NEW ERA AGRICULTURE MAGAZINE

Modulation of Gene Expression by Transposable Elements in Plant Systems

Mukul Kumar[1], K. Rahul Karthik[2], Nikkondi Chandra Prabha[3], Rishabh Singh[4] and Arkadeb Chatterjee[5]

Abstract:

Transposable elements (TEs) make up a significant portion of many plant genomes, and their activity can generate new genetic variation within a species. To ensure normal gene function despite the presence of TEs, plant genomes have developed various adaptive strategies. This review explores how TEs interact with gene expression in plants by examining three primary mechanisms through which they exert influence. Firstly, increasing evidence suggests that TE insertions in introns or untranslated regions (UTRs) are often tolerated, having little effect on gene expression or splicing. However, in some cases, insertions within genes can lead to unusual or entirely new transcript forms. Secondly, TEs can introduce alternative promoters, resulting in altered expression patterns or the production of novel transcripts with different coding capacities. Thirdly, when located near genes, TEs can affect gene regulation in several ways. They may serve as new cis-regulatory elements such as enhancers, or disrupt existing enhancers to modify transcriptional activity. Additionally, TE insertions can influence chromatin structure and associated epigenetic marks around genes, thereby altering gene expression levels. Overall, the relationship between TEs and genes demonstrates that TEs play a more dynamic and integral role in shaping gene function and genome evolution in plants than previously understood, extending beyond their traditional view as merely passive genomic components or insertional disruptors.

Mukul Kumar[1], K. Rahul Karthik[2], Nikkondi Chandra Prabha[3], Rishabh Singh[4] and Arkadeb Chatterjee[5]

[1]Assistant professor-cum-Jr. Scientist, Department of Plant Physiology & Biochemistry, HRC, Kishanganj, BAU, Sabour, Bihar

[2]Research Scholar, Department of Plant Physiology, IARI- New Delhi

[3]M.Sc Scholar, Department of Genetics and Plant Breeding, OUAT, Bhubaneswar

[4]M.Sc Scholar, Department of Agricultural Economics, Kulbhaskar Ashram PG College, Prayagraj, UP

[5]M.Sc. Scholar, Department of Nematology, Assam Agricultural University, Assam

Introduction:

Most eukaryotic genomes are largely made up of non-coding DNA, which includes introns, intergenic regions, and various repetitive sequences. A significant portion of these repetitive sequences originates from high copy transposable elements (TEs). Research on genome functionality and chromatin structure has revealed marked differences between genes and TEs. For instance, recombination events are frequent within genes but relatively rare around TEs. Genes are typically associated with active chromatin marks such as histone acetylation and methylation at lysine 4 on histone H3 (H3K4me), which support active transcription. In contrast, TEs are generally marked by repressive chromatin modifications like CHG DNA methylation, H3K9me2, and H3K27me1, which are linked to transcriptional silencing.

While these generalizations help distinguish genes from TEs, they oversimplify the complex nature of the genome, which is a mosaic of both coding and non-coding elements. There is a finely balanced interaction between these components that ensures the proper functioning of the genome. In this review, we explore the ways in which transposable elements influence and interact with genes in plant genomes.

Intragenic TEs influence genic chromatin and transcript structure

Plant genomes generally tolerate TEs within non-coding portions of genes

Many plant genes have transposable elements (TEs) embedded within their transcribed regions. Interestingly, numerous essential genes carry such TE insertions, indicating that their presence does not necessarily disrupt gene function and can often be tolerated. The integration of TEs within these regions raises important questions. For instance, is the chromatin structure surrounding these intragenic TEs more similar to that of typical genic regions, or does it resemble the chromatin associated with non-genic TEs? Additionally, do these TEs affect gene expression or interfere with proper transcript splicing? While much of the current understanding comes from studies in *Arabidopsis thaliana*, a species known for having relatively few intragenic TE insertions, recent research in other plant species has also started to shed light on the relationships among chromatin structure, TEs, and gene regulation.

Processing of transcripts from genes containing TEs

Transposable elements (TEs) located within introns can interfere with proper transcript processing and present several biological challenges. Studies of mutant alleles in plants with active DNA transposons have revealed cases where TEs inserted into introns create new splice sites, resulting in the

formation of abnormal transcripts. For instance, insertions of the Mutator transposon in introns of the *Adh1-S* and *brown midrib1 (bm1)* genes have led to the generation of novel transcript isoforms. Similarly, a Mutator insertion in an exon of the *zmet2-m1* gene, which encodes a chromomethylase in maize, is partially spliced out, leaving a 40 base pair fragment in the final mRNA. Typically, such splicing between genes and transposons disrupts normal protein production and results in nonfunctional proteins.

There are also documented instances of alternative splicing involving retrotransposons and plant genes. One notable example is the "mantled" trait in oil palm, attributed to somaclonal variation, which has been linked to changes in chromatin at an intronic TE, ultimately affecting splicing. In soybean, unstable alleles that impact seed color have been shown to involve methylation patterns at a TE within the *R* locus intron, influencing proper splicing. Genome-wide studies in maize further suggest that DNA methylation near splice sites may play a role in regulating splicing, though the exact mechanisms remain unclear.

In humans, certain TEs have contributed to the creation of new exons, which often undergo alternative splicing. However, despite these examples, most TE insertions in introns or untranslated regions do not cause splicing anomalies or noticeable phenotypic changes. Many of these intragenic TE insertions likely have minimal impact on gene expression. When TEs do significantly affect gene function, the resulting phenotypic changes are subject to natural selection, potentially leading to a decrease in the frequency of such alleles within populations. Notably, altered splicing due to TE insertions is mostly observed in populations with recently arisen mutant alleles.

TEs influences on regulation of gene expression

Beyond generating new transcripts via alternative splicing, premature transcription termination, or novel promoters, transposable elements (TEs) can also alter the expression levels or spatial/temporal patterns of plant genes. This regulatory influence can occur through several mechanisms, including disruption of existing cis-regulatory elements, modification of chromatin states, or by introducing new regulatory elements.

Although there are known cases where TE insertions near genes are associated with changes in gene expression, pinpointing the exact mechanism behind these effects remains challenging. This difficulty arises largely because the regulatory architecture of most plant genes is not fully understood many cis-regulatory elements and their interacting trans-factors have yet to be identified. As a result,

it's often unclear whether a TE is interfering with normal regulatory sequences or contributing new regulatory input.

Furthermore, TEs typically possess chromatin signatures that differ from those of nearby gene regulatory regions. To assess whether these chromatin differences influence adjacent gene expression, one would ideally compare expression between alleles with and without TE-associated chromatin marks. While such comparisons are possible for some TE families using lines with active or inactive elements, such resources are often unavailable. Another approach involves using mutants that affect chromatin regulation, but severe disruptions in DNA methylation or silencing pathways are often lethal in many plants, limiting their utility.

This section explores the various ways transposons may regulate gene activity and presents examples where specific mechanisms are believed to play a role.

Insertional mutagenesis of gene regulation by transposons

Transposon insertions can interfere with pre-existing regulatory sequences. A substantial part of gene regulation depends on transacting factors interacting with specific cis-acting elements. When transposable elements insert into these cis regulatory regions, they can hinder transcription factor binding, leading to abnormal gene expression. Several well documented cases in maize illustrate how such insertions by Mutator elements disrupt normal regulatory control. For instance, dominant mutations in the *Knotted1*, *Rough Sheath1*, and *Liguleless3* genes each of which encodes homeobox proteins crucial for maintaining meristem identity have been linked to TE insertions within their third intron. Normally, these genes must be repressed during leaf formation to ensure proper leaf development. However, the TE insertions prevent this repression, causing the genes to remain active and resulting in homeotic changes in leaf structure. These findings suggest that key regulatory elements necessary for gene silencing lie within the third intron, and that TE insertions in these regions can block repressor binding, leading to misexpression. Such dominant mutations offer clear examples of how transposons can disturb typical gene regulation by interrupting essential control elements. Transposable element (TE) insertions into enhancer or promoter regions often lead to decreased gene expression, frequently resulting in recessive traits. For example, the absence of pigmentation in some grape cultivars is linked to a retrotransposon insertion in the promoter of a *Myb* transcription factor gene.

In maize, the *tb1* gene illustrates a key allelic change involved in domestication, caused by a TE insertion in a distant regulatory

region. Another example is the *ZmCCT* gene, where a TE insertion in its promoter leads to lower expression levels and altered sensitivity to photoperiod. A significant flowering time QTL in maize, *Vgt1*, was narrowed down to a conserved non-coding sequence located 70 kb from an *Ap2*-like transcription factor. Natural variation at *Vgt1* arises from the insertion of a miniature inverted-repeat transposable element (MITE) in this region. Although the exact mechanism remains uncertain, the MITE may hinder the binding of regulatory proteins, thereby disrupting normal gene activity. Moreover, this insertion is associated with altered DNA methylation, which could further influence gene regulation. These cases demonstrate that TE insertions in regulatory DNA can interfere with gene expression and serve as valuable tools for uncovering mechanisms of gene control in plants.

Chromatin based regulation of gene expression by transposons

Transposable elements (TEs) can affect gene expression by altering the surrounding chromatin environment. Typically, TEs are enriched with heterochromatic modifications, which can interfere with transcriptional activity. If these repressive marks remain restricted to the TE itself, they may limit the function of any enhancer elements within the TE but may not directly affect neighbouring genes. However, when such chromatin modifications extend beyond the TE into adjacent genomic regions, they can impact the expression of nearby genes.

A study by Hollister and Gaut in *Arabidopsis* showed that heavily methylated TEs located in promoter regions are often associated with reduced gene expression. They also observed that highly methylated TEs near genes are subject to purifying selection, suggesting a detrimental effect on nearby gene expression. This implies that the chromatin state of TEs in proximity to genes can play a regulatory role. Alternatively, it's possible that TEs tend to insert near genes that are already expressed at low levels. Studies of natural variation have produced mixed results regarding the direct relationship between TE presence and gene expression. Chromatin profiling around TEs has revealed that heterochromatin can sometimes spread beyond the TE's boundaries, depending on the TE family involved. This heterochromatin spreading could contribute to reduced transcriptional activity of neighbouring genes.

Transposons as a source of novel regulatory information

Transposable elements (TEs) are tightly regulated and depend on specific cis regulatory sequences to maintain this control. Some TEs show transcriptional activation or mobility in response to stress conditions. For instance, the *Tnt1* element in tobacco can be

triggered by both biotic and abiotic stress factors. Likewise, many TEs are known to become active during tissue culture, and the *mPing* DNA transposon in rice is responsive to cold and salt stress. Additionally, there are TEs that display expression patterns tied to developmental stages for example, several TEs in *Arabidopsis* are activated in the vegetative nucleus of pollen cells. In animals, certain transposons contain silencer elements that recruit Polycomb group proteins for transcriptional repression. For TEs to be properly regulated under stress or during development, the necessary cis-regulatory elements must reside within the transposon itself so that the regulatory functionality remains with the element as it moves. Enhancer sequences embedded within the TE must work in conjunction with the TE's own promoter to ensure context specific expression. However, these same regulatory elements may also affect the expression of neighboring genes. This aligns with Barbara McClintock's early hypothesis that TEs may contribute to the genome's regulatory response to environmental stress and influence gene activity.

The *ONSEN* retrotransposon in *Arabidopsis* serves as a prime example of a transposable element with complex regulation that can also impact the expression of nearby genes. Under normal conditions, *ONSEN* remains transcriptionally inactive, but it can be induced by heat stress. This activation depends on the presence of heat shock factor binding sites and is further influenced by specific chromatin states. Notably, *ONSEN* insertions can enhance heat-induced expression of adjacent genes. Genes located near *ONSEN* elements are upregulated in response to heat, whereas natural alleles without these insertions do not exhibit such heat responsiveness. Similar effects have been observed in other plant species. In rice, many *mPing* insertions are found near the 5′ regions of genes. Although these insertions often do not alter gene expression under normal conditions, *mPing* becomes activated during cold stress. When this happens, nearby genes also tend to show increased cold-responsive expression. This pattern of regulation has been consistently noted for rice genes located near *mPing* elements. A genome-wide study in maize further supports this phenomenon, identifying approximately 20 TE families linked to stress-responsive gene expression. In many of these cases, the TEs themselves are also activated by the same stress signals and appear to carry enhancer elements that drive the coordinated stress responsive activation of nearby genes.

References

1. Eichten, S.R., Ellis, N.A., Makarevitch, I., Yeh, C.T., Gent, J.I., Guo, L.,

McGinnis, K.M., Zhang, X., Schnable, P.S., Vaughn, M.W., Dawe, R.K. and Springer, N.M. Spreading of heterochromatin is limited to specific families of maize retrotransposons. *PLoS Genet.*, **8**, e1003127 **(2012)**.

2. Fu, H., Zheng, Z. and Dooner, H.K. Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1082–1087 **(2002)**.
3. Li, Q., Eichten, S.R., Hermanson, P.J., Zaunbrecher, V.M., Song, J., Wendt, J., Rosenbaum, H., Madzima, T.F., Sloan, A.E., Huang, J., Burgess, D.L., Richmond, T.A., McGinnis, K.M., Meeley, R.B., Danilevskaya, O.N., Vaughn, M.W., Kaeppler, S.M., Jeddeloh, J.A. and Springer, N.M. Genetic perturbation of the maize methylome. *Plant Cell*, **26**, 4602–4616 **(2014)**.
4. SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z. and Bennetzen, J.L. Nested retrotransposons in the intergenic regions of the maize genome. *Science*, **274**, 765–768 **(1996)**.
5. Saze, H. and Kakutani, T. Differentiation of epigenetic modifications between transposons and genes. *Curr. Opin. Plant Biol.*, **14**, 81–87 **(2011)**.
6. Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A.H. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.*, **8**, 973–982 **(2007)**.
7. Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L. and Zilberman, D. The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, **153**, 193–205 **(2013)**.