or in the structural components of the genetic material that affect expression/activity of the gene/genome. Epigenetics is the study of such variations affecting gene expression in the cell/organism [10]. Epigenetic changes include methylation of cytosine resulting in the formation of 5-methylcytosine (5-mC) [11], histone protein modifications, and variation in the biogenesis of small-RNA (sRNA) [12]. Growing evidence indicates the involvement of epigenetic regulation during the developmental processes as well as during biotic and abiotic stresses in plants and animals. Epigenetic changes may revert back to the original state soon after normalization of the conditions (Figure 3). Interestingly, some (~30%) of the epigenetic changes may be carried over the next generation that often results in phenotypic variations [13]. Thus, it is has become evident that epigenetic changes play important roles in acclimatization, stress tolerance, adaptation, and evolutionary processes in living organisms [14]. Therefore, it is important to discover the epigenetic machinery of gene regulation for crop improvement towards the development of climate-smart crop plants to meet the challenges of food and nutritional security for the global population. Since the rates of genetic mutations and phenotypic variations are
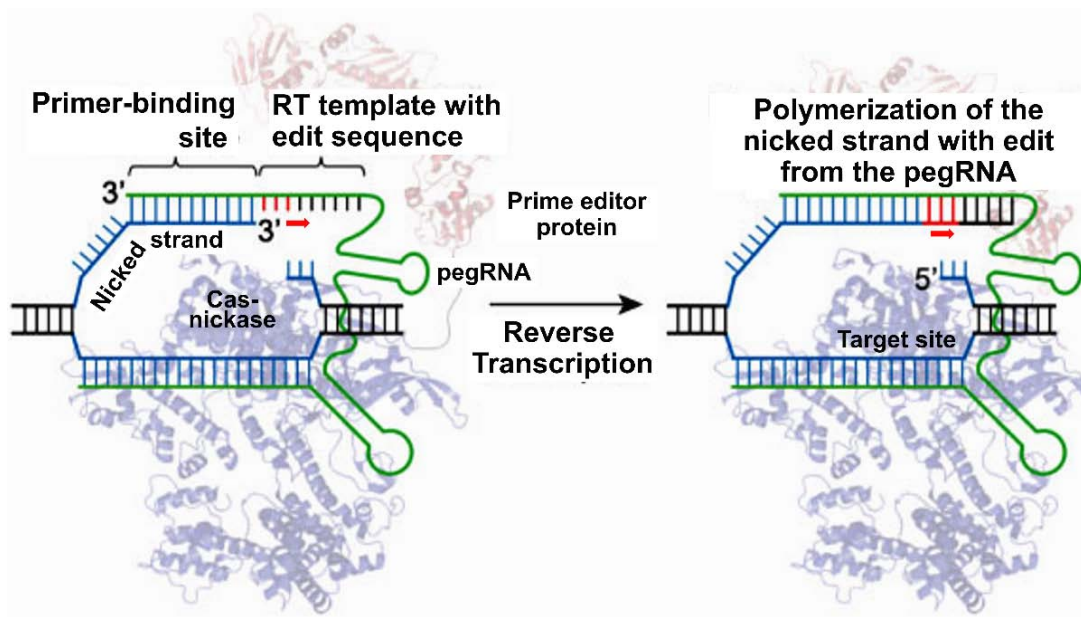
Figure 2. Prime editor with an engineered prime editing guide RNA (pegRNA).
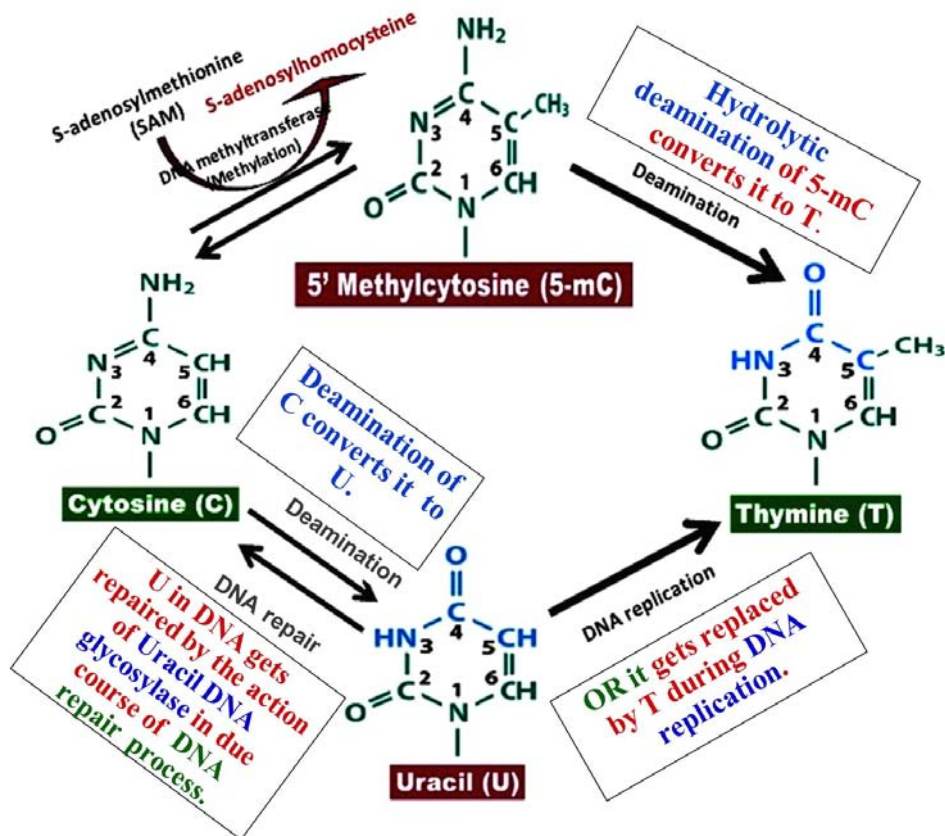

Figure 3. Epigenetic modifications of the genetic material. Conversion of cytosine (C) into 5-methylcytosine (5-mC) and other bases, and its reversal/repair.

considerably different, they cannot be explained merely based on genetics. Additional machinery such as epigenetics can help explaining this enigma [15]. If epigenetics is considered a complementary mechanism, many of the phenotypic variations (e.g. dissimilarity between the clones) can be easily explained.

It has been reported that the rate of spontaneous epimutations is higher in the CG context because these sites are not retargeted by RdDM. DNA methylation generally refers to the addition of a methyl group at the 5th carbon of cytosine as a post-replicative event (Figure 2). While CHH methylation is maintained by Domains Rearranged Methyltransferase 2 (DRM2), it is responsible for *de novo* methylation in all the contexts of cytosine at least in Arabidopsis. DRM2 is recruited to the target loci by a specialized 24 nucleotide small interfering RNA (RNA-directed DNA methylation pathway). Cytosine methylation homeostasis is determined by the DNA methylation and demethylation processes. Demethylation of the promoter and/or coding region may also be required to activate the expression of specific genes under the changing environmental conditions or during the developmental stages of a plant [16]. A variety of histone modifications and their possible combinations (like H3K4me3 & H3K27Ac: activation marks, and H3K9me3 & H3K27me3: repressive marks) affect the transcriptional potential of the gene. Histone methylation can also be reversed by the action of different types of histone demethylases. Studies also indicate that the genome-wide hypo/hyper-methylation induces biogenesis of 24-nt siRNAs, and activates *de novo* (de)methylation pathways. Recent studies reveal a highly cell type-specific nature of epigenetic regulation of genes which indicates the need for new technologies to study the functions of chromatin regulators in a cell-specific manner, at the specific developmental stage, and in proper genomic context. Hence, in-depth studies would be necessary to understand the role of the RdDM pathway and the chromatin regulators in epigenetic regulation of gene expression and its deployment in epigenetic engineering of crop plants. However, epigenetic mechanisms of gene regulation are yet to be fully understood and utilized as epialleles (the alleles that are genetically identical but epigenetically

different due to the epigenetic modifications, showing variable expression) in crop improvement programs.

The conventional approaches may not be adequate to meet the projected food requirements, both in terms of quantity and quality. Moreover, most of the cultivated crops/varieties have reached the yield plateau. Therefore, the need of the day is to deploy modern tools and techniques to further enhance the productivity of crop plants, the nutritional quality of the product, and to explore the possibility of producing novel molecules (molecular farming) with the decreasing availability of natural resources. In view of the biosafety concerns of genetically modified organisms currently being associated with the genetic manipulation of crop plants [3, 17], epigenetic engineering (supposed to have limited biosafety issues) would be a better approach [18]. However, appropriate safety guidelines framed by the regulatory agencies of the country must be followed for personnel, laboratory and environmental safety [5, 19]. Thus, epigenome editing may provide unprecedented opportunities for the manipulation of biological systems in an efficient/effective manner to improve stress tolerance against climatic conditions.

4. Epigenome editing: possibilities and future perspectives

Epigenome editing is a very promising approach that can usher a new era for novel applications of basic research and molecular medicine. Epigenetic editing is based on fusion proteins comprising a designed DNA recognition domain that targets an attached enzymatic domain to defined genomic target sites. Because the target recognition of CRISPR/Cas9 complex is based on Watson/Crick base-pairing between a guide RNA and one DNA strand of the target site, re-targeting of CRISPR/ Cas9 only requires the introduction of a new guide RNA sequence. Histone acetylation and deacetylation play an important role in regulation of gene expression. dCas9 fused with HDAC3 protein can function as a synthetic histone-deacetylase to modulate gene expression. Different groups of scientists are currently working world over to identify the gene(s) involved in epigenetic changes to establish proof of the concept of epigenetic manipulation in plant. However, many

areas of epigenetics remain to be explored. We still know only a little about the factors that regulate targeting of active DNA demethylation during developmental stages. Does DNA (de)methylation interplay with other epigenetic features or chromatin features? Future research should aim at identifying more developmental processes in different species that involve epigenetic regulation. Assessing the contribution of transgenerational epimarks to heritable phenotypic variation has been a major challenge as many of the chromatin (DNA methylation and histone modification) changes and gene expression variants co-segregate with DNA sequence polymorphisms. Nonetheless, there is evidence that plants possess heritable epiallelic variations that can be associated with the trait of interest and utilized for crop improvement. We are still at the beginning of understanding the transgenerational stability of epigenetic variations. Only a little is known to us about the role of the environment in the creation of induced epialleles. We can anticipate that soon epigenome editing will provide a means to assess the role of a QTL in epiallelic variations which may provide an interesting new route for the improvement of crop plants [20]. The proteins involved in DNA (de)methylation, histone modification and the mechanisms of ncRNA-mediated regulation of developmental processes in plants are becoming clear day by day.

CONCLUSION

The discovery of dCas9 has provided a valuable tool for epigenome editing [21]. The recent studies have been focused on regulatory DNA sequences through the recruitment of dCas9 fused to the histone acetyltransferase, and Tet1 DNA demethylase to activate enhancers [22]. Braun *et al.* [7] used FIRE–Cas9 for reversible recruitment of endogenous chromatin complexes to any genomic locus in almost any cell type of mammalian system. The enzymes responsible for writing, erasing, and reading epigenetic marks are multi-protein complexes and they are being identified/characterized day by day. By fusing a single subunit of a chromatin complex with a chemical-induced proximity tag, Frb (FKBP-rapamycin-binding domain of mTOR), Braun *et al.* [7] could rapidly recruit intact multi-subunit complexes to a specific genomic sequence upon rapamycin (RAP) treatment. Locus specificity was obtained through the expression of a complementary dimerizer Fkbp (FK506-binding-protein) fused with a dCas9–MS2 anchor. Focusing on the recruitment of Hp1/Suv39h1 heterochromatin complex and the BAF chromatin-remodeling complex, they could demonstrate possibilities of both gene repression and activation through epigenome editing. This provides new insight into the fine-tuning of epigenetic mechanisms. Recently, Fukushima *et al.* [23] demonstrated *in vivo* epigenome editing using a new construct, dCas9-olEzh2 (Ezh2 from *Oryzias latipes* fused to dCas9) to manipulate H3K27me3. They showed that dCas9-olEzh2 accumulates H3K27me3 at the targeted loci which induced gene repression in Japanese Killifish (*Oryzias latipes*) embryos. These *in vivo* epigenome editing will be very useful for epigenetic regulation of gene expression and heritability of epigenetic modification at targeted genomic loci.

The views expressed here are those of the author only. These may not necessarily be the views of the Institution/Organization the author is associated with.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

The author declares that there are no conflicts of interest.

REFERENCES

1. Kumar, S. 2013, J. Biofertil. Biopestic., 4, e114.
2. Kumar, S. and Singh, A. 2014, J. Biofertil. Biopestic., 5, e121.
3. Kumar, S. 2015, J. Biofertil. Biopestic., 6, e129. doi:10.4172/jbfbp.1000e129.
4. Kumar, S. 2014, Biosafety, 3, e150. doi:10.4172/2167-0331.1000e150.

5. Kumar, S. 2015, Biosafety, 4, e153. doi:10.4172/2167-0331.1000e153.

6. Qi, L. S., Larson, M. H., Gilbert, L. A., Weissman, J. S., Arkin, A. P. and Lim, W. A. 2013, Cell, 152, 1173.

7. Braun, S. M. G., Kirkland, J. G., Chory, E. J., Husmann, D., Calarco, J. P. and Crabtree, G. R. 2017, Nat. Commun., 8, 560.

8. Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J. and Joung, J. 2017, Science, 358, 1019.

9. Kumar, S. 2019, J. Genet. Cell Biol., 3(1), 128.

10. Kumar, S. 2018, Med. Safety Global Health, 7, e145.

11. Kumar, S., Chinnusamy, V. and Mohapatra, T. 2018, Front. Genet., 9, doi:10.3389/fgene. 2018.00640.

12. Wang, X., Li, Q., Yuan, Q., Kumar, S., Li, Y. and Qian, W. 2016, Sci. Rep., 6, e26443. doi:10.1038/srep26443.

13. Kumar, S. 2018, J. Phytochem. Biochem., 2, e102.

14. Kumar, S. 2019, Adv. Biotechnol. Microbiol., 14, e555879. doi:10.19080/AIBM.2019.14. 555879.

15. Kumar, S. 2017, Adv. Plants Agri. Res., 7, e00243.

16. Li, Y., Kumar, S. and Qian, W. 2018, Plant Cell Rep., 37, 77. doi:10.1007/s00299-017-2215-z.

17. Kumar, S., Arul, L., Talwar, D. and Raina, S. K. 2006, Curr. Sci., 91, 930.

18. Kumar, S. 2018, Epigenomes, 2, e6. doi:10.3390/epigenomes2010006.

19. Kumar, S. 2012, Biosafety, 1, e116. doi:10.4172/2167-0331.1000e116.

20. Kumar, S., Singh, A. K. and Mohapatra, T. 2017, Indian J. Genet. Plant Breeding, 77, 445.

21. Thakore, P. I., Black, J. B., Hilton, I. B. and Gersbach, C. A. 2016, Nat. Methods, 13, 127.

22. Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czaudema, X., Shu, J., Dadon, D., Young, R. A. and Jaenisch, R. 2016, Cell, 167, 233.

23. Fukushima, H. S., Takeda, H. and Nakamura, R. 2019, Epigenetics Chromatin, 12, 17.