Review title:

Epigenetic regulation of intragenic transposable elements: A two-edged sword.

Author:

Hidetoshi Saze

Affiliation:

Plant Epigenetics Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Okinawa, 904-0495 Japan.

Running title: Regulation of intragenic transposon

Contact information:

Hidetoshi Saze

1919-1 Tancha, Onna-son, Okinawa, 904-0495 Japan.

TEL +81-98-966-8494

FAX +81-98-966-1064

E-mail hidetoshi.saze@oist.jp

Summary

Genomes of animals and plants contain a large number of transposable elements (TEs). TEs often transpose into genic regions, affecting expression of surrounding genes. Intragenic TEs mostly reside in introns, and in much the same way as intergenic TEs, they are targeted by repressive epigenetic marks for transcriptional silencing. Silenced intragenic TEs generally co-repress expression of associated genes, while in some cases they significantly enhance splicing and transcript elongation. Genomes have evolved molecular mechanisms that allow the presence of silenced TEs within transcriptionally permissive chromatin environments. Epigenetic modulation of intragenic TEs often contributes to gene regulation, phenotypic expression, and genome evolution.

Keywords
Transposable elements/intron/epigenetics/heterochromatin/plant/

Introduction

Transposable elements (TEs) are parasitic genetic elements in the genomes of most organisms that can amplify their own copy numbers in the genome, contributing to genome expansion and evolution (*1, 2*). For instance, TEs comprises nearly 50% of the human genome (*3, 4*) and over 80% of the maize genome (*5*). Since TEs are mutagenic, host genomes transcriptionally inactivate TEs using epigenetic silencing mechanisms, such as DNA cytosine methylation, histone H3K9 methylation, and RNAi (*6, 7*). Nevertheless, substantial numbers of TEs accumulate within transcriptional gene units. In the human genome, 60% of TEs are localized within introns that comprise only 24% of the genome (*8*). Interestingly, these "intragenic" TEs are associated with repressive epigenetic marks and form heterochromatin, in much the same way as intergenic TEs, even within actively transcribed regions. Organisms have evolved specific molecular mechanisms that allow the presence of heterochromatic TE sequences within a transcription-permissive chromatin environment. Recent findings of epigenetic regulation of intragenic TEs and their functional implications, especially in animals and plants, will be reviewed.

Genome evolution and intragenic TEs.

Insertion of TEs into genic regions causes transcriptional attenuation of genes, due to disruption of the coding sequence and creation of novel regulatory elements, such as polyadenylation signals, cryptic promoters, and donor/acceptor sequences for alternative splicing (*3, 9*) (Figure 1). Still, many TEs are found within transcriptional gene units in eukaryote genomes (*10*). In mammals, the majority of intragenic TEs, especially Long Interspersed Nuclear Element 1 (LINE 1; L1) and Short Interspersed Nuclear Element (SINE) families, are enriched in introns (*3, 8*). Their sequences are often integrated as a part of mature transcripts, a state termed "exonization" (*11*), which contributes to splicing variant diversity of transcripts (*12, 13*) (Figure 1).

As in animals, plant genomes also harbor intragenic TEs that contribute to gene organization and genome evolution. For example, about 10% of genes in the maize genome contain >1 kb TEs in their introns (*14*). The genome of Norway spruce (*Picea*

abies, a gymnosperm) is >100-fold larger (~20 Gb) than that of the model plant *Arabidopsis thaliana* (130 Mb), although the number of genes (~28,000) and their exon sizes are comparable. Instead, genes in Norway spruce have longer introns (*15*). This is a result of accumulation of LTR-type retrotransposons in intronic regions over millions of years, without elimination of those sequences. A recent analysis of the *A. thaliana* genome showed that 0.7% of annotated genes contain about 3% of all TEs, the majority of them (~80%) residing within introns (*16*). *Arabidopsis lyrata*, a close relative of *A. thaliana*, has much longer introns than *A. thaliana* due to TE accumulation, suggesting that intron expansion by TE insertion can occur in species-specific manner. These observations indicate that intragenic TEs, especially intronic TEs, are tolerated within transcriptional units in many species, even though they are potentially deleterious for both coding sequences and regulatory elements.

Epigenetic regulation of intragenic TEs

Intragenic TEs are subjected to both genetic and epigenetic regulation. In *A. thaliana*, most intragenic TEs are much shorter than intergenic TEs, suggesting that full-length, intact TEs are selected against and are degraded during evolution (*16-18*). TE sequences as well as their cryptic regulatory sequences, such as promoters and polyA signals, can be eliminated by mechanisms of unequal homologous recombination and illegitimate recombination (*19*). In addition to the genetic changes, epigenetic mechanisms play important roles in inactivation of intragenic TEs. Intragenic TEs are targeted by repressive epigenetic marks, such as DNA methylation and histone H3K9 methylation, even within genes actively transcribed by RNA polymerase II (*20*). Recent studies in mammals showed that repressive histone H3K9 methylation is deposited on L1 family in intronic region for transcriptional silencing, which is likely mediated by the H3K9 methyltransferase SETDB1 (*21*). H3K9 methylation is recruited to L1s by the human silencing hub (HUSH) complex, comprising TASOR, MPP8, and periphilin proteins (*21-23*). In addition, Microrchidia 2 (MORC2; MORC2A in mice), a member of the conserved ATPase domain-containing protein family, is required for deposition of H3K9 methylation to intergenic/intronic TEs, as well as endogenous retroviruses (*21, 24*) (Figure 2A). L1s targeted by HUSH/MORC2 tend to be full-length, evolutionarily

younger, and enriched in the transcriptionally permissive chromatin environment associated with active epigenetic marks, such as H3K4 methylation and H3K27 acetylation (*21, 22*). Transcriptional induction of L1-containing genes results in accumulation of HUSH/MORC2 on intronic L1 sequences, suggesting that RNA polymerase II (PolII) transcription is required for recruitment of the silencing complex to the intronic TEs. The PolII-dependent recruitment of H3K9 methylation is similar to the process for heterochromatin formation at pericentromeric repeats in fission yeast (*25*).

In plants, intronic TEs are also associated with repressive epigenetic marks, such as CG and non-CG methylation, H3K9 methylation, and small interfering RNAs (siRNAs) (*26*), that are required for formation of heterochromatin at TE sequences (*16*). In *A. thaliana*, the heterochromatic state of intronic TEs can be maintained by maintenance CG methylase MET1, histone H3K9 methyltransferases KYP/SUVH4, SUVH5, SUVH6, and the chromatin remodeler DDM1 (*16, 27*). Although detailed molecular events for recognition and initiation of epigenetic silencing of intragenic TEs still remain elusive, TEs newly inserted into intragenic regions can be recognized by host defense mechanisms involving RNA-directed DNA methylation (RdDM) (*28*) (Figure 2B). A recent study using a heat-inducible LTR-type retrotransposon *ONSEN/COPIA78* showed that after transposition into an intron of an active gene, TEs seem to be maintained in active chromatin state for multiple generations (*29*). Interestingly, genetic crosses of *Arabidopsis* strains efficiently induce DNA methylation and silencing of intronic *ONSEN*. This silencing process depends on siRNA and RdDM, since loss of RNA polymerase IV, which is required for production of 24-nt siRNA and *de novo* DNA methylation, abolishes the silencing (Figure 2B). It is likely that siRNAs produced from homologous *ONSEN* copies in other loci can guide the RdDM machinery to the newly introgressed intronic *ONSEN* in the genome. The TE silencing occurs immediately and efficiently in F1 plants, but how this silencing *in trans* takes place during plant development remains unclear. Transgenes inserted in intronic regions also show behavior similar to that of TEs (*30*).

Impact of intragenic TEs on gene expression

In general, repressive epigenetic modifications spread into neighboring chromatin and down-regulate expression of surrounding genes. Indeed, in mammals, heterochromatin formation on intronic L1s by binding of HUSH and MORC2 causes co-

repression of both the TEs and associated genes (*21, 22*) (Figure 2A). It is also reported that genic H3K9 methylation inhibits PolII elongation, causes accumulation of PolII over H3K9 methylated regions, and promotes alternative splicing (*31, 32*). In plants, TEs associated repressive epigenetic marks also have a negative effect on gene expression; therefore, TEs have been evolutionarily purged from genic regions (*33, 34*). However, recent studies in plants indicate that outcomes of heterochromatin formation at intronic TEs may differ from those in mammals, in some respects. For example, expression levels of TE-containing genes in some plant species are comparable to those of genes without TE insertions (*14, 15*). In addition, loss of intronic heterochromatin in *A. thaliana* leads to premature polyadenylation within or before intronic TE sequences, rather than de-repression of associated genes (*16, 27*). These observations suggest that in plants, intronic heterochromatin tends to enhance proper splicing and/or to mask cryptic polyA signals within TE sequences.

Plants seem to have evolved specific epigenetic mechanisms to "skip" obstructive heterochromatic domain within introns. Genetic and biochemical studies in *Arabidopsis* identified factors required for this mechanism. ENHANCED DOWNY MILDEW 2 (EDM2) was identified as a factor required for proper transcription of a disease-resistance gene that contains many heterochromatic TEs in introns (*27, 35*). EDM2 has PHD domains that recognize H3K9 methylation, and a putative methyltransferase domain in its C-terminal part (*36*). Another factor, INCREASE IN BONSAI METHYLATION 2 (IBM2) /ANTI-SILENCING 1 (ASI1) /SHOOT GROWTH 1 (SG1) contains an RNA-Recognition Motif (RRM) and a Bromo-Adjacent Homology (BAH) domain that likely binds to chromatin (*36-39*) (Figure 2B). In addition, several interacting factors of EDM2 have been identified (*40*). Mutants of these factors show similar molecular phenotypes: transcripts from genes containing heterochromatic domains in their introns are prematurely terminated and polyadenylated within the intron. This defect is similar to that in mutants defective in heterochromatin formation, such as *ddm1* and *suvh456* (The triple mutant for KYP/SUVH4, SUVH5, SUVH6) (*16, 27*). However, intronic heterochromatin is still maintained in *ibm2* and *edm2*, indicating that these factors are acting downstream of DNA methylation and H3K9 methylation. In addition, PolII elongation seems not to be affected by loss of these factors (*37*), suggesting their roles in post-transcriptional events, such as splicing and suppression of polyA signals. How these factors are preferentially recruited to intronic TEs, but not to intergenic

TEs remains unclear.

Gene regulatory role of intragenic TEs

Intragenic TE sequences associated with epigenetic modifications often acquire regulatory roles in gene expression, and eventually affect individual phenotypes. In humans, DNA hypermethylation of *Alu* elements in introns of the proopiomelanocortin (*POMC*) gene causes downregulation of *POMC*, that is significantly associated with childhood obesity (*41*). In another case, DNA methylation of L1 in the intron of the *β-globin* gene (*HBB*) spreads into promoter and enhancer regions, that induce *β*-Thalassemia (*42*). In plants, several studies revealed the impact of epigenetic regulation of intronic TEs in gene function. In oil palm, methylation status of the LINE family retrotransposon Karma in the intron of the DEFICIENS gene affects splicing of the gene, where hypomethylation of Karma causes mis-splicing of the gene and results in the mantled phenotype of oil palm (*43*). Some plants require a cold period for transition from vegetative phase to reproductive phase, so called vernalization (*44*). In winter wheat, vernalization treatment induces DNA methylation at TEs in the intronic region of *VRN-A1*, which leads to transcriptional activation and promotion of flowering (*45*). In addition, population genome study of *A. thaliana* showed that the promoter and intron of the floral repressor gene *Flowering locus C* (*FLC*) are identified as a "hotspot" of TE insertion among natural accessions (*46*). Especially enriched are LTR retrotransposons, including *ONSEN/COPIA78*, and accessions having the TE insertion tend to show early flowering phenotypes, suggesting that insertion of the TE attenuates the function of *FLC*. How the epigenetic state of intronic TEs affects *FLC* expression in response to heat stress has yet to be investigated.

Conclusion

Epigenetic silencing of intragenic TEs is a two-edged sword. Aggressive TEs that transpose to transcription-permissive genic regions are a serious threat to gene function and genome integrity. On the other hand, silencing of intragenic TEs attenuates PolII transcription of the associated genes. Observed accumulation of heterochromatic TEs especially in intronic regions may be a result of a trade-off between TE silencing and

protection of gene function. In the case of plants, the heterochromatic state of intronic TEs enhances splicing/elongation of transcripts by specific molecular mechanisms. This would lead to further accumulation of TEs and expansion of intronic sequences in the genome, which would increase costs of maintenance and propagation. Although so far there are few reports that imply that epigenetic modification of intragenic TEs controls gene function, further studies would allow us to understand the impact of epigenetic regulation of intragenic TEs and their contributions to adaptation and evolution.

Acknowledgements

I thank Drs. Kenji Ichiyanagi, Kei Fukuda, and members of PEU for critical reading of the manuscript and comments. I also thank Dr. Steven D. Aird for editing of the manuscript. The author is grateful to OIST Graduate University for generous support of the Plant Epigenetics Unit.

Conflict of interest

The author declares no conflict of interest.

Legends to Figures

Figure 1. Transcriptional defects caused by insertion of transposable elements (TEs) into intronic regions of genes. (A) Cryptic polyA signals encoded by intronic TE cause premature polyadenylation. (B) Internal promoters encoded by TEs can create cryptic transcription start sites. (C) Incorporation of TE sequence as part of mature mRNAs, known as exonization, which changes coding sequences of genes. (D) Epigenetic silencing of intronic TE causes co-repression of associated genes. Black lollipops represent repressive epigenetic marks such as DNA methylation. Light blue boxes represent exon sequences. Red bars denote mRNAs. Arrows indicate transcription start sites.

Figure 2. Epigenetic regulation of intronic TEs in mammals and plants. (A) Epigenetic silencing of intronic L1 elements in mammals. It is suggested that transcription of associated genes can recruit the HUSH/MORC complex to L1 for silencing *(21)*. HUSH/MORC further recruits the H3K9 methylase SETDB1. Heterochromatic state of intronic TEs down-regulates associated genes. (B) Epigenetic silencing of intronic TEs in plants. siRNAs produced from homologous TEs at other loci directs RNA-directed DNA methylation (RdDM) to newly inserted TEs in intronic regions. This process is enhanced by genetic cross. Heterochromatic state of intronic TEs is maintained by DNA/histone methyltransferases such as MET1, SUVHs and chromatin remodeler DDM1, which enhances splicing of heterochromatic intron. IBM2 and EDM2 may bind to intronic heterochromatin via BAH and PHD domains and enhance splicing/elongation. Black lollipops represent repressive epigenetic marks such as DNA methylation. Light blue boxes represent exon sequences. Red bars denote mRNAs and siRNAs. Arrows indicate transcription start sites.

References

(1) Tenaillon, M.I., Hollister, J.D., and Gaut, B.S. (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci*. **15**, 471-478

(2) Kazazian, H.H., Jr. (2004) Mobile elements: drivers of genome evolution. *Science*. **303**, 1626-1632

(3) Cordaux, R., and Batzer, M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat Rev Genet*. **10**, 691-703

(4) Hancks, D.C., and Kazazian, H.H., Jr. (2012) Active human retrotransposons: variation and disease. *Curr Opin Genet Dev*. **22**, 191-203

(5) Schnable, P.S., and Ware, D., and Fulton, R.S., and Stein, J.C., and Wei, F., and Pasternak, S., and Liang, C., and Zhang, J., and Fulton, L., and Graves, T.A., and Minx, P., and Reily, A.D., and Courtney, L., and Kruchowski, S.S., and Tomlinson, C., and Strong, C., and Delehaunty, K., and Fronick, C., and Courtney, B., and Rock, S.M., and Belter, E., and Du, F., and Kim, K., and Abbott, R.M., and Cotton, M., and Levy, A., and Marchetto, P., and Ochoa, K., and Jackson, S.M., and Gillam, B., and Chen, W., and Yan, L., and Higginbotham, J., and Cardenas, M., and Waligorski, J., and Applebaum, E., and Phelps, L., and Falcone, J., and Kanchi, K., and Thane, T., and Scimone, A., and Thane, N., and Henke, J., and Wang, T., and Ruppert, J., and Shah, N., and Rotter, K., and Hodges, J., and Ingenthron, E., and Cordes, M., and Kohlberg, S., and Sgro, J., and Delgado, B., and Mead, K., and Chinwalla, A., and Leonard, S., and Crouse, K., and Collura, K., and Kudrna, D., and Currie, J., and He, R., and Angelova, A., and Rajasekar, S., and Mueller, T., and Lomeli, R., and Scara, G., and Ko, A., and Delaney, K., and Wissotski, M., and Lopez, G., and Campos, D., and Braidotti, M., and Ashley, E., and Golser, W., and Kim, H., and Lee, S., and Lin, J., and Dujmic, Z., and Kim, W., and Talag, J., and Zuccolo, A., and Fan, C., and Sebastian, A., and Kramer, M., and Spiegel, L., and Nascimento, L., and Zutavern, T., and Miller, B., and Ambroise, C., and Muller, S., and Spooner, W., and Narechania, A., and Ren, L., and Wei,

S., and Kumari, S., and Faga, B., and Levy, M.J., and McMahan, L., and Van Buren, P., and Vaughn, M.W., and Ying, K., and Yeh, C.T., and Emrich, S.J., and Jia, Y., and Kalyanaraman, A., and Hsia, A.P., and Barbazuk, W.B., and Baucom, R.S., and Brutnell, T.P., and Carpita, N.C., and Chaparro, C., and Chia, J.M., and Deragon, J.M., and Estill, J.C., and Fu, Y., and Jeddeloh, J.A., and Han, Y., and Lee, H., and Li, P., and Lisch, D.R., and Liu, S., and Liu, Z., and Nagel, D.H., and McCann, M.C., and SanMiguel, P., and Myers, A.M., and Nettleton, D., and Nguyen, J., and Penning, B.W., and Ponnala, L., and Schneider, K.L., and Schwartz, D.C., and Sharma, A., and Soderlund, C., and Springer, N.M., and Sun, Q., and Wang, H., and Waterman, M., and Westerman, R., and Wolfgruber, T.K., and Yang, L., and Yu, Y., and Zhang, L., and Zhou, S., and Zhu, Q., and Bennetzen, J.L., and Dawe, R.K., and Jiang, J., and Jiang, N., and Presting, G.G., and Wessler, S.R., and Aluru, S., and Martienssen, R.A., and Clifton, S.W., and McCombie, W.R., and Wing, R.A., and Wilson, R.K. (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*. **326**, 1112-1115

(6) Heard, E., and Martienssen, R.A. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*. **157**, 95-109

(7) Slotkin, R.K., and Martienssen, R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*. **8**, 272-285

(8) Sela, N., Mersch, B., Gal-Mark, N., Lev-Maor, G., Hotz-Wagenblatt, A., and Ast, G. (2007) Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome. *Genome Biol*. **8**, R127

(9) Li, J., Akagi, K., Hu, Y., Trivett, A.L., Hlynialuk, C.J., Swing, D.A., Volfovsky, N., Morgan, T.C., Golubeva, Y., Stephens, R.M., Smith, D.E., and Symer, D.E. (2012) Mouse endogenous retroviruses can trigger premature transcriptional termination at a distance. *Genome Res*. **22**, 870-884

(10) Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*. **13**, 335-340

(11) Schmitz, J., and Brosius, J. (2011) Exonization of transposed elements: A challenge and opportunity for evolution. *Biochimie*. **93**, 1928-1934

(12) Sela, N., Mersch, B., Hotz-Wagenblatt, A., and Ast, G. (2010) Characteristics of transposable element exonization within human and mouse. *PLoS One*. **5**, e10907

(13) Zarnack, K., Konig, J., Tajnik, M., Martincorena, I., Eustermann, S., Stevant, I., Reyes, A., Anders, S., Luscombe, N.M., and Ule, J. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell*. **152**, 453-466

(14) West, P.T., Li, Q., Ji, L., Eichten, S.R., Song, J., Vaughn, M.W., Schmitz, R.J., and Springer, N.M. (2014) Genomic distribution of H3K9me2 and DNA methylation in a maize genome. *PLoS One*. **9**, e105267

(15) Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.C., Scofield, D.G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hallman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Kaller, M., Luthman, J., Lysholm, F., Niittyla, T., Olson, A., Rilakovic, N., Ritland, C., Rossello, J.A., Sena, J., Svensson, T., Talavera-Lopez, C., Theissen, G., Tuominen, H., Vanneste, K., Wu, Z.Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Garcia Gil, R., Hvidsten, T.R., de Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Thompson, S.L., Van de Peer, Y., Andersson, B., Nilsson, O., Ingvarsson, P.K., Lundeberg, J., and Jansson, S. (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*. **497**, 579-584

(16) Le, T.N., Miyazaki, Y., Takuno, S., and Saze, H. (2015) Epigenetic regulation of intragenic transposable elements impacts gene transcription in Arabidopsis thaliana. *Nucleic Acids Res*. **43**, 3911-3921

(17) Lockton, S., and Gaut, B.S. (2009) The contribution of transposable elements to expressed coding sequence in Arabidopsis thaliana. *J Mol Evol*. **68**, 80-89

(18) Hollister, J.D., Smith, L.M., Guo, Y.L., Ott, F., Weigel, D., and Gaut, B.S.

(2011) Transposable elements and small RNAs contribute to gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. *Proc Natl Acad Sci U S A*. **108**, 2322-2327

(19) Bennetzen, J.L., and Wang, H. (2014) The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annu Rev Plant Biol*. **65**, 505-530

(20) To, T.K., Saze, H., and Kakutani, T. (2015) DNA Methylation within Transcribed Regions. *Plant Physiol*. **168**, 1219-1225

(21) Liu, N., Lee, C.H., Swigut, T., Grow, E., Gu, B., Bassik, M.C., and Wysocka, J. (2018) Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature*. **553**, 228-232

(22) Robbez-Masson, L., Tie, C.H.C., Conde, L., Tunbak, H., Husovsky, C., Tchasovnikarova, I.A., Timms, R.T., Herrero, J., Lehner, P.J., and Rowe, H.M. (2018) The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res*. **28**, 836-845

(23) Tchasovnikarova, I.A., Timms, R.T., Matheson, N.J., Wals, K., Antrobus, R., Gottgens, B., Dougan, G., Dawson, M.A., and Lehner, P.J. (2015) GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science*. **348**, 1481-1485

(24) Fukuda, K., Okuda, A., Yusa, K., and Shinkai, Y. (2018) A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells. *Genome Res*. **28**, 846-858

(25) Martienssen, R., and Moazed, D. (2015) RNAi and heterochromatin assembly. *Cold Spring Harb Perspect Biol*. **7**, a019323

(26) Law, J.A., and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. **11**, 204-220

(27) Tsuchiya, T., and Eulgem, T. (2013) An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proc Natl Acad Sci U S A*. **110**, E3535-3543

(28) Matzke, M.A., and Mosher, R.A. (2014) RNA-directed DNA methylation:

an epigenetic pathway of increasing complexity. *Nat Rev Genet*. **15**, 394-408

(29) Ito, H., Kim, J.M., Matsunaga, W., Saze, H., Matsui, A., Endo, T.A., Harukawa, Y., Takagi, H., Yaegashi, H., Masuta, Y., Masuda, S., Ishida, J., Tanaka, M., Takahashi, S., Morosawa, T., Toyoda, T., Kakutani, T., Kato, A., and Seki, M. (2016) A Stress-Activated Transposon in Arabidopsis Induces Transgenerational Abscisic Acid Insensitivity. *Sci Rep*. **6**, 23181

(30) Osabe, K., Harukawa, Y., Miura, S., and Saze, H. (2017) Epigenetic Regulation of Intronic Transgenes in Arabidopsis. *Sci Rep*. **7**, 45166

(31) Saint-Andre, V., Batsche, E., Rachez, C., and Muchardt, C. (2011) Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol*. **18**, 337-344

(32) Lev Maor, G., Yearim, A., and Ast, G. (2015) The alternative role of DNA methylation in splicing regulation. *Trends Genet*. **31**, 274-280

(33) Hollister, J.D., and Gaut, B.S. (2009) Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res*. **19**, 1419-1428

(34) Wang, X., Weigel, D., and Smith, L.M. (2013) Transposon variants and their effects on gene expression in Arabidopsis. *PLoS Genet*. **9**, e1003255

(35) Eulgem, T., Tsuchiya, T., Wang, X.J., Beasley, B., Cuzick, A., Tor, M., Zhu, T., McDowell, J.M., Holub, E., and Dangl, J.L. (2007) EDM2 is required for RPP7-dependent disease resistance in Arabidopsis and affects RPP7 transcript levels. *Plant J*. **49**, 829-839

(36) Lei, M., La, H., Lu, K., Wang, P., Miki, D., Ren, Z., Duan, C.G., Wang, X., Tang, K., Zeng, L., Yang, L., Zhang, H., Nie, W., Liu, P., Zhou, J., Liu, R., Zhong, Y., Liu, D., and Zhu, J.K. (2014) Arabidopsis EDM2 promotes IBM1 distal polyadenylation and regulates genome DNA methylation patterns. *Proc Natl Acad Sci U S A*. **111**, 527-532

(37) Saze, H., Kitayama, J., Takashima, K., Miura, S., Harukawa, Y., Ito, T., and Kakutani, T. (2013) Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. *Nat Commun*. **4**, 2301

(38) Wang, X., Duan, C.G., Tang, K., Wang, B., Zhang, H., Lei, M., Lu, K., Mangrauthia, S.K., Wang, P., Zhu, G., Zhao, Y., and Zhu, J.K. (2013) RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intronic heterochromatin-containing gene IBM1. *Proc Natl Acad Sci U S A*. **110**, 15467-15472

(39) Coustham, V., Vlad, D., Deremetz, A., Gy, I., Cubillos, F.A., Kerdaffrec, E., Loudet, O., and Bouche, N. (2014) SHOOT GROWTH1 maintains Arabidopsis epigenomes by regulating IBM1. *PLoS One*. **9**, e84687

(40) Duan, C.G., Wang, X., Zhang, L., Xiong, X., Zhang, Z., Tang, K., Pan, L., Hsu, C.C., Xu, H., Tao, W.A., Zhang, H., and Zhu, J.K. (2017) A protein complex regulates RNA processing of intronic heterochromatin-containing genes in Arabidopsis. *Proc Natl Acad Sci U S A*. **114**, E7377-E7384

(41) Kuehnen, P., Mischke, M., Wiegand, S., Sers, C., Horsthemke, B., Lau, S., Keil, T., Lee, Y.A., Grueters, A., and Krude, H. (2012) An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet*. **8**, e1002543

(42) Lanikova, L., Kucerova, J., Indrak, K., Divoka, M., Issa, J.P., Papayannopoulou, T., Prchal, J.T., and Divoky, V. (2013) beta-Thalassemia due to intronic LINE-1 insertion in the beta-globin gene (HBB): molecular mechanisms underlying reduced transcript levels of the beta-globin(L1) allele. *Hum Mutat*. **34**, 1361-1365

(43) Ong-Abdullah, M., Ordway, J.M., Jiang, N., Ooi, S.E., Kok, S.Y., Sarpan, N., Azimi, N., Hashim, A.T., Ishak, Z., Rosli, S.K., Malike, F.A., Bakar, N.A., Marjuni, M., Abdullah, N., Yaakub, Z., Amiruddin, M.D., Nookiah, R., Singh, R., Low, E.T., Chan, K.L., Azizi, N., Smith, S.W., Bacher, B., Budiman, M.A., Van Brunt, A., Wischmeyer, C., Beil, M., Hogan, M., Lakey, N., Lim, C.C., Arulandoo, X., Wong, C.K., Choo, C.N., Wong, W.C., Kwan, Y.Y., Alwee, S.S., Sambanthamurthi, R., and Martienssen, R.A. (2015) Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*. **525**, 533-537

(44) Baulcombe, D.C., and Dean, C. (2014) Epigenetic regulation in plant responses to the environment. *Cold Spring Harb Perspect Biol*. **6**, a019471

(45) Khan, A.R., Enjalbert, J., Marsollier, A.C., Rousselet, A., Goldringer, I., and Vitte, C. (2013) Vernalization treatment induces site-specific DNA hypermethylation at the VERNALIZATION-A1 (VRN-A1) locus in hexaploid winter wheat. *BMC Plant Biol.* **13**, 209

(46) Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., LeBlanc, C., Martienssen, R.A., Jeddeloh, J.A., and Colot, V. (2016) The Arabidopsis thaliana mobilome and its impact at the species level. *Elife.* **5**,

Figure 1

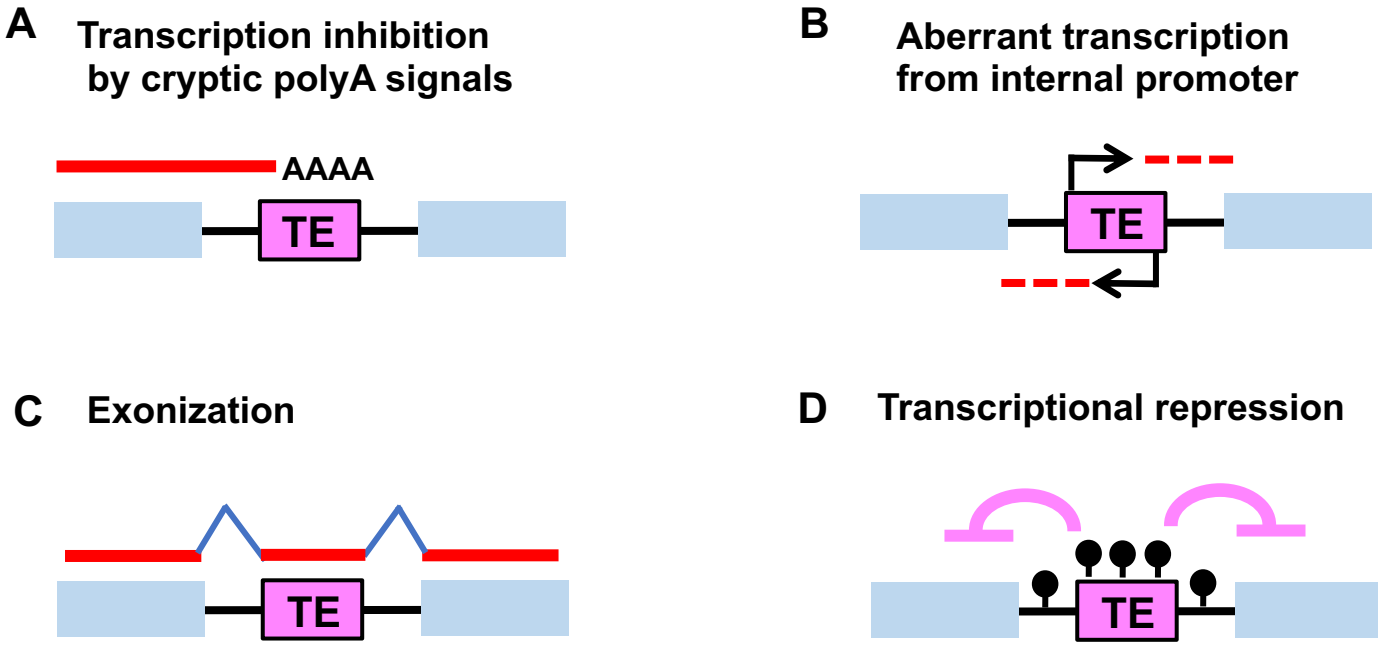
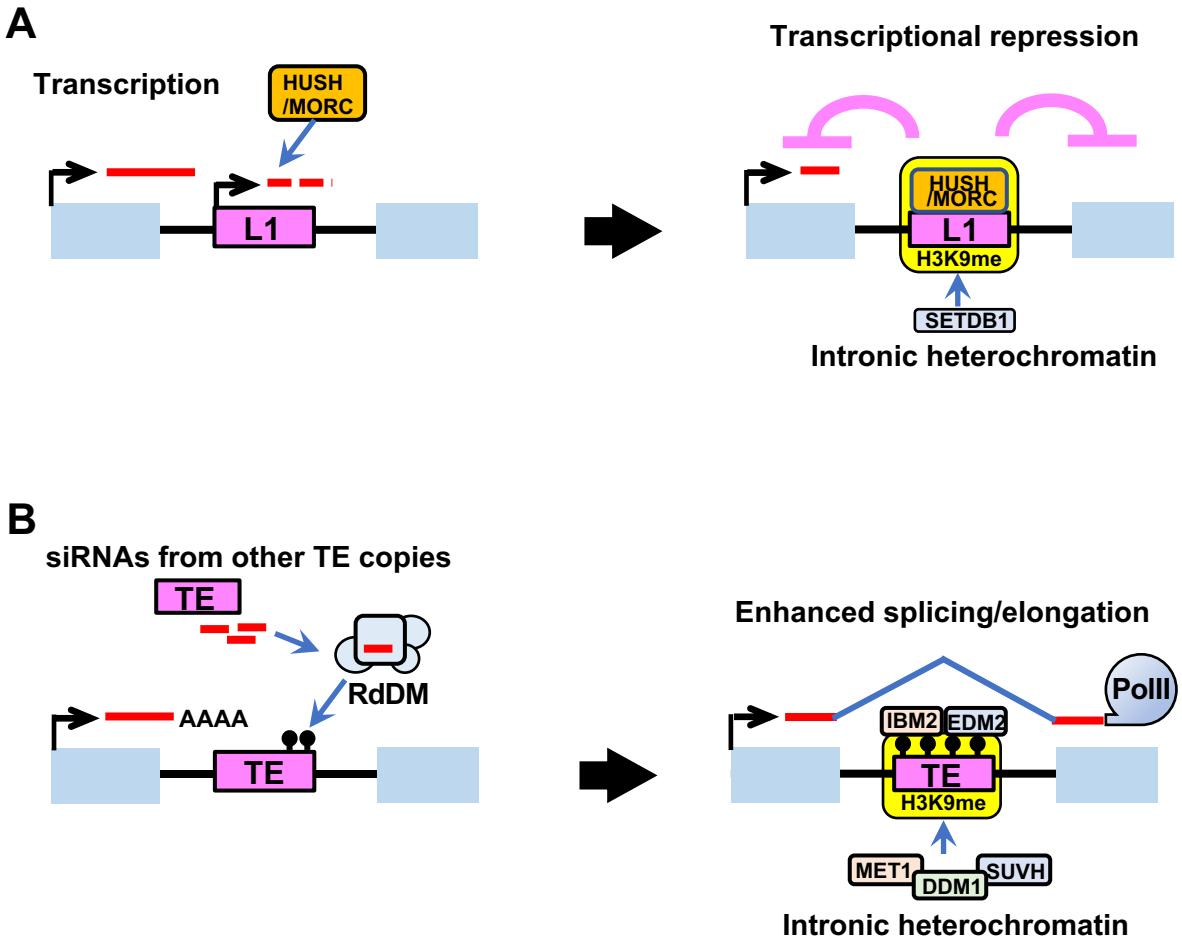

A
Transcription inhibition
by cryptic polyA signals
AAAA
TE
B
Aberrant transcription
from internal promoter
TE
C
Exonization
TE
D
Transcriptional repression
TE

Figure 2


A
Transcription
HUSH
/MORC
L1
Transcriptional repression
HUSH
/MORC
L1
H3K9me
SETDB1
Intronic heterochromatin
B
siRNAs from other TE copies
TE
RdDM
AAAA
TE
Enhanced splicing/elongation
PolII
IBM2
EDM2
TE
H3K9me
MET1
DDM1
SUVH
Intronic heterochromatin