Published in final edited form as:

Curr Opin Plant Biol. 2017 April ; 36: 135–141. doi:10.1016/j.pbi.2017.03.002.

Genetic and epigenetic variation of transposable elements in *Arabidopsis*

Charles J. Underwood^{1,3}, Ian R. Henderson³, and Robert A. Martienssen^{1,2}

¹Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

²HHMI-GBMF, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

³Department of Plant Sciences, Downing Street, University of Cambridge, Cambridge, UK

Abstract

Transposable elements are mobile genetic elements that are prevalent in plant genomes and are silenced by epigenetic modification. Different epigenetic modification pathways play distinct roles in the control of transposable element transcription, replication and recombination. The *Arabidopsis* genome contains families of all of the major transposable element classes, which are differentially enriched in particular genomic regions. Whole genome sequencing and DNA methylation profiling of hundreds of natural *Arabidopsis* accessions has revealed that transposable elements exhibit significant intraspecific genetic and epigenetic variation, and that genetic variation often underlies epigenetic variation.

Together, epigenetic modification and the forces of selection define the scope within which transposable elements can contribute to, and control, genome evolution.

Introduction

Transposable elements (TEs) can drive evolution by controlling gene expression and through mutagenic transposition [1,2]. TEs were first described in maize by Barbara McClintock who named them controlling elements, because of their capacity to regulate gene expression [3]. TEs can be split into retrotransposons (Class I) which replicate via RNA and cDNA intermediates, and DNA transposons (Class II) which replicate through a DNA intermediate [1]. TEs, and TE classes, vary in abundance between genomes, with TEs making up 75% of the maize genome, 40% of the human genome and 20% of the *Arabidopsis thaliana* genome [4]. Generally, transposon frequency positively correlates with genome size within eukaryotes [5]. The factors governing transposon abundance are complex and include frequency of transcription and transposition, insertion preferences, positive and negative selection and population size [1,5].

Studies in *Arabidopsis thaliana* have made significant contributions to our understanding of the nature and control of TEs in plant genomes [4]. It has long been appreciated that TEs and genes show marked epigenetic differentiation in plant genomes [6], and various epigenetic modifications contribute to control of TE transcription and transposition (Figure 1). RNA interference (RNAi) leads to the accumulation of 21–24 nucleotide homologous

small interfering RNAs (siRNAs) at plant TE 5' and 3' repeats which play a role in establishing and maintaining DNA cytosine methylation (Figure 1) [7–10]. DNA cytosine methylation of TEs occurs in CG, CHG and CHH (where H = A, C or T) sequence contexts, and the nature of bases represented by H can influence methylation frequency [11,12]. CHG and CHH methylation are largely interdependent with dimethylation of histone H3 on lysine 9 (H3K9me2) (Figure 1) [13]. Plant TEs are also enriched for monomethylation of histone H3 on lysine 27 (H3K27me1), and the histone variants H3.1 and H2A.W (Figure 1) [14–16]. Despite having a 125 megabase genome, Arabidopsis contains large families of the major TE classes (Figure 2) [17–19]. Here we discuss the significant diversity exhibited between Arabidopsis TE families, with regards to genomic distributions and variation between natural accessions. We also explore how epigenetic modification of Arabidopsis TEs influences their transcription, transposition, replication and recombination.

TE families have diverse genomic distributions

Retrotransposons and DNA transposons exist as autonomous or non-autonomous elements [1,4]. Autonomous elements encode all of the proteins required for their transposition. In contrast, non-autonomous elements possess *cis* sequences required for transposition, but require a cognate autonomous element to provide transposition factors in *trans*. For transposition, DNA transposons do not necessarily require transcription of the transposing element [1]. However, individual retrotransposons must be transcribed in order for transposition to occur via RNA (copy) and cDNA (paste) intermediates [1].

Retrotransposons can be further divided into those that are flanked by long terminal repeats (LTR retrotransposons) and non-LTR retrotransposons [4]. In Arabidopsis, the *GYPSY* and *COPIA* families are the predominant LTR retrotransposons, and *LINE* elements are the predominant non-LTR retrotransposons [17,18]. Interestingly, these three retrotransposon families show distinct chromosomal distributions (Figure 2). Arabidopsis centromeres, which remain to be fully assembled, consist of megabase tandem arrays of the *CEN180* repeat within which the kinetochore binds [17,20]. *GYPSY* elements are most dense in proximity to the centromeres, whereas *COPIA* and *LINE* elements show greater pericentromeric accumulation (Figure 2), although the forces shaping these distributions are unclear. Interestingly, the introduction of the *Transposon of Arabidopsis lyrata 1 (Tall)* *COPIA* element from *Arabidopsis lyrata* into *A.thaliana* showed preferential integration into centromeric repeats [21]. However, a transposition burst of various *COPIA* elements, including *ATCOPIA93/EVADE*, in Arabidopsis DNA hypomethylated mutants showed no preference for centromeric regions [22,23]. Hence, retrotransposons may or may not preferentially integrate into centromeres. In maize and wheat centromeric enrichment of specific retrotransposons is also observed [24,25], and retrotransposons show a high degree of nested insertions in maize [26]. The accumulation of transposons in centromeric regions may arise due to such insertions being more likely to be selectively neutral.

The DNA transposons in the Arabidopsis genome can be broadly split into 6 classes; *MuDR*, *EnSpm/CACTA*, *hAT*, *PIF/HARBINGER*, the related *POGO*, *Tc1* and *MARINER* elements, and rolling circle replicating *HELITRON* elements (Figure 2). The first five classes belong to families that transpose using DDE class nucleases which interact with terminal inverted

repeats (TIRs) [1,27]. Several DNA TE families encode accessory proteins in addition to the transposase which can play important roles regulating activity. For example, VANC of the Arabidopsis *VANDAL21 MuDR* family exhibits anti-silencing functions [28]. Overall DNA elements show greater pericentromeric and chromosome arm abundance compared with RNA elements, which is also evident in the wheat genome (Figure 2) [17,25]. Of the Arabidopsis DNA elements the *EnSpm/CACTA* family shows greatest centromeric enrichment (Figure 2). Transposition and mutagenesis by the *CAC1 EnSpm* element has been observed in *ddm1* [29]. Many DNA elements, for example *HELITRON* and *MuDR/MUTATOR* elements, show high abundance in proximity to genes [30–32]. In the case of *HELITRONs* this is partially driven by the insertion preference of *HELITRONs* for TA dinucleotides [1,31]. While *MUTATOR* element abundance, gene density and levels of meiotic recombination correlate at the broad scale along maize chromosomes [32].

Variation in TEs between natural Arabidopsis accessions

The study of natural Arabidopsis accessions has revealed insights into genetic and epigenetic TE polymorphisms, and variation in TE copy number and sequence [19,33–35]. Furthermore, the identification of novel TE insertion sites, compared to the Col-0 reference, in hundreds of natural accessions has identified *COPIA* and *MuDR* elements as the most active TEs in the recent history of the species [19]. Since their divergence, any two Arabidopsis accessions are predicted to have between 200 and 300 new TE insertions, and the flowering time regulator *FLC* and various NLR immunity genes were identified to be prone to TE insertion and haplotype maintenance [19]. Arabidopsis accessions show significant variation in DNA methylation, which includes a geographical component [35–37]. TE epigenetic variation may be influenced by genetic control by *trans* acting factors, due to genetic variation in *cis* within the element or due to epiallelic variation *in cis* [19,35,36]. Natural genetic variation in the DNA methyltransferase gene *CMT2*, and the RNAi genes *AGO1*, *AGO9* and *NRPD1B* associate with variation in TE DNA methylation [35,36], demonstrating that genetic variation in *trans* can underlie TE epigenetic variation. Genetic control in *cis* is also a major contributor to TE epigenetic variation in Arabidopsis and maize [37,38]. Epiallelic variation of DNA methylation has also been observed [39].

Upon comparison between accessions, the likelihood of genetic deterioration of TEs by insertions and/or deletions is variable between TE classes, with deletions in Class I *GYPSY* elements occurring at a much greater rate than expected by chance [34]. Interestingly, both deletions and insertions within *HELITRON* elements occur at a lower frequency than expected, suggesting that selection may act to maintain these elements [34]. Importantly, *HELITRON* and *MuDR* elements were also identified as associating with genetic and epigenetic variation between accessions, especially in proximity to disease resistance NBS-LRR genes [35]. In addition, DNA demethylation of repeats present in immune response gene promoters have been linked to innate immune response [40]. Together this suggests that TEs may play a key role in shaping adaptive variation in Arabidopsis.

Variation in epigenetic control of TE expression, transposition, replication and meiotic recombination

Expression and Transposition

The critical role of heterochromatic epigenetic modifications in suppressing transposon expression and transposition has been revealed using *Arabidopsis* mutants [13,29,30,41]. Mutations that affect maintenance of CG methylation (*met1*), CG and partially non-CG (*ddm1*), CHG (*cmt3*), CHG/CHH (*drm1 drm2 cmt2 cmt3*) and H3K27me1 (*atxr5 atxr6*) transcriptionally upregulate distinct groups of transposons [13,41]. Loss of CG methylation maintenance in the *met1* and *ddm1* mutants leads to the largest increase in transposon expression [41]. However, transposition is only detectable for a small number of TEs in *ddm1*, including some *MUTATOR* and *CACTA* DNA transposons [29,30], and retrotransposons from specific Gypsy (*ATGP3*) and Copia (*ATCOPIA13*, *ATCOPIA21*, *ATCOPIA93*) subfamilies [22]. *ATCOPIA93/EVADE* transposition is also observed in *met1*-derived epiRILS [23]. In some cases, abundant TE transcription is observed, but with few transposition events, which may be due to compensatory RNAi mechanisms, for example in the form of 'epigenetically activated' siRNAs (easiRNAs) [10]. Interestingly the *ONSEN* retrotransposon, which can transpose upon combined heat exposure and loss of RNAi, does not jump in *ddm1* mutants and is directly controlled by a heat responsive element [42,43]. For both *ONSEN* and epigenetically activated TEs, RNAi likely plays a key role in post-transcriptional and/or translation control of TE-derived transcripts. Given the recent annotation and study of TE differences between natural *Arabidopsis* accessions [19,34–36], an obvious next step will be the perturbation of epigenetic and RNAi pathways in these accessions and the study of differential TE mobilization.

It is important to note that TEs exhibit extreme genetic heterogeneity both within and between TE classes, and therefore the basis of epigenetic control is complex. Current models of TE silencing are largely based on the study of either a particular TE/repeat locus, or an average of thousands of unrelated TEs in the reference Col-0 accession. Differences in TE class, structure (internal or nested repeats), AT:GC content, density of CG, CHG and CHH sequence contexts, nucleosome occupancy, and proximity to flanking genes/TEs, all likely affect the importance of different epigenetic pathways in the silencing of a given TE. These various TE characteristics likely interact to contribute to the differential sensitivity of TEs to different epigenetic perturbations. The current 'generic' models are likely inadequate, which is exemplified by the recent discovery that polymorphism in *MET2a*, a gene with no previous connection to TE silencing, correlates with the activity of seven TE families in natural *Arabidopsis* accessions [19].

DNA replication

In *Arabidopsis* and many other species, heterochromatin replicates late in S phase after euchromatin, and this is likely to involve epigenetic control [44,45]. Importantly, most DNA transposons replicate by cut-and-paste mechanisms, and can therefore only increase in copy number by transposing during S-phase after their own replication into yet to be replicated genomic regions [1,4]. Recent work in maize has demonstrated that *Ac* transposition-mediated DNA re-replication can also lead to the increased copy number of flanking regions

[46]. Therefore delaying the replication of DNA transposons until late S-phase could help to prevent copy number increases.

In Arabidopsis, H3K27me1, mediated by ATXR5 and ATXR6, is thought to play a key role in the control of the replication of TEs and repeated sequences [14]. H3K27me1 is enriched over TEs (Figure 1), and *atxr5 atxr6* double mutants exhibit over replication and loss of transcriptional silencing of specific TE [14,41]. The repetitive DNA that over-replicates in *atxr5 atxr6* is H3-associated and DNA methylated, suggesting it is associated with chromatin and not extrachromosomal [14]. Heterochromatin is decondensed in *atxr5 atxr6* double mutants and the over-replicated regions associate with the DNA double strand break markers RAD51 and gamma-H2AX [47]. ATXR5 and ATXR6 likely associate with the replication fork and are only capable of modifying the replication dependent histone variant H3.1, due to an amino acid difference between H3.1 and H3.3 at residue 31 [48]. Reverse and forward genetic studies have identified a plethora of suppressors of TE over-replication in *atxr5 atxr6*, which have differential effects on the suppression of loss transcriptional silencing [41,48,49] It remains to be seen whether H3K27me1 controls heterochromatic DNA replication directly or indirectly, and also if H3K27me1 plays a role in delaying heterochromatic replication to late S-phase.

Meiotic recombination

Meiotic recombination frequency varies dramatically along plant chromosomes, and at the broad scale is enriched in gene-dense euchromatin and suppressed within centromeres [20,25]. Due to the enrichment of many TEs in proximity to the centromeres, it may be assumed that TEs would be suppressed for meiotic recombination (Figure 2). Indeed recombination between repeated sequences poses a risk for non-allelic recombination and the potential for structural change. Consistent with this crossover hotspots are enriched at gene promoters in Arabidopsis and Tomato [50–52], and artificial introduction of DNA methylation to a known hotspot is sufficient to silence crossover [53]. In Arabidopsis DNA methylation in CG contexts, which is enriched in pericentromeric and centromeric regions, is almost completely lost in *met1* and *ddm1* [9]. Counter intuitively *met1* and *ddm1* mutants have reduced pericentromeric crossovers, and increased euchromatic crossovers [54–57]. It will be interesting to explore the role of other heterochromatic epigenetic modifications to understand their role in pericentromeric crossover suppression.

Despite the suppression of centromeric meiotic recombination, there are a number of positive correlations between meiotic recombination and transposons that suggest a more complex relationship. First, meiotic gene conversion, but not crossover, is observed in maize centromeric *CRM* retroelements, suggesting meiotic DSBs can initiate within repeats but may be repaired by non-crossover pathways [58]. Second, *MUTATOR* elements positively correlate with meiotic crossover frequency in maize [32], and the presence of a *MUTATOR* element was observed to increase crossover frequency at the *a1* hotspot when active MU transposase was also present [59]. In contrast, active *Ac* does not stimulate recombination in Maize meiosis at the *bx* locus, but can stimulate mitotic recombination at the *p1* locus [60,61]. These observations are further consistent with functional diversity between TE elements and families, in this case at the level of their effects on meiotic recombination.

Further work will be required to assess the relationship between specific TE families and meiotic recombination, and also to address the genetic and epigenetic contributions to suppression of recombination in TE-rich regions.

Conclusions and future directions

We have reviewed emerging data that reveals diversity of TEs between Arabidopsis accessions, and how epigenetic pathways control TE expression, replication and recombination. Indeed, it is plausible that genetic and epigenetic variation of TEs between crop varieties may underlie important agronomic traits. For instance, in oil palm DNA hypomethylation of an intronic *LINE* retrotransposon during tissue culture leads to palm mantling and associated drastic yield reduction [62]. Transposition events can also lead to structural rearrangements including inversions and segmental duplications which can have phenotypic effects [63,64]. The advent of economical long-read sequencing technology will make complete genomic assemblies, and the full assessment of structural differences, a reality, even in crops with large TE complements and genome sizes. This will facilitate detailed study of TEs in all plants and allow us to further understand how these controlling elements contribute to the evolution and function of plant genomes.

Acknowledgments

This work was supported by a William Miller Fellowship from the Watson School of Biological Sciences (CJU), the Gatsby Charitable Foundation and the Royal Society (IRH), the Howard Hughes Medical Institute-Gordon and Betty Moore Foundation (RAM) and by grants from the NSF (DBI-1025830) and the NIH (GM067014) (RAM).

References

1. Lambowitz, AM. Gellert, M. Chandler, M. Craig, NL. Sandmeyer, SB., Rice, PA., editors. Mobile DNA III [Internet]. American Society of Microbiology; 2015.
2. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits [Internet]. Nat Rev Genet. 2016; 18:71–86. [PubMed: 27867194]
3. McClintock B. Controlling Elements and the Gene [Internet]. Cold Spring Harb Symp Quant Biol. 1956; 21:197–216. [PubMed: 13433592]
4. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. [Internet]. Nat Rev Genet. 2007; 8:272–85. [PubMed: 17363976]
5. Lynch, M. The origins of genome architecture. Sinauer Associates; 2007.
6. Rabinowicz PD, Schutz K, Dedhia N, Yordan C, Parnell LD, Stein L, McCombie WR, Martienssen RA. Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. [Internet]. Nat Genet. 1999; 23:305–8. [PubMed: 10545948]
7. Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell. 2013; 153:193–205. [PubMed: 23540698]
8. Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. Reconstructing de novo silencing of an active plant retrotransposon. [Internet]. Nat Genet. 2013; 45:1029–39. [PubMed: 23852169]
9. Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome [Internet]. Cell. 2013; 152:352–64. [PubMed: 23313553]
10. Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, Martienssen RA. miRNAs trigger widespread epigenetically activated siRNAs from transposons in Arabidopsis [Internet]. Nature. 2014; 508:411–5. [PubMed: 24670663]

11. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis* [Internet]. *Cell*. 2008; 133:523–36. [PubMed: 18423832]
- 12*. Gouil Q, Baulcombe DC. DNA Methylation Signatures of the Plant Chromomethyltransferases. [Internet]. *PLoS Genet*. 2016; 12:e1006526. This paper demonstrates that the CHG and CHH classifications of DNA methylation hide functional information about plant DNA methylation pathways. [PubMed: 27997534]
13. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. [Internet]. *Nat Struct Mol Biol*. 2014; 21:64–72. [PubMed: 24336224]
14. Jacob Y, Stroud H, Leblanc C, Feng S, Zhuo L, Caro E, Hassel C, Gutierrez C, Michaels SD, Jacobsen SE. Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases [Internet]. *Nature*. 2010; 466:987–91. [PubMed: 20631708]
15. Stroud H, Otero S, Desvoves B, Ramírez-Parra E, Jacobsen SE, Gutierrez C. Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. [Internet]. *Proc Natl Acad Sci U S A*. 2012; 109:5370–5. [PubMed: 22431625]
16. Yelagandula R, Stroud H, Holec S, Zhou K, Feng S, Zhong X, Muthurajan UM, Nie X, Kawashima T, Groth M, et al. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis* [Internet]. *Cell*. 2014; 158:98–109. [PubMed: 24995981]
17. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. [Internet]. *Nature*. 2000; 408:796–815. [PubMed: 11130711]
18. Buisine N, Quesneville H, Colot V. Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. *Genomics*. 2008; 91:467–75. [PubMed: 18343092]
- 19*. Quadrana L, Silveira AB, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V. The *Arabidopsis thaliana* mobilome and its impact at the species level [Internet]. *Elife*. 2016; 5:6919–6931. This paper identifies the most active TEs in the recent history of *Arabidopsis* and shows that genetic variation in *MET2* associates with TE activity.
20. Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S, Lin X, Bevan M, Murphy G, Harris B, Parnell LD, et al. Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science*. 1999; 286:2468–74. [PubMed: 10617454]
21. Tsukahara S, Kawabe A, Kobayashi A, Ito T, Aizu T, Shin-i T, Toyoda A, Fujiyama A, Tarutani Y, Kakutani T. Centromere-targeted de novo integrations of an LTR retrotransposon of *Arabidopsis lyrata* [Internet]. *Genes Dev*. 2012; 26:705–713. [PubMed: 22431508]
22. Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of retrotransposition reproduced in *Arabidopsis* [Internet]. *Nature*. 2009; 461:423–426. [PubMed: 19734880]
23. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O. Selective epigenetic control of retrotransposition in *Arabidopsis* [Internet]. *Nature*. 2009; 461:427–430. [PubMed: 19734882]
24. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. The B73 maize genome: complexity, diversity, and dynamics. *Science*. 2009; 326:1112–5. [PubMed: 19965430]
25. Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, et al. Structural and functional partitioning of bread wheat chromosome 3B [Internet]. *Science*. 2014; 345:1249721. [PubMed: 25035497]
26. San Miguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, et al. Nested retrotransposons in the intergenic regions of the maize genome [Internet]. *Science*. 1996; 274:765–8. [PubMed: 8864112]
27. Feschotte C, Pritham EJ. DNA Transposons and the Evolution of Eukaryotic Genomes [Internet]. *Annu Rev Genet*. 2007; 41:331–368. [PubMed: 18076328]
28. Fu Y, Kawabe A, Etcheverry M, Ito T, Toyoda A, Fujiyama A, Colot V, Tarutani Y, Kakutani T. Mobilization of a plant transposon by expression of the transposon-encoded anti-silencing factor. [Internet]. *EMBO J*. 2013; 32:2407–17. [PubMed: 23900287]

29. Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis [Internet]. *Nature*. 2001; 411:212–4. [PubMed: 11346800]
30. Singer T, Yordan C, Martienssen RA. Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene decrease in DNA Methylation (DDM1). [Internet]. *Genes Dev*. 2001; 15:591–602. [PubMed: 11238379]
31. Kapitonov VV, Jurka J. Rolling-circle transposons in eukaryotes. [Internet]. *Proc Natl Acad Sci U S A*. 2001; 98:8714–9. [PubMed: 11447285]
32. Liu S, Yeh C-T, Ji T, Ying K, Wu H, Tang HM, Fu Y, Nettleton D, Schnable PS. Mu Transposon Insertion Sites and Meiotic Recombination Events Co-Localize with Epigenetic Marks for Open Chromatin across the Maize Genome [Internet]. *PLoS Genet*. 2009; 5:e1000733. [PubMed: 19936291]
33. Wang X, Weigel D, Smith LM, Lander E, Linton L, Birren B, Nusbaum C, Zody M, Messing J, Bharti A, et al. Transposon Variants and Their Effects on Gene Expression in Arabidopsis [Internet]. *PLoS Genet*. 2013; 9:e1003255. [PubMed: 23408902]
34. Stuart T, Eichten SR, Cahn J, Karpievitch YV, Borevitz JO, Lister R. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation [Internet]. *Elife*. 2016;5.
35. Kawakatsu T, Huang S-SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He Y, et al. Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions [Internet]. *Cell*. 2016; 166:492–505. [PubMed: 27419873]
- 36**. Dubin MJ, Zhang P, Meng D, Remigereau M-S, Osborne EJ, Paolo Casale F, Drewe P, Kahles A, Jean G, Vilhjálmsson B, et al. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation [Internet]. *Elife*. 2015; 4:e05255. This paper demonstrates that DNA methylation of TEs in CHH contexts increases with temperature and that natural variation in CHH methylation has a genetic basis. [PubMed: 25939354]
37. Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ, et al. Patterns of population epigenomic diversity [Internet]. *Nature*. 2013; 495:193–198. [PubMed: 23467092]
38. Eichten SR, Briskine R, Song J, Li Q, Swanson-Wagner R, Hermanson PJ, Waters AJ, Starr E, West PT, Tiffin P, et al. Epigenetic and Genetic Influences on DNA Methylation Variation in Maize Populations [Internet]. *Plant Cell*. 2013; 25:2783–2797. [PubMed: 23922207]
39. Seymour DK, Becker C. The causes and consequences of DNA methylome variation in plants. *Curr Opin Plant Biol*. 2017; 36:56–63. [PubMed: 28226269]
40. Yu A, Lepère G, Jay F, Wang J, Bapaume L, Wang Y, Abraham A-L, Penterman J, Fischer RL, Voinnet O, et al. Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. [Internet]. *Proc Natl Acad Sci U S A*. 2013; 110:2389–94. [PubMed: 23335630]
41. Stroud H, Hale CJ, Feng S, Caro E, Jacob Y, Michaels SD, Jacobsen SE. DNA methyltransferases are required to induce heterochromatic re-replication in Arabidopsis. [Internet]. *PLoS Genet*. 2012; 8:e1002808. [PubMed: 22792077]
42. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress [Internet]. *Nature*. 2011; 472:115–119. [PubMed: 21399627]
43. Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, Mittelsten Scheid O. How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation [Internet]. *PLoS Genet*. 2014; 10:e1004115. [PubMed: 24497839]
44. Lee T-J, Pascuzzi PE, Settlege SB, Shultz RW, Tanurdzic M, Rabinowicz PD, Menges M, Zheng P, Main D, Murray JAH, et al. Arabidopsis thaliana chromosome 4 replicates in two phases that correlate with chromatin state. [Internet]. *PLoS Genet*. 2010; 6:e1000982. [PubMed: 20548960]
45. Costas C, de la Paz Sanchez M, Stroud H, Yu Y, Oliveros JC, Feng S, Benguria A, López-Vidriero I, Zhang X, Solano R, et al. Genome-wide mapping of Arabidopsis thaliana origins of DNA replication and their associated epigenetic marks [Internet]. *Nat Struct Mol Biol*. 2011; 18:395–400. [PubMed: 21297636]

- 46*. Zhang J, Zuo T, Wang D, Peterson T. Transposition-mediated DNA re-replication in maize [Internet]. Elife. 2014; 3:e03724. This paper provides direct evidence for transposition-mediated DNA re-replication of an *Ac* element and flanking sequences. [PubMed: 25406063]
47. Feng W, Hale CJ, Over RS, Cokus SJ, Jacobsen SE, Michaels SD. Large-scale heterochromatin remodeling linked to overreplication-associated DNA damage [Internet]. Proc Natl Acad Sci. 2016; doi: 10.1073/pnas.1619774114
48. Jacob Y, Bergamin E, Donoghue MTA, Mongeon V, LeBlanc C, Voigt P, Underwood CJ, Brunzelle JS, Michaels SD, Reinberg D, et al. Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication [Internet]. Science. 2014; 343:1249–53. [PubMed: 24626927]
- 49*. Hale CJ, Potok ME, Lopez J, Do T, Liu A, Gallego-Bartolome J, Michaels SD, Jacobsen SE. Identification of Multiple Proteins Coupling Transcriptional Gene Silencing to Genome Stability in *Arabidopsis thaliana* [Internet]. PLOS Genet. 2016; 12:e1006092. This paper identifies multiple suppressors of over-replication of repetitive sequences in *atxr5 atxr6* and suggests loss of transcriptional silencing may be upstream of replication defects. [PubMed: 27253878]
50. Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver GP, Franklin FCH, et al. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. [Internet]. Nat Genet. 2013; 45:1327–36. [PubMed: 24056716]
51. Wijnker E, Velikkakam James G, Ding J, Becker F, Klasen JR, Rawat V, Rowan BA, de Jong DF, de Snoo CB, Zapata L, et al. The genomic landscape of meiotic crossovers and gene conversions in *Arabidopsis thaliana* [Internet]. Elife. 2013; 2:e01426. [PubMed: 24347547]
52. Demirci S, van Dijk ADJ, Sanchez Perez G, Aflitos SA, de Ridder D, Peters SA. Distribution, position, and genomic characteristics of crossovers in tomato recombinant inbred lines derived from an interspecific cross between *S. lycopersicum* and *S. pimpinellifolium*. [Internet]. Plant J. 2016; doi: 10.1111/tpj.13406
53. Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in *Arabidopsis*. [Internet]. Genes Dev. 2015; 29:2183–202. [PubMed: 26494791]
54. Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J, Grelon M, Copenhaver GP, et al. Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase mutants. [Internet]. PLoS Genet. 2012; 8:e1002844. [PubMed: 22876192]
55. Melamed-Bessudo C, Levy AA. Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in *Arabidopsis*. [Internet]. Proc Natl Acad Sci U S A. 2012; 109:E981–8. [PubMed: 22460791]
56. Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J, Paszkowski J. Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. [Internet]. Proc Natl Acad Sci U S A. 2012; 109:5880–5. [PubMed: 22451936]
57. Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, et al. Features of the *Arabidopsis* recombination landscape resulting from the combined loss of sequence variation and DNA methylation. [Internet]. Proc Natl Acad Sci U S A. 2012; 109:16240–5. [PubMed: 22988127]
58. Shi J, Wolf SE, Burke JM, Presting GG, Ross-Ibarra J, Dawe RK. Widespread Gene Conversion in Centromere Cores [Internet]. PLoS Biol. 2010; 8:e1000327. [PubMed: 20231874]
59. Yandea-Nelson MD, Zhou Q, Yao H, Xu X, Nikolau BJ, Schnable PS. MuDR Transposase Increases the Frequency of Meiotic Crossovers in the Vicinity of a Mu Insertion in the Maize *al* Gene [Internet]. Genetics. 2005; 169:917–929. [PubMed: 15489518]
60. Dooner HK, Martínez-Férez IM. Germinal excisions of the maize transposon activator do not stimulate meiotic recombination or homology-dependent repair at the *bz* locus [Internet]. Genetics. 1997; 147:1923–32. [PubMed: 9409847]
61. Xiao YL, Li X, Peterson T. *Ac* insertion site affects the frequency of transposon-induced homologous recombination at the maize *p1* locus [Internet]. Genetics. 2000; 156:2007–17. [PubMed: 11102391]

62. Ong-Abdullah M, Ordway JM, Jiang N, Ooi S-E, Kok S-Y, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK, et al. Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm [Internet]. *Nature*. 2015; 525:533–537. [PubMed: 26352475]
63. Fransz P, Linc G, Lee C-R, Aflitos SA, Lasky JR, Toomajian C, Ali H, Peters J, van Dam P, Ji X, et al. Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana* [Internet]. *Plant J*. 2016; 88:159–178. [PubMed: 27436134]
64. Zuo T, Zhang J, Lithio A, Dash S, Weber DF, Wise R, Nettleton D, Peterson T. Genes and Small RNA Transcripts Exhibit Dosage-Dependent Expression Pattern in Maize Copy-Number Alterations [Internet]. *Genetics*. 2016; 203:1133–47. [PubMed: 27129738]

Highlights

- Natural Arabidopsis accessions exhibit genetic and epigenetic variation of TEs.
- Variable enrichment of TE classes in euchromatin, pericentromeres and centromeres.
- Different epigenetic pathways control TE expression, replication and recombination.
- TEs play a major role in genome evolution and likely shape adaptive variation.

Figure 1. Epigenetic features of Arabidopsis transposable elements

TEs in the Arabidopsis genome are genetically heterogeneous, and silencing of each individual element likely relies upon the different molecular pathways to different extents. Here we schematically illustrate the key pathways known to play a role in TE silencing in Arabidopsis, at the level of transcription, DNA replication and meiotic recombination. Small interfering RNAs (siRNAs) are the initiators of silencing and direct *de novo* DNA cytosine methylation. DNA cytosine methylation is enriched at TEs and occurs in all sequence contexts. DNA cytosine methylation in non-CG contexts is highly correlated with H3K9me2. H3K27me1 is enriched at TEs and is present only on the histone H3 variant H3.1. Like H3.1, the H2A variant H2A.W is enriched at TEs, suggesting TE associated nucleosomes likely contain H3.1 and H2A.W.

Figure 2. Genomic distributions of *Arabidopsis thaliana* TE families

Plots of transposon density are shown along the *Arabidopsis* genome. The five chromosomes are plotted on a continuous x-axis shown in megabases (Mb). The position of centromeres are indicated by the vertical dotted lines and labels at the top of the figure. The number of elements analysed and the total base pairs of annotation are listed in parentheses [18]. The plot labels are color-coded indicating whether the TE family is a DNA (red) or RNA (blue) element.