Plant Epigenetics in stress adaptations

Pronomita Das

Department of Plant Sciences

University of Hyderabad

Telangana, India

pronomita22@gmail.com

Abstract

Chromatin is likely to undergo modifications through epigenetic regulation mechanisms including DNA methylations, structural modulation of histones and RNA-based control mechanisms. Chromatin structure modulation leads to the alteration of gene expression by the interaction of regulatory proteins like transcription factors with DNA. Nowadays, high throughput sequencing technologies have evolved for a better investigation of epigenomic changes in the genome, which would like to help in manipulating these pathways using biochemical and genetic approaches. These methods work as a promising research tool for investigating the epigenetic variations especially cytosine methylation on the genomic regions in a cell or active cistrons. The correlations between genotype, phenotype and epigenotype are extensively studied using epiGWAS in order to accelerate breeding programs for crop improvement.

Keywords: Chromatin, Epigenetics, DNA methylation, high throughput sequencing

I. Introduction

The epigenome consists of the chemical alterations of DNA and histone proteins which constitutively alter gene expressions, plant growth, development, transposable element suppression and tissue differentiations. Waddington coined the term “epigenetics” and he explained the term as “Gene interactions that affect the phenotypic changes” in 1940s. A more specific definition of epigenetics was later developed as “The study of the heritable alterations in gene expression caused during mitosis/meiosis that does not change the DNA sequence” (Yadav et al. 2018). Epigenetic regulation of gene expression can be of three different types such as DNA methylation, chromatin modifications via histone architecture modulations, and RNA-based control mechanisms. Epigenomics has turned out to be one of the most crucial research topics in plant functional genomics. Environmental stresses influence plant growth and development because of their sessile nature. Plants have developed several stress-signaling strategies to cope with stress conditions. Among which epigenetic stress sensing plays a pivotal role in environmental stress adaptation. Recently, various researchers have had investigated to unravel new hidden information to catapult some finer intricacies and active constituents underpinning plant epigenetics controlling stress adaptations in different plant communities.

II. Epigenetics and epigenomics

Epigenetic alterations of chromatin structures such as DNA methylation, chromatin modifications, and RNA-based control mechanisms help plants adapt to environmental changes (Handy et al. 2011).

A. DNA Methylation

DNA methylation is the formation of 5-methylcytosine by an addition of a methyl (-CH_3) group to the fifth carbon of cytosine. DNA methylation such as adenine methylation is generally occurring in prokaryotes and cytosine methylation is mainly found in both plants and animals (mainly eukaryotes)

(Moore et al. 2012). Bacterial DNA methylation assists bacteria to differentiate between host genomic DNA (self) and phage DNA (non-self) which helps to the cleavage of phage DNA using host's restriction enzymes. DNA methylations are highly conserved mechanisms in plants, animals and fungi. In plants, DNA methyl-transferases catalyses DNA methylations into three different sequences CpNpG, CpG, and CpNpN (where N = A, C or T). Because of the symmetrical nature of CpNpG and CpG methylation, it can possibly be retained after replication (Yadav et al. 2018).

Enzymes involved in DNA methylation

These epigenetic modifications of bases are reversible and enzyme mediated. DNA methylation is directed by two types of enzymes, DNA methyl-transferase and DNA demethylation enzymes. Three definite classes of enzymes regulate cytosine methylation in plants namely DNA Methyl-transferase 1 or MET1, Chromomethylase3, and the other class includes two DNA methyl-transferases, DRM1 (Domain rearranged methylase 1) and DRM2 (Zhong et al. 2014).

DNA methyl-transferase 1 (MET1) maintains symmetric cytosine methylation (CpG) of genome. MET1 is a homologue of mammalian methyl-transferase. On the other hand, chromomethylase3 helps in recruiting a methyl group at CHG sequence in both centromeric repeats and transposons. The third class including two DNA methyltransferase catalyses asymmetrical cytosine methylation at CpNpNp site and de novo methylation (He et al. 2011).

Regulation of DNA methylation

Various developmental, physiological and stress influences DNA methylation in plants. Histone and DNA methylations are correlative mechanisms. This methylation mechanism establishes a condensed state of chromatin called heterochromatin in the downstream of CpG (most likely H3K9 methylation) (Moore et al. 2013).

Demethylation takes place in both active and passive ways. Active demethylation occurs by glycosylase activity which removes the methyl-cytosines in DNA and passive demethylation is a result of de novo methylation inhibition or maintenance inability of the parental imprint after DNA replication. Active demethylation significantly plays a role in inhibiting the production of stable hypermethylated epialleles in plant genome (Yadav et al. 2018) (Figure 1).

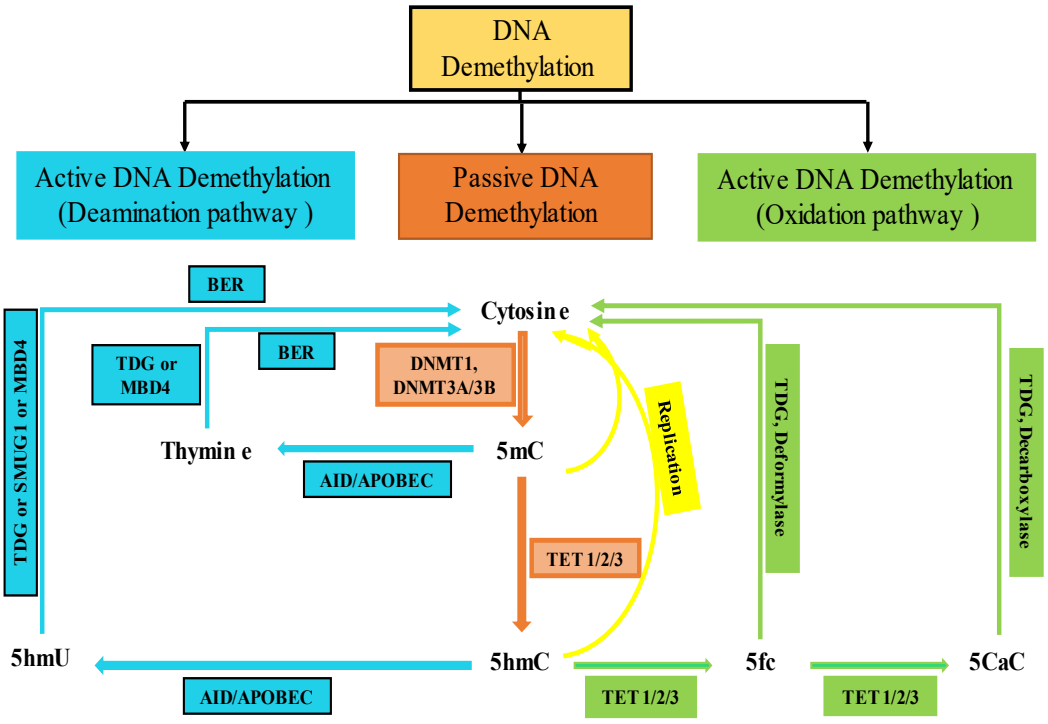

Figure 1: Graphical representation of active and passive demethylation. Passive DNA demethylation is caused by a reduction in activity or absence of DNMTs (orange arrows). Active demethylation via oxidation pathway (green arrows): TET enzymes can hydroxylate methylcytosine (5mC) to form 5-hydroxymethylcytosine (5hmC) ; further oxidation produces 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC can be actively removed by the DNA glycosylases. In addition, a putative deformylase may convert 5fC to C and decarboxylase convert 5caC to C. Active demethylation via deamination pathway (blue arrows): AID/APOBEC family members can deaminate 5mC or 5hmC to form thymidine or 5-hydroxymethyluracil (5hmU). These intermediates are replaced by cytosine during base excision repair (BER) mediated by the uracil-DNA glycosylase (UDG) family, like TDG or SMUG1 as well as MBD4 (specifically recognize thymine and 5hmU). AID activation-induced deaminase .

APOBEC apolipoprotein B mRNA-editing enzyme complex, BER— base excision repair, DNMT1/3A/3B—DNA methyltransferase, MBD4—methyl-binding domain protein 4, SMUG1—single-strand specific monofunctional uracil-DNA glycosylase, TET1/2/3—ten- eleven methylcytosine dioxygenase family, TDG—thymine -DNA glycosylase .

B. Chromatin modifications

Chromatin modifications are covalent post translational modifications of the N-terminal tail of histone protein core in amino acid residues like arginine, lysine, threonine and serine. These covalent modifications are of mainly two types like acetylation and methylation which are regulated by histone modifying enzymes. Acetylation is regulated by acetyl-transferase and de-acetylase, on the other hand, methylation is regulated by methyl transferase and demethylase. Histone modifications regulate transcriptional activity of a gene by altering the chromatin state (figure 2) from loose euchromatin to condense heterochromatin structure or vice-versa (Smith et al. 2009).

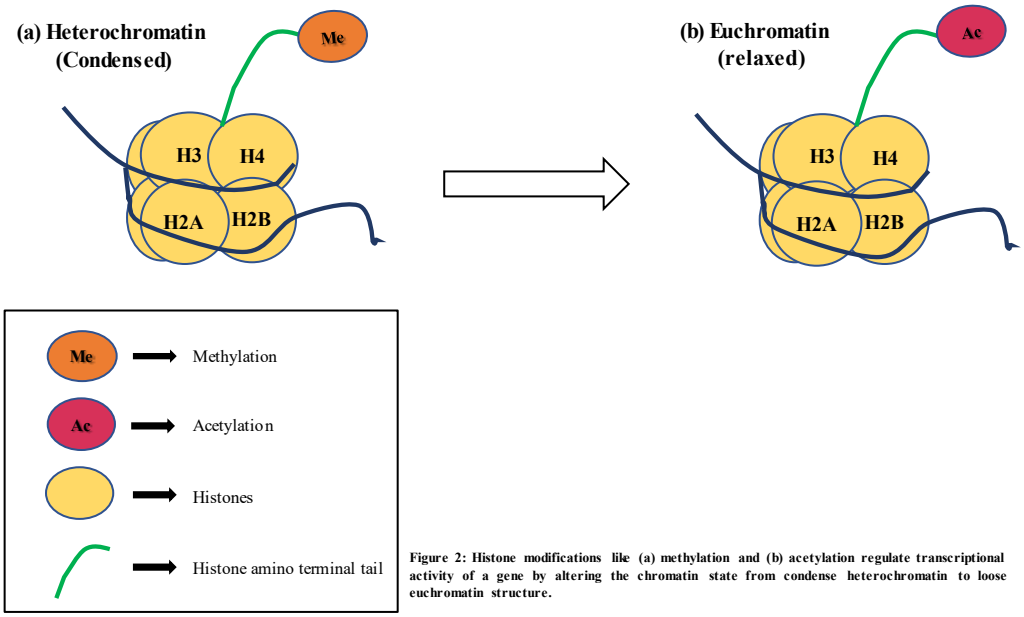

Figure 2: Histone modifications like (a) methylation and (b) acetylation regulate transcriptional activity of a gene by altering the chromatin state from condense heterochromatin to loose euchromatin structure.

Regulation of chromatin modifications

Histone acetyl-transferase (HATs) positively regulates transcription by acetylating 4, 9, 27, 36, and 73 H3 and H4 lysine (K) positions. In contrast, on the contrary, histone deacetylase (HDACs) negatively control transcription by removing acetyl groups from histone backbones. Histone methylation affects transcription depending on the degree and position of methylation. Lysine and arginine methyl-transferases are two methyl-transferase types that direct lysine (K) and arginine (R) methylation, respectively. H3K48me, H3K79me, H3K36me, and H3K4me3 activate the transcription, and in contrast, transcription repression is directed by H3K9, H3K27, H4K20, and H4R3me2. Another type of histone modification constitutively activates the transcription, like H3 phosphorylation at serine and threonine histone residues (Kumar et al. 2021).

C. RNA- based control mechanisms

Wasseneger first discovered RNA based control mechanisms in plants which were then known as RNA-directed DNA methylation (RdDM). RNA based control mechanisms directs de novo DNA methylation at CHG, CG, and CHH sites in a sequence specific manner (Erdmann et al. 2020).

Mechanism and regulation of RNA -based control mechanisms

The first step of siRNA synthesis starts from the double-stranded RNAs (dsRNAs) generation. Transcribed inverted repeats, transposable elements, and intermediates of viral replication likely to act as a source of dsRNAs. RNA polymerase IV (RNA POL IV) initiates RNA-directed DNA methylation (RdDM) by generating single-stranded RNA (ssRNA). The next step is catalysed by RNA-dependent RNA polymerase 2 (RDR2) with the help of chromatin remodeler CLASSY 1 (CLSY1, remodelling

factor). RDR2 generates dsRNA from template ssRNA. DICER LIKE 3 (DCL3 proteins) cleaves dsRNA to produce 24-nt siRNA (small interfering RNA) along with 3` overhangs, which is further methylated by HUA – ENHANCER 1 (HEN1). ARGONAUTE 4 (AGO 4) utilizes a single-stranded methylated siRNA and forms an RNA-induced silencing complex (RISC) – AGO4 complex. Then the RISC-AGO4 complex further directs methylation in homologous loci. Concomitantly, at the target site, DDR complex (DRD1 (DEFECTIVE IN RNA DIRECTED DNA METHYLATION 1), DMS3 (DEFECTIVE IN MERISTEM SILENCING 3), RDM1 (REQUIRED FOR DNA METHYLATION 1), and DMS4 helps PolV to transcribe long noncoding RNA (lncRNA). It was reported that DDR unwinds DNA for transcription. The association of (RISC)-AGO4 complex and PolV are formed by base-pairing between lncRNA and siRNA. Subunits of PolV-Nuclear RNA Polymerase E1 (NRPE1), NRPE2, and KTF1 (Kow domain-containing transcription factor) stabilize this association by interacting with AGO4. Cytosine methylation occurs at the target site with the help of RDM1 binding with AGO4 and DRM2 (de novo methyltransferase). (Erdmann et al. 2020) (Figure 3).

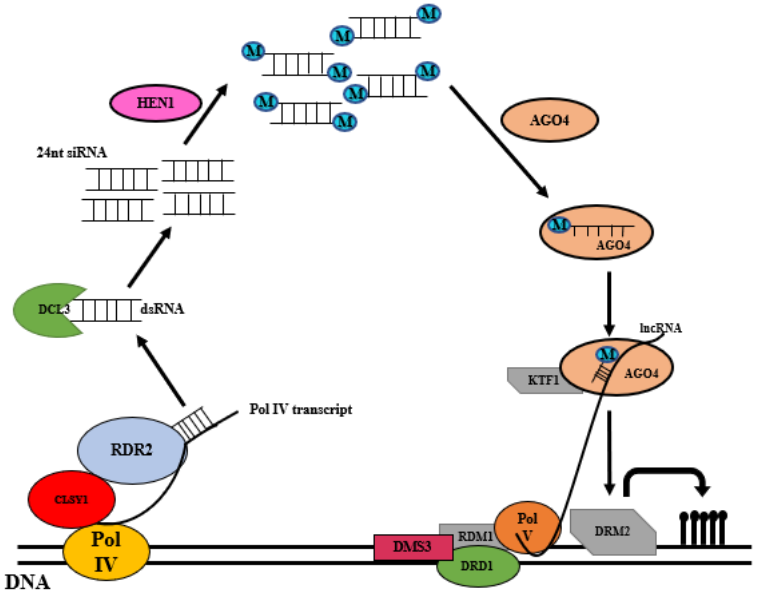

Figure 3: Pictorial representation of the mechanism of RNA – directed DNA methylation (RdDM) – Firstly, RDR2 interacts with POL IV and converts POL IV transcripts to double stranded RNA (dsRNA) with the help of chromatin remodeler CLSY 1. The dsRNAs further forms 24-nt siRNAs by DCL3 action, and the guide strand is loaded onto AGO4, which then enters the POL V-mediated pathway of de novo DNA methylation.

III. Epigenetic stress memory

Researchers have put forward the transgenerational inheritance of DNA methylation patterns in plants which have been implied to help in environmental stress tolerance. Epigenetic regulations help in long term stress adaptation and immediate counter-response of an organism towards stress alleviations. Preservation of altered DNA methylation in selfed progenies is reported in contrasting cultivars of rice (*Oryza sativa*) under alkaline and salt stress condition. This phenomenon has also been seen in short term stress adaptation. As example, Global hypermethylation could be seen in stress treated plant progenies in the absence of stress whereas these epigenetic effects had not persevered in further successive generations. In *Arabidopsis*, a stress-treatment experiment has reported the quantification data of cytosine DNA methylation (Boyko et al. 2010) for two generations. Both treated and untreated *Arabidopsis* progeny plants showed the presence of higher 5mC levels. In unaffected progeny plants of the same generation, during absence of stress, DNA methylation decreases.

IV. Genome wide epigenetic analysis by sequencing technology

Genome wide analytic techniques to detect the patterns of DNA methylation and chromatin modifications in the genome can be divided into four distinct categories such as bisulphite conversion, methylation sensitive restriction enzyme digestions, chromatin immunoprecipitation mediated high throughput sequencing (ChIP- sequencing) and small RNA-mediated methylation. These categories are classified on the basis of the principle of differentiating methylation and demethylation of DNA. Generally, promoter regions in the genome are CpG rich region which are being methylated. Various environmental factors are found to influence a hypermethylated state in the promoter regions having CpG islands to inactivate the concern genes. Therefore, quantification of differentially methylated regions (DMRs) in the plant genome during several stress response becomes handy.

A. Bisulfite sequencing

Bisulfite sequencing method includes sodium Bisulfite treatment onto genomic DNA which leads to the transformation of unmethylated cytosine into uracil, on the other hand this treatment does not affect methylated cytosine residues. Then PCR amplification of converted DNA is done by using sequence-specific primers which helps in understanding the methylation status of DNA. So, Bisulfite treatment can provide promising information about detection of specific modifications in the genomic DNA sequence. This information depends on the methylation conditions thus providing the single nucleotide resolution information on the methylation status of a DNA sequence. Thus, several bioinformatic analyses ensure the recovery of the information for nucleotide resolution.

The whole genome bisulfite sequencing can also help in understanding the methylation level in UTRs (untranslated regions), promoters and other protein-coding regions of the active cistron. It can also be used in the quantification of small RNAs and transcriptome sequencing and their associations with the abundance of DNA methylation with small RNAs (Shaoke et al. 2014). Strand specific mRNA sequencing provides information about altered transcript amount of distinct genomic regions such as intergenic regions, transposons and gene changes in the abundant transcripts of transposons, hundreds of genes and unannotated intergenic transcripts upon changing their DNA methylation state. Briefly all these data sets help in understanding the complex epigenetic relationship between DNA methylation and transcription. At present a large amount of research focus has been concentrated on whole genome bisulfite sequencing in generating methylomes maps ranges from *Arabidopsis* to *Zea mays*.

Reduced-representation bisulfite sequencing (RRBS) approach is an alternative, effective and cost reduced bisulfite sequencing developed by Meissner et al. (2008) which helps in investigating the regulation of mammalian methylome. The RRBS method includes Msp1 restriction enzyme digestion of genome which later undergoes bisulfite conversion and subsequent next generation sequencing to conclude about methylation patterns of specific fragments. Along with web-based tools, several bioinformatic analyses are also available which provides evidence in the quantitative analysis of bisulfite sequencing data obtained from methylation status in plants. For example, Akalin et al. (2012) developed methylKit which is a user-friendly tool having multi-threaded R package with quick analysing and data characterization power from a set of methylation experiments. Similarly, Jiang et al. (2014) developed another methylome analysis pipeline named as Methy-Pipe which is an efficient software package for methylation data analysis along with downstream analysis. Krueger and Andrews (2011) developed a flexible tool (Bismark) for the analysis of bisulfite sequencing data.

B. Methylation sensitive restriction enzyme digestion

Methylation-sensitive amplified polymorphism (MSAP) is a quantification technique of DNA methylation without knowing any DNA sequence information. This method helps to recognize the cytosine methylation pattern in the genomes. Methylation-sensitive two isoschizomers such as Hpa II and Msp I are involved in this identification method. Their function differs in their methylation level sensitivity to some recognition sequences (5`-CCGG-3`). Methylated strands (both) of double stranded cytosine methylations disables HpaII restriction but enables the restriction activity when external cytosine is hemimethylated (Fulneček J et al. 2014). On the contrary, MspI actively cuts fully or hemi-methylated C5mCGG but not 5mCCGG. The locus specific differentiation between unmethylated and methylated DNA sequences can be analyzed based on restriction sites. Recently, investigation of

methylation/demethylation status of plants can be done using MSAP technique. Various research groups utilize this quantification analysis of differentially methylated regions in agronomic traits by comparing the methylation pattern in contrasting cultivars.

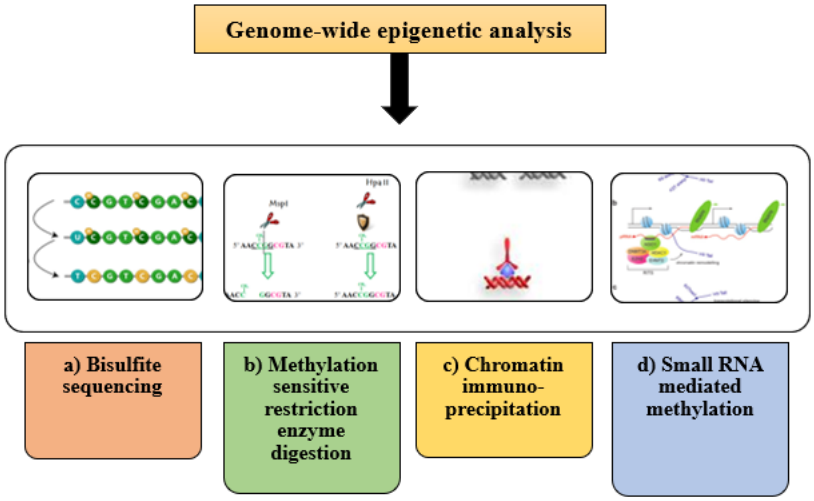

Figure 4: Genome-wide epigenetic analysis by sequencing technology – a) Bisulfite sequencing, b) Methylation sensitive restriction enzyme, c) Chromatin immuno-precipitation, d) Small RNA mediated methylation

C. Chromatin immunoprecipitation

Epigenomic dynamic alterations of chromatin are diverse for developmental stages, disease states and distinct tissue types for different environmental stress responses. Genome wide high throughput technologies are used to analyse these chromatin states or epigenetic phenomena (Kidder et al. 2011). Recently the most popular technique to study epigenomics is Chromatin immunoprecipitation (ChIP assays or ChIP sequencing). ChIP seq discovers genome wide modifying state of chromatin complex along with transcription factors and other proteins. The approach analyses the regulation of histone or other DNA-protein interactions in different cell types, developmental stages and environment effects. The ChIP seq method includes crosslinking, isolation and chromatin fragmentation along with protein-DNA complex capturing by antibodies against the transcription factors or histone proteins. The immunoprecipitated protein-DNA complexes are reversely cross-linked. Further analysis can be done using the purified DNA by hybridization on to microarrays, i.e., high throughput sequencing ChIP-seq or ChIP-chip assay (Mundade et al. 2014).

This technique involves the crosslinking of formaldehyde with the histones and DNA -protein nucleosome complexes, subsequently with extraction and fragmentation of chromatin. Then finally those fragmented samples are accredited for chromatin immunoprecipitation (ChIP) with antibodies which are modification specific. Then PCR amplification can be done to obtain proper amount of DNA which is then undergoes denaturation to produce the single stranded DNA (ssDNA). These ssDNAs are fluorescent-labelled for sample differentiation. Finally, fluorescent-labelled fragments are subjected to hybridization with the target single stranded sequences on to the DNA microarray surface representing the genomic region of interest (Milne et al. 2009). The target sequences will hybridize with complementary labelled fragments on the chip array to form dsDNA followed by fluorescence illumination. Finally, the fluorescent signals are generalized with administrate signals along with statistical tools to check out the methylated regions. Then to find the physical positions, reference genome can be used for mapping the coordinates of microarray probes.

A strong strategy is the use of next generation sequencing of ChIP fragments for producing the high-resolution, high genomic coverage and low noise compared with ChIP-on-chip assays. The resolution of

ChIP-on-chip depends on both compactness and size of chromatin fragments which are used for ChIP and probes on the array. In contrast, the resolution of ChIP-seq depends on the production of same fragment size and the details of reads while sequencing (Mahony et al. 2015). ChIP-seq is relatively cost-effective in achieving plant genome derived nucleosome complexes resolution than ChIP-on-chip assay.

D. Small RNA mediated methylation

Small interfering RNA (SiRNA) can target chromatin to similar sequence genomic regions. The most acceptable method of sRNA based epigenomic investigation is dependent on the size choosing of total RNA population from diverse genotypes, tissue types, mutants and subspecies. Size fractionation is done after total RNA isolation from required samples. Subsequently, DNA adaptors are ligated at both the sRNA ends, which function as primer binding sites during PCR amplification and reverse transcription. Then, sequencing can be done using NGS approach followed by genome wide large-scale reads (Yadav et al. 2018).

V. Conclusions and future perspectives

Nowadays, epigenetic inheritance in plants is most sought-after area of research. Epigenetics indeed has a possibility of facilitating novel and better approaches to environmental stress tolerance and crop improvement. Advance technologies are now growing up for the fast and efficient investigation of both epigenotype and genotype. These contribute in garnering valuable clues and leads for investigating the character of epigenetics in influencing essential phenotypes and the counterproductive environmental signal responses. Epigenetic regulations such as DNA methylation, chromatin modifications, and gene silencing by small RNAs alter gene activity which is a hallmark molecular marker for the environmental stress responses related to development, cell-fate determination, and cellular proliferation. Using this pillar of evidence, mutation of alleles of genes can be used to the identification of DNA methylation and sRNA pathways for the development of an aberrant phenotype. This identification on the other hand helps the researchers to focus on suitable genes, genomic regions and transcription factors to develop desired phenotypes for further crop improvement programs.

References:

1. Akalin, A., Kormaksson, M., Li, S. *et al.* (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* **13,** R87. https://doi.org/10.1186/gb-2012-13-10-r87
2. Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollunder J, Meins F Jr, Kovalchuk I. (2010). Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. PLoS One;5(3):e9514. doi: 10.1371/journal.pone.0009514.
3. Erdmann RM, Picard CL. (2020). RNA-directed DNA Methylation. PLoS Genet. 2020 Oct 8;16(10): e1009034. doi: 10.1371/journal.pgen.1009034.
4. Fulneček J, Kovařík A. (2014). How to interpret methylation sensitive amplified polymorphism (MSAP) profiles? BMC Genet.; 15:2. doi: 10.1186/1471-2156-15-2. PMID: 24393618; PMCID: PMC3890580.
5. Handy DE, Castro R, Loscalzo J. (2011). Epigenetic modifications: basic mechanisms and role in cardiovascular disease. Circulation. 17;123(19):2145-56. doi: 10.1161/CIRCULATIONAHA.110.956839
6. He XJ, Chen T, Zhu JK. (2011). Regulation and function of DNA methylation in plants and animals. Cell Res. 21(3):442-65. doi: 10.1038/cr.2011.23.
7. Jiang P, Sun K, Lun FMF, Guo AM, Wang H, Chan KCA, et al. (2014) Methy-Pipe: An Integrated Bioinformatics Pipeline for Whole Genome Bisulfite Sequencing Data Analysis. PLoS ONE 9(6): e100360. https://doi.org/10.1371/journal.pone.0100360
8. Kidder BL, Hu G, Zhao K. (2011). ChIP-Seq: technical considerations for obtaining high-quality data. Nat Immunol. ;12(10):918-22. doi: 10.1038/ni.2117.

9. Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics (Oxford, England)*, *27*(11), 1571–1572. https://doi.org/10.1093/bioinformatics/btr167
10. Kumar, V., Thakur, J.K. & Prasad M. (2021). Histone acetylation dynamics regulating plant development and stress responses. Cell. Mol. Life Sci. 78, 4467–4486. DOI: https://doi.org/10.1007/s00018-021-03794-x
11. Lou, S., Lee, H. M., Qin, H., Li, J. W., Gao, Z., Liu, X., Chan, L. L., Kl Lam, V., So, W. Y., Wang, Y., Lok, S., Wang, J., Ma, R. C., Tsui, S. K., Chan, J. C., Chan, T. F., & Yip, K. Y. (2014). Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. Genome biology, 15(7), 408. https://doi.org/10.1186/s13059-014-0408-0
12. Mahony S, Pugh BF. (2015). Protein-DNA binding in high-resolution. Crit Rev Biochem Mol Biol. ;50(4):269-83. doi: 10.3109/10409238.2015.1051505.
13. Milne TA, Zhao K, Hess JL. (2009). Chromatin immunoprecipitation (ChIP) for analysis of histone modifications and chromatin-associated proteins. Methods Mol Biol; 538:409-23. doi: 10.1007/978-1-59745-418-6_21.
14. Moore LD, Le T, Fan G. (2013). DNA methylation and its basic function. Neuropsychopharmacology. 38(1):23-38. doi: 10.1038/npp.2012.112.
15. Mundade R, Ozer HG, Wei H, Prabhu L, Lu T. (2014). Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. Cell Cycle;13(18):2847-52. doi: 10.4161/15384101.2014.949201.
16. Smith, B. C., & Denu, J. M. (2009). Chemical mechanisms of histone lysine and arginine modifications. Biochimica et biophysica acta, 1789(1), 45–57. https://doi.org/10.1016/j.bbagrm.2008.06.005
17. Yadav, C. B., Pandey, G., Muthamilarasan, M., & Prasad, M. (2018). Epigenetics and Epigenomics of Plants. Advances in biochemical engineering/biotechnology, 164, 237–261. https://doi.org/10.1007/10_2017_51
18. Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE. (2014). Molecular mechanism of action of plant DRM de novo DNA methyltransferases. Cell.157(5):1050-60. doi: 10.1016/j.cell.2014.03.056.