Review

Roles of Transposable Elements in the Different Layers of Gene Expression Regulation

Denise Drongitis [1], Francesco Aniello [2], Laura Fucci [2] and Aldo Donizetti [2,*]

[1] Institute of Genetics and Biophysics "Adriano Buzzati Traverso", Consiglio Nazionale delle Ricerche, 80131 Naples, Italy; denise.drongitis@igb.cnr.it

[2] Department of Biology, University of Naples Federico II, 80126 Naples, Italy; faniello@unina.it (F.A.); fucci@unina.it (L.F.)

* Correspondence: aldo.donizetti@unina.it

check for
updates

Abstract: The biology of transposable elements (TEs) is a fascinating and complex field of investigation. TEs represent a substantial fraction of many eukaryotic genomes and can influence many aspects of DNA function that range from the evolution of genetic information to duplication, stability, and gene expression. Their ability to move inside the genome has been largely recognized as a double-edged sword, as both useful and deleterious effects can result. A fundamental role has been played by the evolution of the molecular processes needed to properly control the expression of TEs. Today, we are far removed from the original reductive vision of TEs as "junk DNA", and are more convinced that TEs represent an essential element in the regulation of gene expression. In this review, we summarize some of the more recent findings, mainly in the animal kingdom, concerning the active roles that TEs play at every level of gene expression regulation, including chromatin modification, splicing, and protein translation.

Keywords: transposable elements; gene expression regulation; TEs co-option

1. Introduction

The mobile genetic elements known as transposable elements (TEs) were discovered in the late 1940s thanks to Barbara McClintock's [1,2] seminal work on maize, a discovery that revolutionized the previous view of genes as stable entities on chromosomes. TEs are an important driving force in genome evolution, and bursts of TEs have been connected with significant evolutionary events such as the speciation process [3]. TEs are DNA fragments that have the ability to mobilize within a host genome, often creating new copies of themselves during the mobilization process.

1.1. Transposable Elements Classification

A detailed classification has been extensively reviewed elsewhere [4]. Briefly, TEs are divided into two large classes called Class I retrotransposons and Class II DNA transposons, which differ from each other in their mobilization mechanisms. Class I retrotransposons are transcribed into an RNA intermediate that can then be reverse transcribed into DNA and reintegrated as an additional copy in the genome. In contrast, Class II DNA transposons encode a transposase enzyme that cuts the parental sequence from the genomic sequence and then mediates its reintegration into another location in the genome [4,5]. Class I retrotransposons are further subdivided into five orders depending on their mechanistic features, organization, and reverse transcriptase phylogeny [4]. Class I includes long terminal repeats (LTR) retrotransposons, such as the endogenous retroviruses (ERVs) superfamily, and long and short interspersed nuclear elements (LINEs and SINEs, respectively) [4,6,7]. Class II DNA transposons can be divided into two subclasses on the basis of the number of DNA strand that

are cut during transposition, and includes the order of terminal inverted repeat (TIR) transposons in the subclass I [4].

1.2. The Deleterious Effects of the Transposable Elements Mobilization

Many elements cooperate in TE integration such as the genomic sequence, chromatin, and nuclear contexts, which accounts for the diversity in insertion-site distribution and evolutionary strategies [8]. The ability of TEs to move through the genome can be detrimental to the host if not properly controlled. The deleterious effects of mobile element activation had long been assumed to be harmful by causing chromosome breakage or acting as a mutagen when targeted to protein-coding genes [9]. However, TE transposition can also occur during germline development and, less frequently, in somatic cells [10,11]. The deleterious effects of germline transpositions have been observed for some time, with different examples coming from different organisms and different transposon classes. For example, LINE1 (L1; Class I) insertions in human haemophilia A and the *P*-element DNA transposon (Class II) in *Drosophila* that is the basis for dysgenic traits [12,13]. Somatic transpositions have also been receiving much attention, as a growing body of evidence is accumulating that associates detrimental biological effects with the somatic transposition of TE. One of the most interesting examples of this is represented by the mariner-Mos1 element of Drosophila (Class II transposon), which can be transposed during the Drosophila life cycle and negatively affects behavioral activities and embryonic viability [14]. Probably one of the most characterized pathological states linked to a new wave of transpositions in somatic cells is cancer, including ovarian [15], colorectal [16], and Fanconi anemia [17] cancer, to name just a few. Transposable elements can escape from their epigenetic silencing, as has been shown in most cancers, where lower methylation levels (hypomethylation) and the dysregulated chromatin modification of L1 retrotransposons results in their integration into novel sites (insertional mutagenesis), a process that has been reported for esophageal squamous cell carcinoma [18] and pancreatic ductal adenocarcinoma [19]. In addition to cancer, it has become clear that TEs are implicated in the rise of different brain disorders, including schizophrenia [20], and autism [21]. In schizophrenia patients, whole-genome sequencing has revealed specific L1 insertions in the brain, probably triggered by environmental and/or genetic risk factors and, preferentially localized to synapse- and schizophrenia-related genes, this has led to the hypothesis that the hyperactivation of L1 retrotransposons may contribute to the susceptibility and pathophysiology of schizophrenia [22]. The aberrant activation and mobilization of TEs have also been associated with the pathophysiology of neurodegenerative diseases such as Alzheimer disease (AD). Guo et al. [23] have shown an interesting association between the differential expression of several TEs and the neurofibrillary tangle burden in post-mortem human brains, indicating an association between TE activation and genomic instability in Tau-mediated AD mechanisms. In addition, the silencing of TEs mobilization through epigenetic mechanisms has been hypothesized as a neuroprotective mechanism of cortical spreading depression (CSD), which is an evolutionarily conserved phenomenon that involves a slow, self-propagating depolarization wave that is associated with the spontaneous depression of electrical neuronal activity, which can promote neuroprotection by inducing preconditioning [24]. The effects of the epigenetic control of LINE sequence silencing by DNA methylation together with the histone modifications in CSD-induced tolerance might counteract genome destabilization, thus preventing disabling phenomena such as senescence [25,26].

1.3. Mechanisms to Control the Transposable Elements Mobilization

Organisms evolved into different mechanisms to control the mobilization of TEs in order to counteract the damages caused by their transposition. One well-known mechanism in germline is based on interactions between the TEs and a plethora of non-coding RNAs. The PIWI-interacting RNAs (piRNAs), which interact with TEs at numerous levels, are the most studied of the regulatory RNAs. These RNAs can induce silencing at the transcriptional level or cause epigenetic modifications that alter the protein accessibility to DNA required for transcription. At the post-transcriptional level, they can

cause TE transcript degradation through the formation of double-stranded RNAs (dsRNAs) that can be cleaved into small-interfering RNAs (siRNAs), and they are associated with the RNA-induced silencing complex (RISC). In general, piRNAs act in the gonads to protect male and female germ-line genomes from transposable elements. Through a mechanism conserved among different species, piRNAs repress TE expression by modifying the histone 3 lysine 9 (H3K9) methylation state [27]. Specifically, nascent transcripts derived from full-length LINEs are recognized by MIWI2 protein/piRNAs complex that recruits a histone methyltransferase, resulting in the deposition of the histone 3 trimethyl lysine 9 (H3K9me3) mark on LINE repeats in the germ-line genome [27]. In an ovarian somatic cell line, piRNAs also formed complexes with H1/H3K9me3 and heterochromatin protein 1a (HP1a), modulating chromatin accessibility and thus affecting TE transcription [28]. There is also evidence for the piRNA-induced silencing of TEs in *Drosophila* somatic cells [29]. Other molecular mechanisms have also evolved, such as the distribution of TEs in regions of low gene density and the packaging of TEs into transcriptionally silent heterochromatin [30]. DNA methylation has largely been considered a strategy for blocking transposon mobility and preserving genome integrity in a wide range of cells, tissues, organisms, and biological contexts. The strict relationship between TEs and their silencing by DNA methylation is reflected by the fact that, because TEs represent such a large fraction of the mammalian genome, TE methylation analyses are often used as surrogates for global DNA methylation analyses [31,32].

2. Transposable Elements as an Essential Part of Gene Expression Regulation

TE mobilization, expression, and insertion can be deleterious if not property controlled, with negative effects on cell physiology. However, at the same time, they can be normally involved in shaping genomic structure and controlling gene expression. This is, for instance, evident in adult brain tissue, where a pervasive mobilization of the mobilization of the L1, Alu, and SINE-R/VNTR/Alu (SVA) transposons can reshape the gene regulatory networks of different kinds of neurons, particularly those in the hippocampus [33,34]. Similarly, other studies have also indicated a functional role for L1 transposition in neuronal precursor cells [11,35]. The expression of TEs in the brain is emblematic of this "double-edged sword" phenomenon, given that TEs can play a dual and conflicting role in the correct differentiation and development of neural mosaicism and in the onset of neurological disease. These are interesting examples of how TE evolution includes not only mechanisms for silencing the deleterious effects of these invading sequences, but also mechanisms for them to take an active and positive role in genome function. It is becoming clear that TEs are an integral part of the regulatory toolkit of the genome [36]. Repetitive elements have emerged as potential regulators for various biological processes [37], such as gene transcription, alternative splicing, and RNA translation. Even more evidence has accumulated showing how TEs can be instrumentally involved in the control of gene expression.

2.1. Transposable Elements and the Epigenetic Regulation of Gene Expression

Historically, the relationship between TEs and epigenetic modifications has often been seen as consisting of silencing TEs by DNA methylation and histone modification, with the distribution of TEs seemingly having evolved in parallel with the strategies to control their expression [38]. However, transposable elements are also emerging as an important source of epigenetic marks that can influence gene expression, leading to the idea that TE insertion may be involved in targeting epigenetic modifications to a specific locus. The causal effect of TE mobilization and distribution on the epigenetic regulation of gene expression is still a complex and unresolved issue. Evolution has favored that active intragenic TEs be preferentially inserted in an antisense direction with respect to the genes, producing a "trap" mechanism that can stop the invasion of these sequences during demethylation waves along the genome. To silence the TEs, these overlapping sense/antisense transcripts enter an endo-siRNA pathway that is regulated by DICER and Argonaute 2 (AGO2) and, when activated by global demethylation, leads to an increase in the repressive histone marks [39]. The molecular pathway that

underlies this link between the repression of TEs and the increase in histone repressive marks remains yet to be revealed [39]. New research in naïve murine embryonic stem cells pre-implantation indicates a central role for the KZFP/KAP1 (Krüppel associated box (KRAB) zinc finger protein/KRAB-associated protein 1) complex in maintaining heterochromatin, with DNA methylation marks at TEs protecting the loci from ten eleven translocation (TET)-mediated demethylation [40]. Ecco et al. [41] found that in differentiated tissues, two members of the KRAB/ZFP (KRAB/Zinc Finger Protein) family regulate TE targets through histone-based mechanisms that are not always correlated with the DNA methylation status of the loci. It is clear from the study that the interactions between the TEs and their KRAB-ZFP controllers influence nearby gene expression. This interaction is even clearer in human neural progenitor cells where it has been demonstrated that primate-specific ERVs act as docking platforms for the co-repressor protein KAP1 (also known as TRIM28) to establish local heterochromatin [42]. KAP1 binds and represses the ERVs, and consequently, regulates the expression of neighboring genes important for brain development [42]. Another example of the regulation of neighbouring genes by TEs includes the interactions of the transcriptional regulators human silencing hub (HUSH) and microrchidia family CW-type zinc finger 2 (MORC2) with evolutionarily young full-length L1s located in the transcriptionally permissive euchromatic region, which promotes the deposition of histone H3K9me3, a specific mark for transcriptional silencing. This MORC2/HUSH-bound L1 specific effect can spread to the neighboring genes, inducing a decrease in mRNA expression and probably influencing the RNA polymerase II (POL II) elongation rate [43]. Being a source of chromatin variation, specific classes of TEs (especially younger LINEs) have been found to impact chromatin accessibility in the livers of different inbred mice strains. This shows that TEs can regulate tissue-specific genes, which could result in downstream phenotypic diversity among populations [44]. Transposable elements can also exert a long-range regulation of gene expression by actively reshaping chromatin organization. Generally, and across different tissues, about 10% of TE families have been found to be enriched in active genomic regions. Among these, SINEs and DNA transposons are the most frequently enriched classes in the active chromatin regions, while L1 LINEs and ERV LTRs are the most frequently enriched TE classes in the repressed regions targeted with the H3K9me3 epigenetic mark [45].

Mechanisms of the Transposable Elements Co-option in the Epigenetic Control of Gene Expression

Intriguingly, the epigenetic influence of TEs is particularly evident in open euchromatin regions. For instance, the enrichment of repressive epigenetic marks around euchromatic TEs is due to the presence of TEs, rather than due to the preferential insertion of TEs into genomic regions already enriched with repressive epigenetic marks, as has been demonstrated by comparing the epigenomes of two *D. melanogaster* strains. This distribution accounted for finding higher histone 3 dimethyl lysine 9 (H3K9me2) enrichment and lower transcript levels for TE-flanking alleles than for the homologous alleles lacking nearby TE insertions [46]. Similarly, the analysis of epigenetic marks in flies with and without *Bari-Jheh*, a natural transposon that affects the expression of nearby genes, showed significant differences in histone 3 trimethyl lysine 4 (H3K4me3), H3K9me3, and histone 3 trimethyl lysine 27 (H3K27me3) histone marking in relation to oxidative stress conditions, highlighting that this TE element influences gene expression by affecting the local chromatin state [47]. These examples suggest that the different distribution of TEs occurring in the germlines of different organisms/strains and species strongly affect the gene expression of nearby genes (Figure 1A).

Interestingly, different TE distributions due to somatic transpositions affecting gene expression are also relevant in the same individual, leading to the idea that TEs have been co-opted, and their distributions co-evolved with gene expression regulation [48]. In fact, TE enrichment varies across tissues and TEs contain binding sites for tissue-specific master regulators of transcription [45]. One element of TE targeting comes directly from the fact that integration is restricted to open chromatin regions. This is apparent in mammalian brains, where somatic LINE insertions are enriched in the proximity of neuronal genes [11–33]. If we consider that TEs can be the target of epigenetic marks, the non-random and targeted tissue-specific distribution of TEs can be considered

a mechanism to genetically fix a landmark, which could induce an epigenetic control of nearby gene expression (Figure 1B). The finding that different environmental factors induce L1 transposition through different basic helix-loop-helix PER-ARNT-SIM (bHLH/PAS) proteins opens the possibility of the differential targeting of L1 insertions during different types of stresses [49]. Experimental findings hint that TE insertions are targeted beyond the apparent mechanistic necessity of open chromatin [50]. In *Arabidopsis thaliana*, TEs located in proximity to environmentally induced genes are silenced via enhanced DNA methylation, establishing a temporal and functional hierarchy of transcriptional and epigenomic changes in response to stress [51]. It is possible to argue that the mobilization and insertion of TEs can also be controlled in adult tissues and post-mitotic cells after specific stimuli for the purpose of driving the epigenetic regulation of specific genes (Figure 1C).
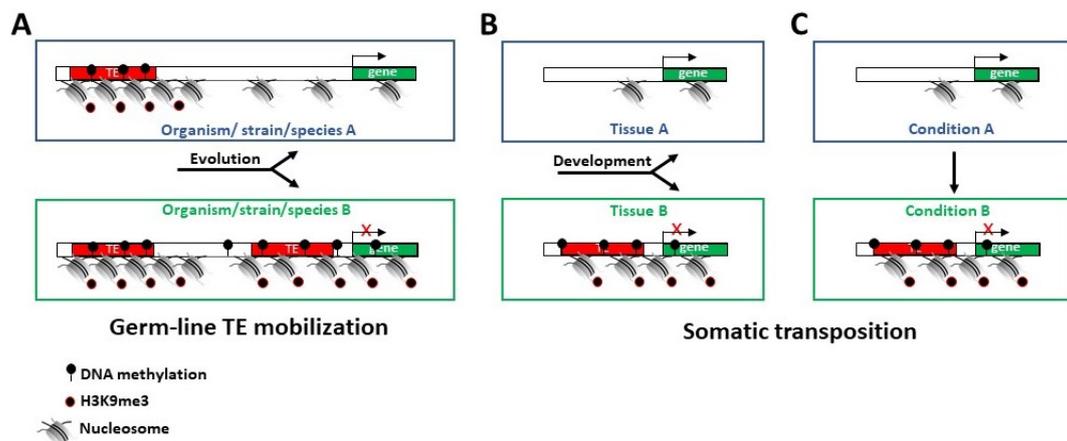

Figure 1. Effect of TE localization on the epigenetic regulation of gene expression. (**A**) The different distribution of TEs during evolution reshapes the epigenetic regulation of a specific gene. (**B**) The differential redistribution of TEs in different cells/tissues of the same organism during the development affects the expression of a specific gene. (**C**) The relocalization of TEs sequence in the same cell after a specific stimulus/condition affects the expression of a specific gene.

2.2. Transposable Elements and Long-Range Regulation

Throughout evolution, different families of TEs have originated multiple binding sites for transcription factors, producing different transcriptome landscapes. The ENCODE (Encyclopedia Of DNA Elements) data contains about 2 million transcription factor binding sites (TFBSs) that overlap with putatively regulation-competent human retrotransposons, which are located in a 5-kb gene promoter neighborhood composed of ~44% SINEs, ~33% LINEs, and 23% LR/ERVs [52]. The data show that LINE-derived transcription factor binding sites (TFBSs) are more numerous than SINEs outside the 5-kb region near the transcription start site, while the reverse occurs within the region. Moreover, the pathways most strongly impacted by the different retrotransposon distributions have been implicated in important processes, such as the cell stress and immune responses, ribosome biogenesis, chromatin remodeling, and DNA replication as well as mitotic spindle organization and cell cycle progression [52]. Nishihara et al. [53] were the first to report that an evolutionarily conserved genomic region named AS3_9, comprised of three TEs inserted side-by-side, functions as a distal enhancer for *wnt5a* expression during the morphogenesis of the mammalian secondary palate. Functional analyses have shown that, during mammalian evolution, the AmnSINE1, X6b_DNA, and MER117 retrotransposons have been co-opted together by a retroposition/transposition mechanism, with the acquisition inside of X6b_DNA sequence of a specific binding site for Msx1 protein, which together with Wnt5a is involved in the palatogenesis. This study suggested that the combination of different TEs located within the same DNA fragment could have played a significant role in the evolution of the high diversity of many *cis*-regulatory elements [53].

2.3. mRNA Decay and Splicing

Transposable Elements can also be involved in gene expression regulation, affecting mRNA stability through non-sense mediated decay as well as the activity of microRNAs, circular RNAs, and potentially long non-coding RNAs (lncRNAs). The Alu sequences, a subfamily of SINEs elements, are often embedded in pre- and mature mRNAs, usually as a part of the 3'-untranslated region (UTRs) or introns [54]. The Alu elements contained in mRNA sequences (and lncRNAs) can play a role in Staufen-mediated decay (SMD). It has been reported that STAU1 binding sites can be formed by imperfect base pairing between an Alu element within the 3'-UTR of an SMD target and another Alu element within a cytoplasmic and polyadenylated lncRNA. These sites can engage the STAU protein, triggering SMD through a downregulated mRNA expression profile [55] (Figure 2).
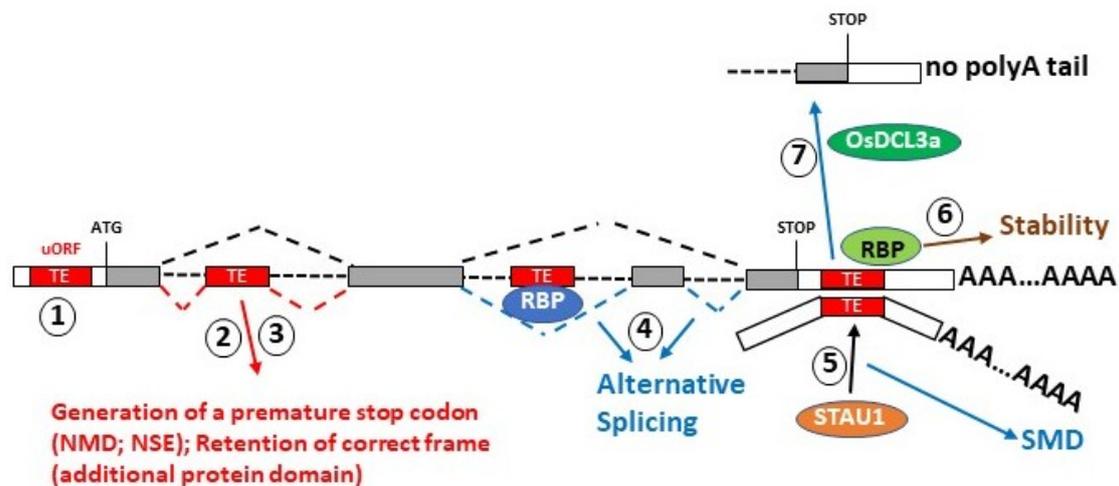
Figure 2. TE-dependent post-transcriptional regulation of gene expression. (1) TE inside the 5'-UTR cause the presence of an upstream ORF that participates in the regulation of the main ORF translation; (2) the exonization of a TE and thus its translation determines the presence of an additional domain inside the encoded protein; (3) the exonization of TE inside the coding region of an mRNA can result in the presence of a premature stop codon, thus resulting in Nonsense-mediated Decay process. (4) TE sequence can interact with RNA-binding proteins (RBP) and affects alternative splicing; (5) the presence of a TE sequence in the 3'-UTR of an mRNA can trigger STAU-mediated decay, (6) serves as a docking station for RBP involved in RNA stability such as HuR protein, (7) or triggers translational repression through the generation of a shorter mRNA without poly-A tail.

Lineage-specific 3'-UTR SINEs can function to regulate the levels of mRNAs from orthologous genes in different species (including human) by directing SMD, demonstrating the specific role that lineage-specific SINEs have played in the convergence of gene expression profiles between species [56]. Transposable element-derived sequences localized in RNA transcripts can also play a role in the regulation of mRNA abundance and alternative splicing. In a process called exonization, TEs inserted into introns are recognized by the splicing machinery and recruited into RNA transcripts as exons. TEs contain numerous splice-donor and -acceptor sites that contribute to alternative splicing. Thus, they can interact with a plethora of RNA binding proteins (RBPs) with composition site preferences that drive them to specific regions of TEs, for example Human antigen R (HuR) or Fused in Sarcoma (FUS) proteins, that prefer to bind to U-rich motifs [56]. However, the depletion of TE binding sites for different RBPs indicates these sites affect transcript abundance and splicing in a way similar to that seen in gene-binding sites located in non-repetitive sequences [56]. It has been also noticed that in a few specific cases, the effect of RBP binding can change depending on the specific TE family bound. For example, the RBP HuR confers transcript stability unless it is bound to an Alu element in a U-rich region [57] (Figure 2). The Alu-containing RNAs may also be capable of forming stable structural domains, which would likely give rise to novel biological functions [58]. Kralovicova et al. [59]

have shown that an intronic transposed element highly similar to medium reiterated frequency repeat family 51 (MER51A) can influence gene expression, modulating the inclusion levels of many nonsense-mediated RNA decay switch exon (NSEs), pseudo-exons derived by the activation of cryptic splice sites that act as a buffer to counteract Alu-mediated NSE activation (Figure 2).

2.4. Transposable Elements and Noncoding RNAs

Another class of small RNAs implicated in TE regulation is the circular RNAs (circRNA), which are a new class of small noncoding RNAs with gene regulation functions. Recent mammalian studies have shown that the flanking regions of circRNAs are enriched with transposons that have the potential to mediate circRNA formation via reverse complementary pairing [60,61]. In humans, the presence of different circRNAs derived from a single gene locus is associated with a variety of Alu pairings in human introns, suggesting a role for competition between the pairings in the formation of alternative circularizations [62]. A new study in maize that sequenced circRNA showed that LINE1-like elements and their reverse complementary Pairs (LLERCPs) are significantly enriched in the flanking regions of circRNAs [63]. Interestingly, when LLERCP transcription increases, the accumulation of circRNAs varies, leading to a decrease in linear transcripts [63]. Jung et al. [64] described how TEs can affect the expression of another class of long non-coding RNAs called cis-natural antisense transcript (cis-NATs) in different ways. First, NATs can originate from TEs and can be transcribed by alternative promoters depending on the TE. They can also be exonized by TEs, with the newly formed exon being complementary to the exon of a sense protein-coding gene [64]. Thus, the expression of sense transcripts may be mediated by NATs through dsRNA formations, that can participate in the RNA interference or adenosine deaminase acting on RNA (ADAR) pathways.

2.5. Transposable Elements and Protein Translation

Retrotransposons have evolved side by side with genes, inserting into different positions along the gene bodies and producing a plethora of effects. Transposable element insertions into 5′-UTRs or 3′-UTRs of mRNAs affects the protein expression of many genes in different ways (Figure 2). Kitano et al. [65] showed that TEs are also implicated in the upstream open reading frame (uORF)-mediated translational regulation of a large number of genes. The major families of retrotransposons, such as LINEs and SINEs can operate as cis-acting elements to inhibit or promote the translation of canonical ORFs located downstream of the uORFs in eukaryotic mRNAs. Using the human RefSeq mRNA sequence database, Kitano et al. [65] demonstrated that about 10% of human uORFs are generated and regulated by TEs localized in the 5′-UTR of mRNAs. While previous studies have suggested that retrotransposons act as translational regulators, the role of DNA transposons in influencing protein-host translation is still not clear. Miniature inverted-repeat transposable elements (MITEs) are sub-family of DNA transposons widely distributed in plant and animal genomes. It has been reported that their localization in the 3′-UTRs of rice mRNAs can exert a regulatory function via a translational repression mechanism. One example is the Ghd2 gene, a member of the CCT (CONSTANS (CO), CO-LIKE and TIMING OF CAB1) gene family in rice that regulates important agronomic traits, such as grain number, plant height, and heading date [66]. The mechanism underlying the translational repression of Ghd2 by the MITEs results on the Dicer-like 3a (OsDCL3a) pathway which can generate shortened mRNA lacking a poly-A tail by processing the MITE nascent transcripts [66]. However, which mRNA translation phase is blocked and how the mRNA translation is repressed by MITEs remains unknown [66]. When in the coding region of genes, TEs can be implicated in generating new alternative mRNA splicing isoforms. This can be seen as an intermediate step in the rise of new genes during evolution. For example, the mammalian thymopoietin (TMPO) and zinc finger protein 451 (ZNF451) genes both code for splice isoforms containing lamina-associated polypeptide 2alpha (LAP2alpha) domains, which are related to the first ORF from a Dictyostelium intermediate repeat sequence 1 (DIRS1)-like retrotransposon. Both mRNAs produce the canonical protein and a new

non-canonical protein isoform. Specifically, the LAP2alpha specific isoform of TMPO (LAP2a) was co-opted during evolution to support a new and important role in the cell [67].

The evolutionary insertion of TEs in gene coding regions has also produced chimeras in a process called domestication. The activity regulated cytoskeleton associated Protein (Arc) is a good example of retrotransposon domestication, with much of the evidence suggesting it is derived from a vertebrate lineage of Ty3/gypsy retrotransposons [68]. In particular, Arc is a cellular immediate-early gene required for learning and memory, and whose mRNA is localized at the synaptic junction. Interestingly, the regulation of Arc mRNA resembles that of viral RNA, as Arc contains an internal ribosomal entry site that enables cap-independent translation [69]. Crystallography has shown that protein structure of the Arc subdomains forms a bi-lobar architecture similar to the capsid domain of the human immunodeficiency virus gag protein [69]. These findings suggest that evolution have repurposed Gag containing elements to mediate intercellular communication in the nervous system [69].

3. Conclusions

We have provided an overview of several new findings that corroborate the ever more evident active role of TEs in genome function, emphasizing their effects on gene expression regulation. While in many examples, it is evident that TE co-option has been positively selected for during evolution, with TE sequences becoming fixed in the genome, other examples have argued the possibility that more complex molecular mechanisms have evolved to control mobilization and co-option in the same individual during specific periods of embryonic development or under specific environmental stimuli. Although interesting questions about these mechanisms remain unraveled, it is evident that TEs represent a powerful arsenal in the many layers of gene expression regulation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. USA* **1950**, *36*, 344–355. [CrossRef]
2. McClintock, B. Chromosome organization and genic expression. *Cold Spring Harb. Symp. Quant. Biol.* **1951**, *16*, 13–47. [CrossRef] [PubMed]
3. Belyayev, A. Bursts of transposable elements as an evolutionary driving force. *J. Evol. Biol.* **2014**, *27*, 2573–2584. [CrossRef] [PubMed]
4. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982. [CrossRef] [PubMed]
5. Gifford, W.D.; Pfaff, S.L.; Macfarlan, T.S. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* **2013**, *23*, 218–226. [CrossRef] [PubMed]
6. Garcia-Perez, J.L.; Widmann, T.J.; Adams, I.R. The impact of transposable elements on mammalian development. *Development* **2016**, *143*, 4101–4114. [CrossRef] [PubMed]
7. Lee, S.I.; Kim, N.S. Transposable elements and genome size variations in plants. *Genom. Inform.* **2014**, *12*, 87–97. [CrossRef]
8. Sultana, T.; Zamborlini, A.; Cristofari, G.; Lesage, P. Integration site selection by retroviruses and transposable elements in eukaryotes. *Nat. Rev. Genet.* **2017**, *18*, 292–308. [CrossRef]
9. Belancio, V.P.; Deininger, P.L.; Roy-Engel, A.M. LINE dancing in the human genome: Transposable elements and disease. *Genome Med.* **2009**, *27*, 97. [CrossRef]
10. Haig, D. Transposable elements: Self-seekers of the germline, team-players of the soma. *Bioessays* **2016**, *38*, 1158–1166. [CrossRef]
11. Muotri, A.R.; Chu, V.T.; Marchetto, M.C.; Deng, W.; Moran, J.V.; Gage, F.H. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **2005**, *435*, 903–910. [CrossRef] [PubMed]

12. Kazazian, H.H., Jr.; Wong, C.; Youssoufian, H.; Scott, A.F.; Phillips, D.G.; Antonarakis, S.E. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* **1988**, *332*, 164–166. [CrossRef] [PubMed]

13. Bingham, P.M.; Kidwell, M.G.; Rubin, G.M. The molecular basis of P-M hybrid dysgenesis: The role of the P element, a P-strain-specific transposon family. *Cell* **1982**, *25*, 995–1004. [CrossRef]

14. Pereira, C.M.; Stoffel, T.J.R.; Callegari-Jacques, S.M.; Hua-Van, A.; Capy, P.; Loreto, E.L.S. The somatic mobilization of transposable element mariner-Mos1 during the Drosophila lifespan and its biological consequences. *Gene* **2018**, *679*, 65–72. [CrossRef]

15. Tang, Z.; Steranka, J.P.; Ma, S.; Grivainis, M.; Rodić, N.; Huang, C.R.; Shih, I.M.; Wang, T.L.; Boeke, J.D.; Fenyö, D.; et al. Human transposon insertion profiling: Analysis, visualization and identification of somatic LINE-1 insertions in ovarian cancer. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E733–E740. [CrossRef]

16. Pitkänen, E.; Cajuso, T.; Katainen, R.; Kaasinen, E.; Välimäki, N.; Palin, K.; Taipale, J.; Aaltonen, L.A.; Kilpivaara, O. Frequent L1 retrotranspositions originating from TTC28 in colorectal cancer. *Oncotarget* **2014**, *5*, 853–859. [CrossRef]

17. Brégnard, C.; Guerra, J.; Déjardin, S.; Passalacqua, F.; Benkirane, M.; Laguette, N. Upregulated LINE-1 Activity in the Fanconi Anemia Cancer Susceptibility Syndrome Leads to Spontaneous Pro-inflammatory Cytokine Production. *EBioMedicine* **2016**, *8*, 184–194. [CrossRef]

18. Doucet-O'Hare, T.T.; Sharma, R.; Rodić, N.; Anders, R.A.; Burns, K.H.; Kazazian, H.H., Jr. Somatically Acquired LINE-1 Insertions in Normal Esophagus Undergo Clonal Expansion in Esophageal Squamous Cell Carcinoma. *Hum. Mutat.* **2016**, *37*, 942–954. [CrossRef]

19. Rodić, N.; Steranka, J.P.; Makohon-Moore, A.; Moyer, A.; Shen, P.; Sharma, R.; Kohutek, Z.A.; Huang, C.R.; Ahn, D.; Mita, P.; et al. Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. *Nat. Med.* **2015**, *21*, 1060–1064. [CrossRef]

20. Doyle, G.A.; Crist, R.C.; Karatas, E.T.; Hammond, M.J.; Ewing, A.D.; Ferraro, T.N.; Hahn, C.G.; Berrettini, W.H. Analysis of LINE-1 Elements in DNA from Postmortem Brains of Individuals with Schizophrenia. *Neuropsychopharmacology* **2017**, *42*, 2602–2611. [CrossRef]

21. Shpyleva, S.; Melnyk, S.; Pavliv, O.; Pogribny, I.; Jill James, S. Overexpression of LINE-1 Retrotransposons in Autism Brain. *Mol. Neurobiol.* **2018**, *55*, 1740–1749. [CrossRef] [PubMed]

22. Bundo, M.; Toyoshima, M.; Okada, Y.; Akamatsu, W.; Ueda, J.; Nemoto-Miyauchi, T.; Sunaga, F.; Toritsuka, M.; Ikawa, D.; Kakita, A.; et al. Increased l1 retrotransposition in the neuronal genome in schizophrenia. *Neuron.* **2014**, *81*, 306–313. [CrossRef] [PubMed]

23. Guo, C.; Jeong, H.H.; Hsieh, Y.C.; Klein, H.U.; Bennett, D.A.; De Jager, P.L.; Liu, Z.; Shulman, J.M. Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep.* **2018**, *23*, 2874–2880. [CrossRef] [PubMed]

24. Stetler, R.A.; Leak, R.K.; Gan, Y.; Li, P.; Zhang, F.; Hu, X.; Jing, Z.; Chen, J.; Zigmond, M.J.; Gao, Y. Preconditioning provides neuroprotection in models of CNS disease: Paradigms and clinical significance. *Prog. Neurobiol.* **2014**, *114*, 58–83. [CrossRef] [PubMed]

25. Rana, G.; Donizetti, A.; Virelli, G.; Piscopo, M.; Viggiano, E.; De Luca, B.; Fucci, L. Cortical spreading depression differentially affects lysine methylation of H3 histone at neuroprotective genes and retrotransposon sequences. *Brain Res.* **2012**, *1467*, 113–119. [CrossRef] [PubMed]

26. Drongitis, D.; Rainone, S.; Piscopo, M.; Viggiano, E.; Viggiano, A.; De Luca, B.; Fucci, L.; Donizetti, A. Epigenetics and cortical spreading depression: Changes of DNA methylation level at retrotransposon sequences. *Mol. Biol Rep.* **2016**, *43*, 755–760. [CrossRef] [PubMed]

27. Pezic, D.; Manakov, S.A.; Sachidanandam, R.; Aravin, A.A. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev.* **2014**, *28*, 1410–1428. [CrossRef]

28. Iwasaki, Y.W.; Murano, K.; Ishizu, H.; Shibuya, A.; Iyoda, Y.; Siomi, M.C.; Siomi, H.; Saito, K. Piwi Modulates Chromatin Accessibility by Regulating Multiple Factors Including Histone H1 to Repress Transposons. *Mol. Cell.* **2016**, *63*, 408–419. [CrossRef]

29. Malone, C.D.; Brennecke, J.; Dus, M.; Stark, A.; McCombie, W.R.; Sachidananda, R.; Hannon, G.J. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **2009**, *137*, 522–535. [CrossRef]

30. Lippman, Z.; Gendrel, A.V.; Black, M.; Vaughn, M.W.; Dedhia, N.; McCombie, W.R.; Lavine, K.; Mittal, V.; May, B.; Kasschau, K.D.; et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **2004**, *430*, 471–476. [CrossRef]

31. Yang, G.; Lee, Y.H.; Jiang, Y.; Shi, X.; Kertbundit, S.; Hall, T.C. A two-edged role for the transposable element Kiddo in the rice ubiquitin2 promoter. *Plant Cell* **2005**, *17*, 1559–1568. [CrossRef] [PubMed]

32. Baba, Y.; Yagi, T.; Sawayama, H.; Hiyoshi, Y.; Ishimoto, T.; Iwatsuki, M.; Miyamoto, Y.; Yoshida, N.; Baba, H. Long Interspersed Element-1 Methylation Level as a Prognostic Biomarker in Gastrointestinal Cancers. *Digestion* **2018**, *97*, 26–30. [CrossRef] [PubMed]

33. Baillie, J.K.; Barnett, M.W.; Upton, K.R.; Gerhardt, D.J.; Richmond, T.A.; De Sapio, F.; Brennan, P.M.; Rizzu, P.; Smith, S.; Fell, M.; et al. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* **2011**, *479*, 534–537. [CrossRef] [PubMed]

34. Upton, K.R.; Gerhardt, D.J.; Jesuadian, J.S.; Richardson, S.R.; Sánchez-Luque, F.J.; Bodea, G.O.; Ewing, A.D.; Salvador-Palomeque, C.; van der Knaap, M.S.; Brennan, P.M.; et al. Ubiquitous L1 mosaicism in hippocampal neurons. *Cell.* **2015**, *161*, 228–239. [CrossRef] [PubMed]

35. Coufal, N.G.; Garcia-Perez, J.L.; Peng, G.E.; Yeo, G.W.; Mu, Y.; Lovci, M.T.; Morell, M.; O'Shea, K.S.; Moran, J.V.; Gage, F.H. L1 retrotransposition in human neural progenitor cells. *Nature* **2009**, *460*, 1127–1131. [CrossRef]

36. Slotkin, R.K.; Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **2007**, *8*, 272–285. [CrossRef]

37. De Souza, F.S.; Franchini, L.F.; Rubinstein, M. Exaptation of transposable elements into novel cis-regulatory elements: Is the evidence always strong? *Mol. Biol. Evol.* **2013**, *30*, 1239–1251. [CrossRef]

38. Hollister, J.D.; Gaut, B.S. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **2009**, *19*, 1419–1428. [CrossRef]

39. Berrens, R.V.; Andrews, S.; Spensberger, D.; Santos, F.; Dean, W.; Gould, P.; Sharif, J.; Olova, N.; Chandra, T.; Koseki, H.; et al. An endosiRNA-Based Repression Mechanism Counteracts Transposon Activation during Global DNA Demethylation in Embryonic Stem Cells. *Cell Stem Cell* **2017**, *21*, 694–703. [CrossRef]

40. Coluccio, A.; Ecco, G.; Duc, J.; Offner, S.; Turelli, P.; Trono, D. Individual retrotransposon integrants are differentially controlled by KZFP/KAP1-dependent histone methylation, DNA methylation and TET-mediated hydroxymethylation in naïve embryonic stem cells. *Epigenetics Chromatin* **2018**, *11*, 7. [CrossRef]

41. Ecco, G.; Cassano, M.; Kauzlaric, A.; Duc, J.; Coluccio, A.; Offner, S.; Imbeault, M.; Rowe, H.M.; Turelli, P.; Trono, D. Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Dev. Cell.* **2016**, *36*, 611–623. [CrossRef]

42. Brattås, P.L.; Jönsson, M.E.; Fasching, L.; Nelander Wahlestedt, J.; Shahsavani, M.; Falk, R.; Falk, A.; Jern, P.; Parmar, M.; Jakobsson, J. TRIM28 Controls a Gene Regulatory Network Based on Endogenous Retroviruses in Human Neural Progenitor Cells. *Cell Rep.* **2017**, *18*, 1–11. [CrossRef] [PubMed]

43. Liu, N.; Lee, C.H.; Swigut, T.; Grow, E.; Gu, B.; Bassik, M.C.; Wysocka, J. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* **2018**, *553*, 228–232. [CrossRef]

44. Du, J.; Leung, A.; Trac, C.; Lee, M.; Parks, B.W.; Lusis, A.J.; Natarajan, R.; Schones, D.E. Chromatin variation associated with liver metabolism is mediated by transposable elements. *Epigenetics Chromatin.* **2016**, *9*, 28. [CrossRef] [PubMed]

45. Trizzino, M.; Kapusta, A.; Brown, C.D. Transposable elements generate regulatory novelty in a tissue-specific fashion. *BMC Genom.* **2018**, *19*, 468. [CrossRef]

46. Lee, Y.C.G.; Karpen, G.H. Pervasive epigenetic effects of Drosophila euchromatic transposable elements impact their evolution. *Elife* **2017**, *6*, e25762. [CrossRef] [PubMed]

47. Guio, L.; Vieira, C.; González, J. Stress affects the epigenetic marks added by natural transposable element insertions in Drosophila melanogaster. *Sci. Rep.* **2018**, *8*, 12197. [CrossRef]

48. Rebollo, R.; Romanish, M.T.; Mager, D.L. Transposable elements: An abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* **2012**, *46*, 21–42. [CrossRef] [PubMed]

49. Ishizaka, Y.; Okudaira, N.; Tamura, M.; Iijima, K.; Shimura, M.; Goto, M.; Okamura, T. Modes of retrotransposition of long interspersed element-1 by environmental factors. *Front. Microbiol.* **2012**, *3*, 191. [CrossRef]

50. Mourier, T.; Nielsen, L.P.; Hansen, A.J.; Willerslev, E. Transposable elements in cancer as a by-product of stress-induced evolvability. *Front. Genet.* **2014**, *5*, 156. [CrossRef]

51. Secco, D.; Wang, C.; Shou, H.; Schultz, M.D.; Chiarenza, S.; Nussaume, L.; Ecker, J.R.; Whelan, J.; Lister, R. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *Elife* **2015**, *4*, e09343. [CrossRef] [PubMed]

52. Nikitin, D.; Penzar, D.; Garazha, A.; Sorokin, M.; Tkachev, V.; Borisov, N.; Poltorak, A.; Prassolov, V.; Buzdin, A.A. Profiling of Human Molecular Pathways Affected by Retrotransposons at the Level of Regulation by Transcription Factor Proteins. *Front. Immunol.* **2018**, *9*, 30. [CrossRef] [PubMed]

53. Nishihara, H.; Kobayashi, N.; Kimura-Yoshida, C.; Yan, K.; Bormuth, O.; Ding, Q.; Nakanishi, A.; Sasaki, T.; Hirakawa, M.; Sumiyama, K.; et al. Coordinately Co-opted Multiple Transposable Elements Constitute an Enhancer for wnt5a Expression in the Mammalian Secondary Palate. *PLoS Genet.* **2016**, *12*, e1006380. [CrossRef] [PubMed]

54. Daniel, C.; Lagergren, J.; Öhman, M. RNA editing of non-coding RNA and its role in gene regulation. *Biochimie* **2015**, *117*, 22–27. [CrossRef] [PubMed]

55. Gong, C.; Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **2011**, *470*, 284–288. [CrossRef]

56. Lucas, B.A.; Lavi, E.; Shiue, L.; Cho, H.; Katzman, S.; Miyoshi, K.; Siomi, M.C.; Carmel, L.; Ares, M., Jr.; Maquat, L.E. Evidence for convergent evolution of SINE-directed Staufen-mediated mRNA decay. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 968–973. [CrossRef]

57. Kelley, D.R.; Hendrickson, D.G.; Tenen, D.; Rinn, J.L. Transposable elements modulate human RNA abundance and splicing via specific RNA-protein interactions. *Genome Biol.* **2014**, *15*, 537. [CrossRef]

58. Kim, E.Z.; Wespiser, A.R.; Caffrey, D.R. The domain structure and distribution of Alu elements in long noncoding RNAs and mRNAs. *RNA* **2016**, *22*, 254–264. [CrossRef]

59. Kralovicova, J.; Moreno, P.M.; Cross, N.C.; Pêgo, A.P.; Vorechovsky, I. Antisense Oligonucleotides Modulating Activation of a Nonsense-Mediated RNA Decay Switch Exon in the ATM Gene. *Nucleic Acid Ther.* **2016**, *26*, 392–400. [CrossRef]

60. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2013**, *19*, 141–157. [CrossRef]

61. Zhang, X.O.; Wang, H.B.; Zhang, Y.; Lu, X.; Chen, L.L.; Yang, L. Complementary sequence-mediated exon circularization. *Cell* **2014**, *159*, 134–147. [CrossRef] [PubMed]

62. Chen, L.L.; Yang, L. Regulation of circRNA biogenesis. *RNA Biol.* **2015**, *12*, 381–388. [CrossRef] [PubMed]

63. Chen, L.; Zhang, P.; Fan, Y.; Lu, Q.; Li, Q.; Yan, J.; Muehlbauer, G.J.; Schnable, P.S.; Dai, M.; Li, L. Circular RNAs mediated by transposons are associated with transcriptomic and phenotypic variation in maize. *New Phytol.* **2018**, *217*, 1292–1306. [CrossRef] [PubMed]

64. Jung, J.; Lee, S.; Cho, H.S.; Park, K.; Ryu, J.W.; Jung, M.; Kim, J.; Kim, H.; Kim, D.S. Bioinformatic analysis of regulation of natural antisense transcripts by transposable elements in human mRNA. *Genomics* **2019**, *111*, 159–166. [CrossRef]

65. Kitano, S.; Kurasawa, H.; Aizawa, Y. Transposable elements shape the human proteome landscape via formation of cis-acting upstream open reading frames. *Genes Cells* **2018**, *23*, 274–284. [CrossRef]

66. Shen, J.; Liu, J.; Xie, K.; Xing, F.; Xiong, F.; Xiao, J.; Li, X.; Xiong, L. Translational repression by a miniature inverted-repeat transposable element in the 3' untranslated region. *Nat. Commun.* **2017**, *8*, 14651. [CrossRef]

67. Abascal, F.; Tress, M.L.; Valencia, A. Alternative splicing and co-option of transposable elements: The case of TMPO/LAP2α and ZNF451 in mammals. *Bioinformatics* **2015**, *31*, 2257–2261. [CrossRef]

68. Volff, J.N. Cellular genes derived from Gypsy/Ty3 retrotransposons in mammalian genomes. *Ann. N. Y. Acad. Sci.* **2009**, *1178*, 233–243. [CrossRef]

69. Pastuzyn, E.D.; Day, C.E.; Kearns, R.B.; Kyrke-Smith, M.; Taibi, A.V.; McCormick, J.; Yoder, N.; Belnap, D.M.; Erlendsson, S.; Morado, D.R.; et al. The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell* **2018**, *173*, 275. [CrossRef]