Biochimica et Biophysica Acta xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Transposable element influences on gene expression in plants☆

Cory D. Hirsch [a], Nathan M. Springer [b,*]

[a] Department of Plant Pathology, University of Minnesota, Saint Paul, MN 55108, USA
[b] Department of Plant Biology, University of Minnesota, Saint Paul, MN 55108, USA

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form 17 May 2016
Accepted 18 May 2016
Available online xxxx

Keywords:
Transposable element
gene expression
chromatin modifications

ABSTRACT

Transposable elements (TEs) comprise a major portion of many plant genomes and bursts of TE movements cause novel genomic variation within species. In order to maintain proper gene function, plant genomes have evolved a variety of mechanisms to tolerate the presence of TEs within or near genes. Here, we review our understanding of the interactions between TEs and gene expression in plants by assessing three ways that transposons can influence gene expression. First, there is growing evidence that TE insertions within introns or untranslated regions of genes are often tolerated and have minimal impact on expression level or splicing. However, there are examples in which TE insertions within genes can result in aberrant or novel transcripts. Second, TEs can provide novel alternative promoters, which can lead to new expression patterns or original coding potential of an alternate transcript. Third, TE insertions near genes can influence regulation of gene expression through a variety of mechanisms. For example, TEs may provide novel *cis*-acting regulatory sites behaving as enhancers or insert within existing enhancers to influence transcript production. Alternatively, TEs may change chromatin modifications in regions near genes, which in turn can influence gene expression levels. Together, the interactions of genes and TEs provide abundant evidence for the role of TEs in changing basic functions within plant genomes beyond acting as latent genomic elements or as simple insertional mutagens. This article is part of a Special Issue entitled: Plant Gene Regulatory Mechanisms and Networks, edited by Dr. Erich Grotewold and Dr. Nathan Springer.

1. Introduction

1.1. Variation in plant genome architecture and the spatial organization of genes and transposons

1.1.1. Genomes as a composite of transposons and genes

Most eukaryotic genomes are predominantly composed of non-coding sequences. These non-coding sequences can include introns, intergenic DNA, and repetitive elements. A large proportion of the repetitive elements are actually derived from high-copy transposable elements (TEs). Many studies of genome function and chromatin modifications suggest there are substantial differences between genes and TEs [1]. Recombination tends to occur commonly within genes, but rarely at TEs [2]. Genes tend to have a set of chromatin modifications such as histone acetylation and methylation of histone H3 at lysine 4 (H3K4me) that are associated with active transcription [1]. In contrast, TEs are enriched for chromatin modifications such as CHG (where H is any nucleotide other than G) DNA methylation, H3K9me2, and H3K27me1, which are associated with transcriptional silencing [1]. While these are useful simplifications in describing genomes they fail to capture the nuances in a genome that is a composite of these types of features. There is an established yet delicate interplay between coding and non-coding sequences, which allows for necessary and essential functions to be carried out. In this review we will consider how TEs interact with plant genes.

1.1.1.1. Classification and annotation of TEs in plant genomes. Throughout this review we will often collectively refer to all transposable elements as TEs. However, there are a number of different types of TEs in plant genomes. Wicker et al. [3] provided an excellent classification scheme for transposable elements. Briefly, TEs are often broken into two different type classes based on their transposition mechanism. Class I elements, generally referred to as retrotransposons, are the most common feature in plant genomes and often contain long terminal repeats (LTRs) and transpose via an RNA intermediate using a copy and paste mechanism. This class of elements can be divided into five orders, each of which can be subdivided into multiple superfamilies based on their features and structure [3]. Class II elements, known as DNA transposons, often have terminal inverted repeats (TIRs) and utilize a DNA intermediate for movement that employs a "cut and paste" mechanism. DNA transposons are usually present at much lower levels within plant genomes than retrotransposons. In addition to TIR elements, the class II transposons

☆ This article is part of a Special Issue entitled: Plant Gene Regulatory Mechanisms and Networks, edited by Dr. Erich Grotewold and Dr. Nathan Springer.

* Corresponding author at: Department of Plant Biology, 140 Gortner Laboratory, 1479 Gortner Ave, Saint Paul, MN 55108, United States.
E-mail address: springer@umn.edu (N.M. Springer).

http://dx.doi.org/10.1016/j.bbagrm.2016.05.010
1874-9399/

also include Helitron elements that lack TIRs. Class I and class II elements include both autonomous and non-autonomous elements. Autonomous TEs encode all the necessary factors to facilitate transposition of an element itself or for other members of the same family. Non-autonomous elements require an autonomous element of the same family to provide the necessary factors for transposition. The patterns of chromatin modifications often vary for different types of stransposons [4,5] or for transposons of varying length [5,6].

As TEs proliferate in genomes and attain higher copy numbers the repetitive nature of the sequences can lead to difficulties in annotation due to their repetitive nature. In addition, many TEs have rearranged or truncated sequences due to abortive transposition events that lead to local rearrangements and deletions of internal sequences or nested insertions within other elements, which can lead to difficulties in recognition and annotation of full-length insertions [7]. The efforts of researchers to compare the genomic and epigenomic properties of genes and TEs in different plant genomes can be complicated by the fact that TE annotations are not conducted in a uniform fashion for all plant species. Adding increased difficulty to the problem is the differences in TE history among plant genomes. For example, a comparison of the *Arabidopsis thaliana*, *Arabidopsis lyrata* and *Arabis alpina* genomes reveals major differences in age distributions of TEs [8]. These differences in TE age structures can influence classification of TEs when different annotation pipelines are applied to related species. Currently, there are two main approaches for identifying TEs in fully sequenced genomes. One is based on sequence similarity to previously identified TEs. The other is a *de novo* approach usually based on searching for repetitive sequences and/or structural attributes common in TEs [9]. Each of these methods has limitations. Homology searching, for example using RepeatMasker, is limited in finding previously unknown TEs or TEs that are in low copy, whereas, *de novo* approaches can lead to numerous false-positive annotations. We note the difficulties of TE annotation and the variety of methods used for different species at the beginning of this review as a caveat to our comparisons of different species in the sections below. Obtaining consistent annotations for both genes and TEs in different plant genomes will increase our ability to make clear comparisons among plant species.

1.1.1.2. Variation in arrangement of genes and TEs. Plant genomes exhibit substantial diversity in the organization of TEs and genes (See Fig. 1). For example, only ~18% of the *Arabidopsis* genome has homology to TE sequences [10,11] while nearly 85% of the *Zea mays* (maize) genome has homology to TEs [12]. This is even more striking when genome size is considered. The *Arabidopsis* genome contains only ~20 Mbp of TEs, while the maize genome contains almost 2000 Mbp of TEs. In addition, the distribution of genes and TEs along chromosomes can vary substantially as well. Nearly 60% of the *Glycine* max (soybean) genome is composed of transposons [11]. However, the majority of TEs occur in the middle portion of chromosomes while the gene-rich chromosome arms have relatively few TEs [13]. In contrast, the maize genome has high levels of TEs throughout the entire chromosome [12,14]. While the central portion of maize chromosomes are ~85–90% TEs, the "gene-rich" regions of the chromosome arms are still 50% TE sequences [14]. A large portion (86%) of maize genes have TEs located within 1 kb, while only 36% of Arabidopsis genes include a TE within 1 kb [15]. In *Arabidopsis*, fewer than 100 genes contain large TE insertions within introns [16]. In contrast there are several thousand rice genes with large TE insertions within introns [16] and over 10% of maize genes contain at least one TE insertion over 1 kb within an intron [17].

TE insertions within the coding regions of a gene are generally mutagenic and result in strong loss-of-function alleles. However, insertions into introns or intergenic regions have less clear effects on gene expression. In this review we will evaluate three possible ways that TEs might influence gene expression and function beyond insertional mutagenesis of coding sequences. These include changes in transcript splicing/processing, providing novel promoters, and positive and negative influences on gene expression levels.

2. Intragenic TEs influence genic chromatin and transcript structure

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

Many plant genes contain TEs located within transcribed regions. In many cases genes with important functions contain TE insertions, suggesting that these TE insertions can be tolerated without compromising the function of the plant gene [17–19]. The presence of TEs within transcribed regions of plant genes raises several questions. First, is the chromatin of these intragenic TEs similar to other non-genic TEs or is it more similar to genic chromatin? Second, do these intragenic TEs have any influence on the proper expression and splicing of the genic transcript? To date, much of the information to address these questions has been from *Arabidopsis*, a species with relatively few TE insertions within genes. However, several recent studies have also provided information on chromatin, TEs, and gene expression in other plant species.

2.1.1. Chromatin modifications found at intragenic TEs

In general, plant genes and TEs contain quite distinct patterns for chromatin modifications. Plant genes tend to have very low levels of CHG and CHH methylation within the coding regions, but many

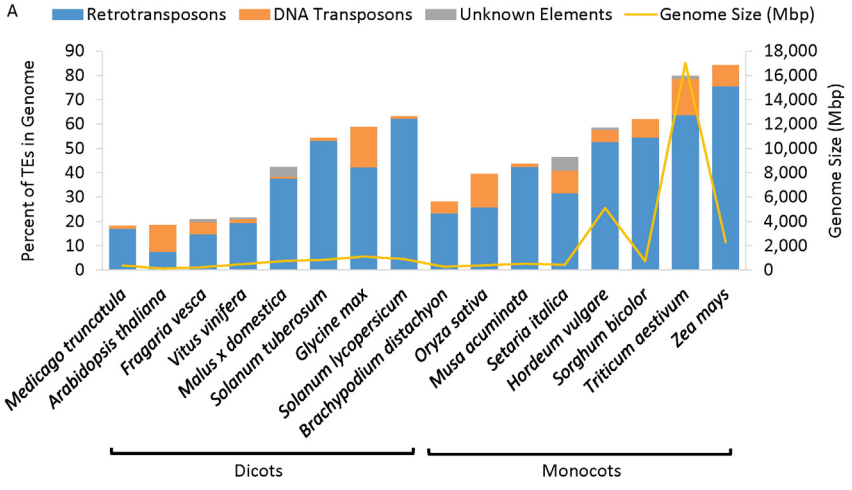

Fig. 1. Percent of transposon types and genome organization in representative plant genomes. For a variety of plant genome the proportion of the genome (as determined by [11]) derived from retrotransposons (blue), DNA transposons (orange) or unclassified transposons (gray) is shown. The genome size for each species is shown as a yellow line.

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

expressed plant genes contain moderate levels of CG methylation over the coding regions [20–23]. This gene body methylation tends to be found at genes with moderate expression levels and at similar genes in different species [19]. However, there is limited evidence for a functional role of this methylation [23]. While CHG methylation is rare in plant genes, it can be observed in some species, particularly when introns are assessed [16,17,19,24]. In *Arabidopsis*, the CHG methylation within genes was often present in relatively long introns [16]. In addition, there were even more examples of long introns containing high levels of CHG methylation in rice, and in many cases these introns contained a TE [16]. Similarly, the presence of intragenic CHG methylation in maize genes is primarily due to the fact that ~10% of maize genes have at least one intronic TE insertion and these intronic TEs often contain high levels of CHG methylation and H3K9me2 [17]. The presence of CHG or CHH methylation within soybean genes is associated with the presence of TEs in these genes [25]. There is evidence that the TEs located within *Arabidopsis* genes have slightly lower levels of DNA methylation than intergenic TEs [26], but the levels of DNA methylation are substantially higher than other genic regions. In contrast, intergenic and intragenic TEs have very similar patterns of chromatin modifications in maize [17]. The observation of elevated CHG methylation and H3K9me2 at intragenic TEs is somewhat surprising as these modifications are traditionally associated with heterochromatin. This suggests plants genes are able to tolerate these chromatin modifications within transcribed regions. When these chromatin modifications are present within promoter regions they are generally associated with gene silencing. The functional significance of the presence for these heterochromatin marks within genes will be examined in the next section.

2.1.2. Expression of genes containing TEs

Data from several plant species suggests that plants are able to express genes containing heterochromatic TEs within introns. In both maize and Norway spruce (*Picea abies*), there is evidence that genes with heterochromatin TE insertions can be expressed at quite high levels [17,18]. However, there is evidence that genes with TEs in *Arabidopsis* and soybean tend to be expressed at slightly lower levels on average [25,26]. Some *Arabidopsis* genes with TEs have altered transcript levels, as measured by the abundance of sequencing reads from exons downstream of the TE insertion, in plants having compromised epigenetic regulation [25]. However, the majority of *Arabidopsis* genes containing TEs insertions do not show different gene expression in these mutant plants [26]. These results suggest that the TE insertions and cognate chromatin can influence expression of some genes, but in many examples gene expression levels are not affected. However, changes in the chromatin state for TEs located within genes may influence gene expression. An intragenic TE located in an intron in the VRN-A1 gene in winter wheat (*Triticum aestivum*) exhibits changes in non-CG DNA methylation following vernalization [27]. These changes in DNA methylation for this intronic TE are associated with expression changes in the VRN-A1 gene that encodes a MADS transcription factor [27]. In soybean, a comparison of the paralogs derived from a recent whole genome duplication (WGD) event found that approximately 6% of pairs were differentially targeted by CHG methylation [25]. In many cases the paralog with higher CHG methylation contained a TE and sometimes this gene is expressed at levels lower than the other paralog. A comparison of retained soybean and common bean methylation patterns also provided evidence for changes in CHG methylation levels that are associated with TE insertions [25]. A study focused on class II DNA transposons in four grass species found that different TE families had varying relationships with gene expression [28]. The *Tc1*/*Mariner* and *PIF*/*Harbinger* TEs are enriched in genes with high expression, while *CACTA* elements are in genes with significantly lower expression than genes without TEs [28]. The exact mechanism or chromatin signal that allows for certain TE families to target active genes has not been fully elucidated. However, there is evidence that some TE proteins contain domains that can interact with histone modifications or RNA polymerase subunits in order to target integration to active or inactive chromatin [29,30]. TE families appear to exhibit diverse targeting mechanisms that allow some families to potentially reduce negative fitness consequences for the host through insertion into silenced heterochromatin, while other families may preferentially insert within active chromatin in order to escape silencing and increase their own fitness. The studies on expression of plant genes containing TEs suggest that plant genes can often tolerate heterochromatin TEs, but in some cases these may reduce expression of the gene.

2.1.3. Processing of transcripts from genes containing TEs

The presence of TEs within introns leads to several challenges for proper processing of transcripts. The analysis of mutant alleles derived from stocks with active DNA transposable elements has found examples of mutant alleles that are caused by the presence of a TE within an intron that contains splice sites leading to novel transcripts. For example, *Mutator* transposon insertions in the introns of *Adh1-S* [31] or *brown midrib1* (*bm1*) [32] can produce novel isoforms. Transposon insertions within exons can also be partially spliced out. A mutant allele of a maize chromomethylase, *zmet2-m1*, contains a *Mutator* insertion within an exon that can be partially spliced out, leaving a 40 bp insertion in the mature transcript [33]. In most cases, isoforms containing splicing between a gene and transposon will result in the production of non-functional proteins. There are also examples of alternative splicing between retrotransposons and plant genes [34]. A recent study, on the basis of the mantled trait that arises due to somaclonal variation in palm oil, found evidence that changes in chromatin at an intronic TE led to alternative splicing [35]. In soybean, an analysis of unstable alleles influencing seed color revealed the methylation patterns for a TE located within an intron of the *R* locus influenced whether splicing occurred properly [36]. Genome-wide analysis of splicing in maize suggests DNA methylation near splice sites may influence splicing patterns, although through unknown mechanisms [37]. There are examples of new exons in the human genome that are derived from TEs and these are frequently subjected to alternative splicing [38,39]. While there are examples of intronic TEs that can lead to altered splicing, it is worth noting that the majority of TE insertions in introns or untranslated regions do not lead to altered splicing and have no documented phenotypic consequences [40]. Many of the intragenic TE insertions in extant plant genomes are likely to have limited impact on transcripts. If TE insertions have strong effects on gene function it would likely result in phenotypic differences that could be subject to selective pressures resulting in reduced allele frequency for the TE-containing alleles. Many of the examples of altered splicing patterns are observed in populations segregating for new mutant alleles.

Plants have active mechanisms for tolerating the presence of heterochromatic TEs within transcribed regions [16,41]. An RNA-binding protein containing a bromo-adjacent homology (BAH) domain named *INCREASE IN BONSAI METHYLATION2* (*IBM2*) or *ANTI-SILENCING1* (*ASI1*) was identified in independent genetic screens in Arabidopsis [16,41]. This gene is required for proper expression of the H3K9 demethylase *IBM1*. *IBM1* contains a heterochromatic TE insertion within the seventh intron [16] that is normally tolerated, but in an *ibm2* mutant background the transcripts are prematurely terminated at the region containing the TE. In addition, other *Arabidopsis* genes with heterochromatin TEs within introns also exhibit premature termination in *ibm2* mutant plants [16,41]. The analysis of natural alleles without TE insertions provided evidence for the role of IBM2 to processing specific transcripts with intragenic heterochromatin although the exact mechanism by which IBM2 enables proper expression through heterochromatin has not been elucidated. There is also evidence for a role of *ENHANCED DOWNY MILDEW2* (*EDM2*) for proper expression of genes with intronic heterochromatin in *Arabidopsis* [42,43]. *EDM2* is a plant homeodomain (PHD) containing protein can bind to histones containing both repressive and active marks such that it might recognize transcribed regions containing heterochromatin marks [44,45]. As with *IBM2*, the exact mechanism by which *EDM2* influences proper expression through intronic TEs that contain heterochromatic

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

chromatin modification has not been determined. Homologs of *IBM2* and *EDM2* are found in other plant species and in many cases there are more genes with heterochromatin TEs in these species. It is possible that loss-of-function mutations in homologs of these genes would be lethal in other plant species with more examples of heterochromatic TEs within genes.

3. TEs as novel promoters

3.1. Transposons can provide novel promoters to drive expression of plant genes

In the previous section we examined the chromatin and functional impacts of TEs located within transcribed regions. In this section we will examine how TEs may impact promoter function. In this review we use the term promoter to describe the sequence that dictates the transcription start site (TSS) while the terms enhancer or regulatory regions are used to describe sequences that influence the expression level or pattern of nearby genes. The impact of TEs on regulatory regions will be examined in Section 4 of this review.

3.1.1. Frequent transcription start sites within TEs

There are several lines of evidence to suggest that transposons may provide novel promoters that would create novel TSSs for nearby genes. The first line of evidence is derived from the mapping of TSS sites within sequences that have homology to TEs. There is support for TE sequences within many eukaryotic promoters [46–48], and the analysis of TSSs has found that many mammalian TSSs actually occur within retroelements [49,50]. Several recent papers have mapped the TSSs in *Arabidopsis* and maize [51–53] providing an opportunity to assess whether certain transcripts initiate within TEs. A recent study experimentally determined the TSSs for over 17,000 maize genes and found that 180 of these TSSs are within TEs [53, E. Grotewold, unpublished observation]. As additional TSS mapping data becomes available for diverse plant species it will be interesting to monitor how TEs have contributed to TSS evolution among plants.

3.1.2. Cryptic promoters within transposons

The analysis of TE induced allelic variants has provided evidence that TEs can provide cryptic promoters to drive expression of nearby genes. In many cases these cryptic promoters have variable activity depending on the chromatin state of the TE. One of the best characterized examples of epigenetic inheritance is observed at the *Agouti viable yellow* (*Avy*) allele in mice [54,55]. At this locus an *IAP* element is inserted 5′ of the *Agouti* gene. The genetically defined allele can exist in two epigenetic states. In most cases the *IAP* element has high levels of DNA methylation and is transcriptionally silent. However, in epialleles with low methylation of the *IAP* element an outward reading promoter, will become active and drive expression of a novel transcript that includes the full Agouti coding region [55]. This leads to ectopic production of the Agouti protein and visible phenotypes. There is also evidence for *IAP* elements providing a novel promoter for the *Cabp* gene in mice depending on the chromatin state of the *IAP* element [56].

In plant species there are several well-characterized examples of transposons providing a cryptic promoter for nearby genes. Natural variation for fruit pigmentation in citrus species is due to allelic variation at the *Ruby* locus, which encodes a MYB transcriptional activator of anthocyanins [57]. Alleles that provide fruit color, as observed in blood oranges, are due to the insertion of a retrotransposon in the *Ruby* regulatory regions. Expression of *Ruby* is driven by a promoter within the retrotransposon LTR element [57]. Perhaps the best examples are observed for recessive suppressible alleles generated by *Mutator* insertions (illustrated in Fig. 2). A *Mutator* insertion within the promoter of the *hcf106* gene in maize can provide an outward reading promoter [58,59]. This promoter is only active when *Mutator* elements are silenced, leading to a suppressible phenotype. When *Mutator* elements are active the outward reading promoter is silenced and the mutant phenotype is observed. Similar mechanisms likely contribute to other suppressible alleles [60–63] and provide examples of transposons that can provide conditional promoters to drive expression of nearby genes. The analysis of phenotypes within *Mutator* active and inactive populations suggests that there are a substantial number of mutant alleles with suppressible phenotypes [64]. There are also examples of dominant *Mutator*-suppressible alleles but these appear to utilize a distinct mechanism for suppression [65].

4. TEs influences on regulation of gene expression

In addition to the potential to create new transcripts through alternative splicing, premature termination, or novel promoters, TEs can influence the expression level or expression pattern of plant genes [48,50,66–68]. TEs can influence gene expression through several potential mechanisms including disruption of *cis*-regulatory sequences, altered chromatin or providing novel regulatory information (Fig. 3). While there are examples of alleles with different expression patterns that are linked to insertions of nearby TEs it can be difficult to determine the exact mechanism through which the TE influences gene expression. In part, this is due to our limited understanding of the specific mechanisms of regulation for many plant genes. For most genes, the cis-acting regulatory sequences and *trans*-acting factors that influence expression have not been fully characterized and therefore it is difficult to know whether the TE insertion is disrupting this normal

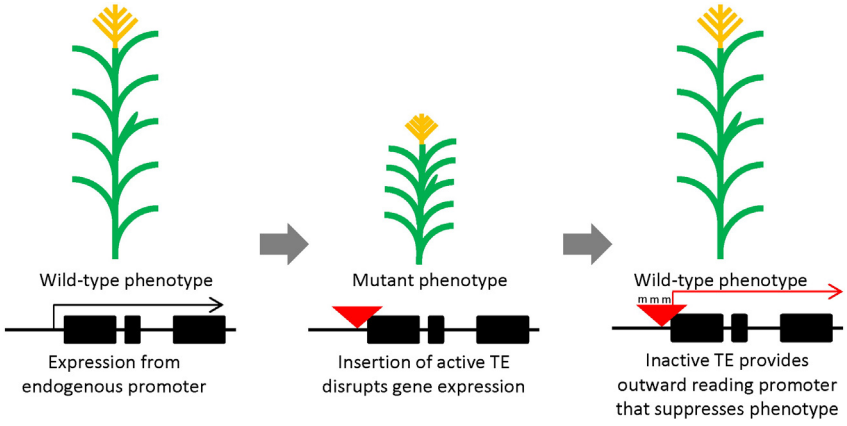

Fig. 2. Suppressible TE alleles with outward reading promoters. Example of a suppressible phenotype induced by a TE insertion. A TE insertion within the promoter or UTR region of a gene blocks expression when the TE is active. When the TE is silenced by chromatin modifications (m) an outward reading promoter within the TE becomes active and produces a transcript that can suppress the mutant phenotype.

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

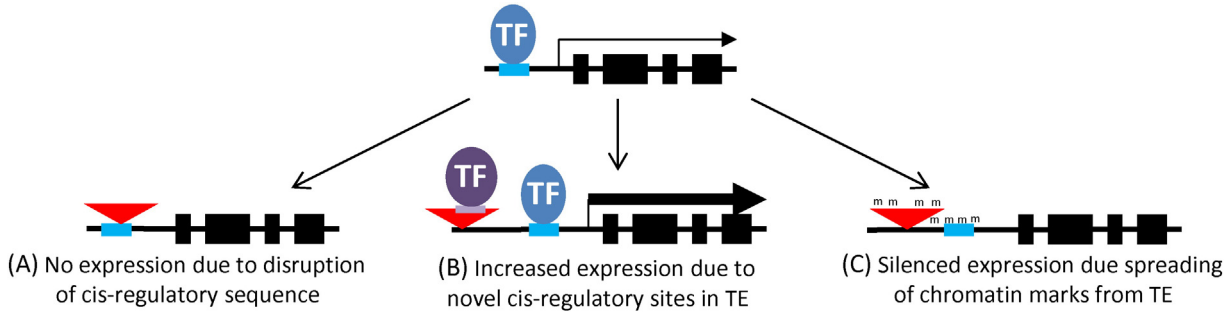

Fig. 3. Mechanisms by which TEs can influence expression of plant genes. A typical plant gene is diagramed with a cis-acting regulatory site that provides a binding site (blue square) for a transcription factor (blue circle) and moderate expression (indicated by black arrow). A TE (red triangle) insertion can have different potential influences on the expression of this gene. In (A) the TE inserts within the regulatory sequence and prevents binding of the TF which reduces expression of the gene. In (B) the TE inserts upstream of the gene and contains a novel cis-regulatory site (purple rectangle) that is bound by another transcription factor which results in increased expression or expression in novel tissues or cell types. In (C) the TE is subject to chromatin modifications (m) that can spread to cover nearby regulatory sequences or the promoter resulting in reduced expression of the gene.

regulation. In addition, most TEs have chromatin modifications that are distinct from the chromatin state of nearby regulatory regions. In order to determine whether the chromatin modifications are important for influencing expression of nearby genes it is important to be able to compare gene expression between alleles with and without the chromatin modifications. For some TE families this can be done by comparing gene expression in stocks with active or inactive TEs but in most cases we do not have access to stocks with active TEs. Alternatively, gene expression could be surveyed in plants that have mutations affecting chromatin state. In many plant species mutations that strongly affect DNA methylation or other silencing modifications are not viable which limits researchers' ability to probe the influence of TE chromatin on expression of nearby genes [6,69]. In this section we will describe different mechanisms by which transposons may influence gene regulation and will highlight examples in which a particular mechanism is suspected.

4.1. Insertional mutagenesis of gene regulation by transposons

Transposon insertions can disrupt existing regulatory information (Fig. 3A). A significant portion of transcriptional regulation occurs via the binding of trans-acting factors to *cis*-acting regulatory sequences. When TEs insert within these cis-acting regulatory sequences they can disrupt the transcription factor binding and result in aberrant gene expression. Several dominant alleles resulting from *Mutator* insertions provide nice examples of how TE insertions can disrupt proper gene regulation. The *Knotted1*, *Rough Sheath1* and *Liguleless3* genes of maize encode homeobox genes that are important for maintaining meristem cell identity, and must be repressed to allow for proper development of the leaf [70–72]. These genes were all identified through the analysis of dominant mutations in which *Mutator* transposons inserted within the third intron. In all three cases the dominant mutant allele is not properly repressed during leaf initiation and prolonged expression of these meristem identity genes results in homeotic transformations within the leaf. Presumably, there are important regulatory regions within the third intron of these genes that are required for proper repression of this set of homeobox genes during leaf initiation and the TE insertions in this region interfere with this repression, resulting in ectopic expression. These dominant mutant alleles triggered by TE insertions into binding sites for repressors provide an insight into how TEs may disrupt normal gene regulation.

It is likely even more common for TE insertions within enhancers or promoters to reduce gene expression and result in recessive alleles (Fig. 3B–C). The loss of pigmentation in certain grape varieties is due to the insertion of a retrotransposon in the promoter region of a Myb transcription factor gene [73]. The *tb1* locus in maize provides an example of an allelic variant that was critical for domestication [74] that is due to the insertion of a TE within a distal regulatory region [75]. A TE insertion within the promoter of the *ZmCCT* locus of maize results in reduced expression for this gene and alters photoperiod sensitivity [76]. A major QTL influencing flowering time in maize, *Vgt1*, was fine-mapped to a conserved non-coding region located 70 kb from an Ap2-like transcription factor [77]. The natural variation at *Vgt1* is due to the presence of a MITE element that has inserted within this region [77]. *Vgt1* provides an example in which the exact mechanism is unclear. The MITE insertion may disrupt the normal function of this regulatory region by interfering with proper binding of *trans*-acting factors. However, the analysis of DNA methylation at this region has found that the MITE insertion is also associated with changes in DNA methylation that may affect proper regulation [78]. These examples provide evidence that TEs can disrupt normal regulation of gene expression through insertions in *cis*-regulatory regions, and highlight how TE mutagenesis can reveal mechanisms of endogenous gene regulation.

4.2. Chromatin-based regulation of gene expression by transposons

Transposons can influence expression of genes by changing the local chromatin states. TEs are generally marked with high levels of heterochromatic marks that may interfere with transcriptional regulation. If these marks are confined to the transposon itself, then they will interfere with the potential for enhancers within the transposon to exert influence on the nearby genes. However, if the chromatin marks are not confined to the transposon, but can spread to flanking sequences they may influence gene regulation. Hollister and Gaut [79] noted that heavily methylated transposons in Arabidopsis promoters are associated with reduced gene expression levels. They also found evidence that heavily methylated transposons near genes are often subjected to purifying selection [79]. These results suggest that the chromatin state of TEs near genes may influence expression of nearby genes. An alternative explanation is that TEs may preferentially insert into genes with low expression levels [80]. The analysis of natural variation has provided conflicting evidence on whether the presence of the TE is associated with expression [79,80]. The analysis of chromatin modifications near TEs has found evidence for the ability of heterochromatin to spread outside the borders of the transposon [4,81]. This spreading of heterochromatin varies for different TE families [4] and could result in reduced expression for nearby genes.

There are a number of examples in which the chromatin state for a transposon near a gene has been linked to expression for the nearby gene. Hypermethylation of a LINE element can lead to silencing of a nearby Anaphase-Promoting Complex (APC) 13 gene in *Arabidopsis* [82]. Sex determination in melons is associated with a methylated TE inserted near the *CmWIP* gene and revertants are associated with a loss of methylation for this TE [83]. The chromatin state for a retrotransposon insertion can influence expression for the nearby *FAE1* gene in yellow mustard [84]. One particularly intriguing example of how chromatin modifications at a TE can affect gene expression is found at the *ROS1*

locus in *Arabidopsis* [85,86]. The expression of *ROS1*, a DNA demethylase gene in *Arabidopsis*, is influenced by the DNA methylation level of a helitron element located upstream of ROS1. This provides a feedback loop for sensing DNA methylation levels and adjusting expression of a demethylase [85,86].

The FWA locus in *Arabidopsis* provides an example of how TE insertions and chromatin might influence expression of different epialleles during development. *FWA* is normally silenced throughout vegetative growth in plants, but is expressed in an imprinted fashion (solely from the maternal allele) in endosperm tissue [87]. Epialleles of *FWA* due to reduced methylation of a SINE element in the promoter are expressed in vegetative tissues and result in a late-flowering phenotype [88,89]. The imprinted expression of the maternal allele of *FWA* likely results from demethylation of the adjacent TE in the central cell [87]. Indeed, many genes with imprinted expression are located near TEs that undergo targeted demethylation in the central cell [90–94]. The analysis of allelic variation for imprinted expression within *Arabidopsis* provides evidence that epigenetic variation at nearby TE insertions can result in variable imprinting [95]. These imprinted genes provide examples for how developmentally induced changes in chromatin of TEs can influence expression of nearby genes.

Recent studies have found that changes in the chromatin state of TEs near genes may play a role in response to environmental conditions. Many stress-responsive genes in *Arabidopsis* contain TEs in the promoter region. In addition, these TEs are actively demethylated when the genes are up-regulated [96]. Many rice genes that are induced by low phosphate stress exhibit increased levels of DNA methylation at transposons near the promoter [97]. However, in this study it appears that the change of chromatin at the nearby TE is a result of altered expression rather than a cause of the altered gene expression.

4.3. Transposons as a source of novel regulatory information

Changes in gene expression are a critical source of phenotypic diversity and adaptation [98]. A fundamental question is how genes acquire novel *cis*-regulatory sequences that provide new expression patterns in response to environmental or developmental cues. In many cases, it would be difficult for single nucleotide polymorphisms (SNPs) to create new binding sites for transcription factors. TEs are one possible sources of novel *cis*-regulatory information that could influence expression of nearby genes (Fig. 3).

TEs themselves are subject to complex regulation and require proper *cis*-regulatory sequences for this regulation. TEs can exhibit stress-responsive transcription or mobilization [99–102]. The tobacco *Tnt1* element can be induced by biotic or abiotic stress [100]. Many TEs can be activated by tissue culture [103,104]. The *mPing* DNA transposon in rice can be activated in response to cold or salt stress [105,106]. There are examples of TEs with developmentally regulated gene expression. Many TEs in *Arabidopsis* are activated in the vegetative nucleus of pollen cells [107]. Some animal transposons contain silencers that can trigger recruitment of Polycomb proteins [108]. For a TE to maintain proper regulation in response to stress or developmental cues the cis-acting regulatory sites would need to be located within the TE itself such that the cis-acting regulation would move with the TE. Local enhancers located within the TE would need to act upon the TE promoter to condition proper expression of the transposon. However, it is quite possible that these cis-acting regulatory sites would also be able to influence expression of nearby genic promoters as well. The concept that TEs may act to influence expression of nearby genes fits with McClintock's suggestion that TEs can play important roles in the response of genomes to environmental stress and controlling gene expression [109–111].

The *ONSEN* retrotransposable element of *Arabidopsis* provides an excellent example of a TE with complex regulation that can influence the regulation of nearby genes. *ONSEN* elements are transcriptionally silent but can be activated by heat-stress [112,113]. The activation of *ONSEN* requires DNA binding sites for heat-shock factors [100] and is also influenced by chromatin modifications [112]. *ONSEN* elements can also affect the heat-responsive expression for nearby genes [112]. When *ONSEN* elements are inserted nearby genes will be up-regulated in response to heat, but natural variant alleles lacking the *ONSEN* insertion do not show heat-responsive expression [112]. There is evidence for TEs influencing expression of nearby genes in other plant species as well. Rice harbors numerous *mPing* insertions near the 5′ ends of genes and many of these have little effect on expression in control conditions [105]. However, mPing itself is up-regulated in response to cold stress and nearby genes also exhibit cold-responsive expression when *mPing* elements are present [105]. Similar regulation of expression has been observed for rice genes located near *mPing* insertions [105]. A genome-wide analysis of TEs near stress-responsive genes in maize found evidence for a ~ 20 TE families that are associated with stress-responsive expression of nearby genes [114]. In most cases, expression of the TEs themselves is responsive to the same stress conditions and they appear to provide enhancers that influence the expression of nearby genes [114].

We have highlighted several mechanisms by which transposons can influence plant gene expression that are well supported in the literature. However, there are additional mechanisms by which transposons may influence gene expression that have not been as well characterized to date. For example, some transposable elements have been "domesticated" by plant genomes and now play important roles in plant development or gene regulation [115,116]. There is also evidence that siRNAs generated from TEs can influence the expression of genic mRNAs [117]. In both of these examples the TEs can act in trans to affect the expression of genes located at unlinked genomic positions through providing regulatory protein or siRNA sequences.

5. Concluding remarks

Transposons have the potential to play diverse roles in shaping plant genomes. While many TE insertions within genes can result in loss-of-function mutations there are also opportunities for TEs to influence transcripts. Here, we have evaluated a variety of mechanisms by which TEs can interact with plant genes to shape transcripts and expression levels. TEs located within plant genes bring a set of chromatin modifications associated with heterochromatin, but are largely tolerated. Plants have evolved mechanisms to allow for proper transcription and processing through heterochromatin TEs located within introns. However, changes in the chromatin state of these intragenic TEs can result in aberrant splicing patterns and novel function for genes. TEs can also affect gene expression through providing novel promoters or influencing regulation. It is possible that a substantial portion of gene regulatory variation originates from transposon-mediated changes. Over time, the regulatory information may be subject to selection while the rest of the transposon sequence decays past the point of detection as a TE. Understanding the mechanisms by which TEs influence gene expression could shed light on the evolutionary basis for regulatory diversity in flowering plants. A more complete understanding of the roles of TEs in generating expression diversity could also allow for novel approaches for generating expression diversity for response to abiotic stress or developmental cues in crop plants.

Conflict of interest

We declare that we do not have any conflicts of interest relevant to the submitted manuscript.

Acknowledgements

We are grateful to Candice N. Hirsch, Amanda J. Waters and Qing Li for reading the manuscript and providing useful comments. N.M.S. and C.D.H. are supported by a grant from USDA-NIFA 2016-67013-24747.

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

References

[1] H. Saze, T. Kakutani, Differentiation of epigenetic modifications between transposons and genes, Curr. Opin. Plant Biol. 14 (2011) 81–87.
[2] H. Fu, Z. Zheng, H.K. Dooner, Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 1082–1087.
[3] T. Wicker, F. Sabot, A. Hua-Van, J.L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A.H. Schulman, A unified classification system for eukaryotic transposable elements, Nat. Rev. Genet. 8 (2007) 973–982.
[4] S.R. Eichten, N.A. Ellis, I. Makarevitch, C.T. Yeh, J.I. Gent, L. Guo, K.M. McGinnis, X. Zhang, P.S. Schnable, M.W. Vaughn, R.K. Dawe, N.M. Springer, Spreading of heterochromatin is limited to specific families of maize retrotransposons, PLoS Genet. 8 (2012), e1003127.
[5] A. Zemach, M.Y. Kim, P.H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S.L. Harmer, D. Zilberman, The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin, Cell 153 (2013) 193–205.
[6] Q. Li, S.R. Eichten, P.J. Hermanson, V.M. Zaunbrecher, J. Song, J. Wendt, H. Rosenbaum, T.F. Madzima, A.E. Sloan, J. Huang, D.L. Burgess, T.A. Richmond, K.M. McGinnis, R.B. Meeley, O.N. Danilevskaya, M.W. Vaughn, S.M. Kaeppler, J.A. Jeddeloh, N.M. Springer, Genetic perturbation of the maize methylome, Plant Cell 26 (2014) 4602–4616.
[7] P. SanMiguel, A. Tikhonov, Y.K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P.S. Springer, K.J. Edwards, M. Lee, Z. Avramova, J.L. Bennetzen, Nested retrotransposons in the intergenic regions of the maize genome, Science 274 (1996) 765–768.
[8] E. Willing, V. Rawat, T. Mandáková, F. Maumus, G.V. James, K.J.V. Nordström, C. Becker, N. Warthmann, C. Chica, B. Szarzynska, M. Zytnicki, M.C. Albani, C. Kiefer, S. Bergonzi, L. Castaings, J.L. Mateos, M.C. Berns, N. Bujdoso, T. Piofczyk, L. de Lorenzo, C. Barrero-Sicilia, I. Mateos, M. Piednoël, J. Hagmann, R. Chen-Min-Tao, R. Iglesias-Fernández, S.C. Schuster, C. Alonso-Blanco, F. Roudier, P. Carbonero, J. Paz-Ares, S.J. Davis, A. Pecinka, H. Quesneville, V. Colot, M.A. Lysak, D. Weigel, G. Coupland, K. Schneeberger, Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation, Nat. Plants 1 (2015) 14023.
[9] E. Lerat, Identifying repeats and transposable elements in sequenced genomes: how to find your way through the dense forest of programs, Heredity (Edinb) 104 (2010) 520–533.
[10] Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, Nature 408 (2000) 796–815.
[11] K.R. Oliver, J.A. McComb, W.K. Greene, Transposable elements: powerful contributors to angiosperm evolution and diversity, Genome Biol. Evol. 5 (2013) 1886–1901.
[12] P.S. Schnable, D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T.A. Graves, P. Minx, A.D. Reily, L. Courtney, S.S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S.M. Rock, E. Belter, F. Du, K. Kim, R.M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S.M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M.J. Levy, L. McMahan, P. Van Buren, M.W. Vaughn, K. Ying, C.T. Yeh, S.J. Emrich, Y. Jia, A. Kalyanaraman, A.P. Hsia, W.B. Barbazuk, R.S. Baucom, T.P. Brutnell, N.C. Carpita, C. Chaparro, J.M. Chia, J.M. Deragon, J.C. Estill, Y. Fu, J.A. Jeddeloh, Y. Han, H. Lee, P. Li, D.R. Lisch, S. Liu, Z. Liu, D.H. Nagel, M.C. McCann, P. SanMiguel, A.M. Myers, D. Nettleton, J. Nguyen, B.W. Penning, L. Ponnala, K.L. Schneider, D.C. Schwartz, A. Sharma, C. Soderlund, N.M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T.K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J.L. Bennetzen, R.K. Dawe, J. Jiang, N. Jiang, G.G. Presting, S.R. Wessler, S. Aluru, R.A. Martienssen, S.W. Clifton, W.R. McCombie, R.A. Wing, R.K. Wilson, The B73 maize genome: complexity, diversity, and dynamics, Science 326 (2009) 1112–1115.
[13] J. Schmutz, S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q. Song, J.J. Thelen, J. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J. Du, Z. Tian, L. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, S.A. Jackson, Genome sequence of the palaeopolyploid soybean, Nature 463 (2010) 178–183.
[14] R.S. Baucom, J.C. Estill, C. Chaparro, N. Upshaw, A. Jogi, J.M. Deragon, R.P. Westerman, P.J. Sanmiguel, J.L. Bennetzen, Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome, PLoS Genet. 5 (2009), e1000732.
[15] Q. Li, J.I. Gent, G. Zynda, J. Song, I. Makarevitch, C.D. Hirsch, C.N. Hirsch, R.K. Dawe, T.F. Madzima, K.M. McGinnis, D. Lisch, R.J. Schmitz, M.W. Vaughn, N.M. Springer, RNA-directed DNA methylation enforces boundaries between heterochromatin and euchromatin in the maize genome, Proc. Natl. Acad. Sci. U. S. A. (2015).
[16] H. Saze, J. Kitayama, K. Takashima, S. Miura, Y. Harukawa, T. Ito, T. Kakutani, Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin, Nat. Commun. 4 (2013) 2301.
[17] P.T. West, Q. Li, L. Ji, S.R. Eichten, J. Song, M.W. Vaughn, R.J. Schmitz, N.M. Springer, Genomic distribution of H3K9me2 and DNA methylation in a maize genome, PLoS One 9 (2014), e105267.
[18] B. Nystedt, N.R. Street, A. Wetterbom, A. Zuccolo, Y.C. Lin, D.G. Scofield, F. Vezzi, N. Delhomme, S. Giacomello, A. Alexeyenko, R. Vicedomini, K. Sahlin, E. Sherwood, M. Elfstrand, L. Gramzow, K. Holmberg, J. Hallman, O. Keech, L. Klasson, M. Koriabine, M. Kucukoglu, M. Kaller, J. Luthman, F. Lysholm, T. Niittyla, A. Olson, N. Rilakovic, C. Ritland, J.A. Rossello, J. Sena, T. Svensson, C. Talavera-Lopez, G. Theissen, H. Tuominen, K. Vanneste, Z.Q. Wu, B. Zhang, P. Zerbe, L. Arvestad, R. Bhalerao, J. Bohlmann, J. Bousquet, R. Garcia Gil, T.R. Hvidsten, P. de Jong, J. MacKay, M. Morgante, K. Ritland, B. Sundberg, S.L. Thompson, Y. Van de Peer, B. Andersson, O. Nilsson, P.K. Ingvarsson, J. Lundeberg, S. Jansson, The Norway spruce genome sequence and conifer genome evolution, Nature 497 (2013) 579–584.
[19] S. Takuno, J. Ran, B.S. Gaut, Evolutionary patterns of genic DNA methylation vary across land plants, Nat. Plants 2 (2016) 15222.
[20] R.K. Tran, J.G. Henikoff, D. Zilberman, R.F. Ditt, S.E. Jacobsen, S. Henikoff, DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes, Curr. Biol. 15 (2005) 154–159.
[21] X. Zhang, J. Yazaki, A. Sundaresan, S. Cokus, S.W. Chan, H. Chen, I.R. Henderson, P. Shinn, M. Pellegrini, S.E. Jacobsen, J.R. Ecker, Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis, Cell 126 (2006) 1189–1201.
[22] D. Zilberman, M. Gehring, R.K. Tran, T. Ballinger, S. Henikoff, Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription, Nat. Genet. 39 (2007) 61–69.
[23] A.J. Bewick, L. Ji, C.E. Niederhuth, E. Willing, B.T. Hofmeisterb, X. Shia, L. Wang, Z. Lua, N.A. Rohra, B. Hartwig, C. Kiefer, R.B. Deal, J. Schmutz, J. Grimwood, H. Stroud, S.E. Jacobsen, K. Schneeberger, X. Zhang, R.J. Schmitz, On the Origin and Evolutionary Consequences of Gene Body DNA Methylation, Biorxiv (2016), http://dx.doi.org/10.1101/045542.
[24] C.E. Niederhuth, A.J. Bewick, L. Ji, M.S. Alabady, K.D. Kim, J.T. Page, Q. Li, N.A. Rohr, A. Rambani, J.M. Burke, J.A. Udall, C. Egesi, J. Schmutz, J. Grimwood, S.A. Jackson, N.M. Springer, R.J. Schmitz, Widespread natural variation of DNA methylation within angiosperms, Biorxiv (2016), http://dx.doi.org/10.1101/045880.
[25] K.D. Kim, M. El Baidouri, B. Abernathy, A. Iwata-Otsubo, C. Chavarro, M. Gonzales, M. Libault, J. Grimwood, S.A. Jackson, A Comparative Epigenomic Analysis of Polyploidy-Derived Genes in Soybean and Common Bean, Plant Physiol. 168 (2015) 1433–1447.
[26] T.N. Le, Y. Miyazaki, S. Takuno, H. Saze, Epigenetic regulation of intragenic transposable elements impacts gene transcription in *Arabidopsis thaliana*, Nucleic Acids Res. 43 (2015) 3911–3921.
[27] A.R. Khan, J. Enjalbert, A.C. Marsollier, A. Rousselet, I. Goldringer, C. Vitte, Vernalization treatment induces site-specific DNA hypermethylation at the VERNALIZATION-A1 (VRN-A1) locus in hexaploid winter wheat, BMC Plant Biol. 13 (2013) (209–2229–13-209).
[28] Y. Han, S. Qin, S.R. Wessler, Comparison of class 2 transposable elements at superfamily resolution reveals conserved and distinct features in cereal grass genomes, BMC Genomics 14 (2013) (71–2164–14-71).
[29] X. Gao, Y. Hou, H. Ebina, H.L. Levin, D.F. Voytas, Chromodomains direct integration of retrotransposons to heterochromatin, Genome Res. 18 (2008) 359–369.
[30] A. Bridier-Nahmias, A. Tchalikian-Cosson, J.A. Baller, R. Menouni, H. Fayol, A. Flores, A. Saib, M. Werner, D.F. Voytas, P. Lesage, Retrotransposons. An RNA polymerase III subunit determines sites of retrotransposon integration, Science 348 (2015) 585–588.
[31] K.R. Luehrsen, V. Walbot, Insertion of Mu1 elements in the first intron of the Adh1-S gene of maize results in novel RNA processing events, Plant Cell 2 (1990) 1225–1238.
[32] W. Chen, N. VanOpdorp, D. Fitzl, J. Tewari, P. Friedemann, T. Greene, S. Thompson, S. Kumpatla, P. Zheng, Transposon insertion in a cinnamyl alcohol dehydrogenase gene is responsible for a brown midrib1 mutation in maize, Plant Mol. Biol. 80 (2012) 289–297.
[33] C.M. Papa, N.M. Springer, M.G. Muszynski, R. Meeley, S.M. Kaeppler, Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation, Plant Cell 13 (2001) 1919–1928.
[34] M.J. Varagona, M. Purugganan, S.R. Wessler, Alternative splicing induced by insertion of retrotransposons into the maize waxy gene, Plant Cell 4 (1992) 811–820.
[35] M. Ong-Abdullah, J.M. Ordway, N. Jiang, S.E. Ooi, S.Y. Kok, N. Sarpan, N. Azimi, A.T. Hashim, Z. Ishak, S.K. Rosli, F.A. Malike, N.A. Bakar, M. Marjuni, N. Abdullah, Z. Yaakub, M.D. Amiruddin, R. Nookiah, R. Singh, E.T. Low, K.L. Chan, N. Azizi, S.W. Smith, B. Bacher, M.A. Budiman, A. Van Brunt, C. Wischmeyer, M. Beil, M. Hogan, N. Lakey, C.C. Lim, X. Arulandoo, C.K. Wong, C.N. Choo, W.C. Wong, Y.Y. Kwan, S.S. Alwee, R. Sambanthamurthi, R.A. Martienssen, Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm, Nature 525 (2015) 533–537.
[36] G. Zabala, L.O. Vodkin, Methylation affects transposition and splicing of a large CACTA transposon from a MYB transcription factor regulating anthocyanin synthase genes in soybean seed coats, PLoS One 9 (2014), e111959.
[37] M. Regulski, Z. Lu, J. Kendall, M.T. Donoghue, J. Reinders, V. Llaca, S. Deschamps, A. Smith, D. Levy, W.R. McCombie, S. Tingey, A. Rafalski, J. Hicks, D. Ware, R.A. Martienssen, The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA, Genome Res. 23 (2013) 1651–1662.
[38] Y. Xing, C. Lee, Alternative splicing and RNA selection pressure–evolutionary consequences for eukaryotic genomes, Nat. Rev. Genet. 7 (2006) 499–509.
[39] S. Ayarpadikannan, H.E. Lee, K. Han, H.S. Kim, Transposable element-driven transcript diversification and its relevance to genetic disorders, Gene 558 (2015) 187–194.
[40] C.R. Dietrich, F. Cui, M.L. Packila, J. Li, D.A. Ashlock, B.J. Nikolau, P.S. Schnable, Maize Mu transposons are targeted to the 5' untranslated region of the gl8 gene and sequences flanking Mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome, Genetics 160 (2002) 697–716.
[41] X. Wang, C.G. Duan, K. Tang, B. Wang, H. Zhang, M. Lei, K. Lu, S.K. Mangrauthia, P. Wang, G. Zhu, Y. Zhao, J.K. Zhu, RNA-binding protein regulates plant DNA methylation

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

by controlling mRNA processing at the intronic heterochromatin-containing gene IBM1, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 15467–15472.
[42] T. Eulgem, T. Tsuchiya, X.J. Wang, B. Beasley, A. Cuzick, M. Tor, T. Zhu, J.M. McDowell, E. Holub, J.L. Dangl, EDM2 is required for RPP7-dependent disease resistance in Arabidopsis and affects RPP7 transcript levels, Plant J. 49 (2007) 829–839.
[43] T. Tsuchiya, T. Eulgem, An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E3535–E3543.
[44] T. Tsuchiya, T. Eulgem, The PHD-finger module of the *Arabidopsis thaliana* defense regulator EDM2 can recognize triply modified histone H3 peptides, Plant Signal. Behav. 9 (2014), e29202.
[45] M. Lei, H. La, K. Lu, P. Wang, D. Miki, Z. Ren, C.G. Duan, X. Wang, K. Tang, L. Zeng, L. Yang, H. Zhang, W. Nie, P. Liu, J. Zhou, R. Liu, Y. Zhong, D. Liu, J.K. Zhu, Arabidopsis EDM2 promotes IBM1 distal polyadenylation and regulates genome DNA methylation patterns, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 527–532.
[46] I.K. Jordan, I.B. Rogozin, G.V. Glazko, E.V. Koonin, Origin of a substantial fraction of human regulatory sequences from transposable elements, Trends Genet. 19 (2003) 68–72.
[47] A.E. Peaston, A.V. Evsikov, J.H. Graber, W.N. de Vries, A.E. Holbrook, D. Solter, B.B. Knowles, Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos, Dev. Cell 7 (2004) 597–606.
[48] C. Feschotte, Transposable elements and the evolution of regulatory networks, Nat. Rev. Genet. 9 (2008) 397–405.
[49] G.J. Faulkner, Y. Kimura, C.O. Daub, S. Wani, C. Plessy, K.M. Irvine, K. Schroder, N. Cloonan, A.L. Steptoe, T. Lassmann, K. Waki, N. Hornig, T. Arakawa, H. Takahashi, J. Kawai, A.R. Forrest, H. Suzuki, Y. Hayashizaki, D.A. Hume, V. Orlando, S.M. Grimmond, P. Carninci, The regulated retrotransposon transcriptome of mammalian cells, Nat. Genet. 41 (2009) 563–571.
[50] P. Mita, J.D. Boeke, How retrotransposons shape genome regulation, Curr. Opin. Genet. Dev. 37 (2016) 90–100.
[51] Y.Y. Yamamoto, T. Yoshitsugu, T. Sakurai, M. Seki, K. Shinozaki, J. Obokata, Heterogeneity of Arabidopsis core promoters revealed by high-density TSS analysis, Plant J. 60 (2009) 350–362.
[52] T. Morton, J. Petricka, D.L. Corcoran, S. Li, C.M. Winter, A. Carda, P.N. Benfey, U. Ohler, M. Megraw, Paired-end analysis of transcription start sites in Arabidopsis reveals plant-specific promoter signatures, Plant Cell 26 (2014) 2746–2760.
[53] M. Mejia-Guerra, W. Li, N.F. Galeano, M. Vidal Gray, D.J. A.I., E. Grotewold, Core promoter plasticity between maize tissues and genotypes contrasts with predominance of sharp transcription initiation sites, Plant Cell (2015).
[54] E.J. Michaud, M.J. van Vugt, S.J. Bultman, H.O. Sweet, M.T. Davisson, R.P. Woychik, Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage, Genes Dev. 8 (1994) 1463–1472.
[55] H.D. Morgan, H.G. Sutherland, D.I. Martin, E. Whitelaw, Epigenetic inheritance at the agouti locus in the mouse, Nat. Genet. 23 (1999) 314–318.
[56] R. Druker, T.J. Bruxner, N.J. Lehrbach, E. Whitelaw, Complex patterns of transcription at the insertion site of a retrotransposon in the mouse, Nucleic Acids Res. 32 (2004) 5800–5808.
[57] E. Butelli, C. Licciardello, Y. Zhang, J. Liu, S. Mackay, P. Bailey, G. Reforgiato-Recupero, C. Martin, Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges, Plant Cell 24 (2012) 1242–1255.
[58] A. Barkan, R. Martienssen, Inactivation of Maize Transposon Mu Suppresses a Mutant Phenotype by Activating an Outward-Reading Promoter Near the End of Mu1, PNAS 88 (1991) 3502–3506.
[59] A.M. Settles, A. Baron, A. Barkan, R.A. Martienssen, Duplication and suppression of chloroplast protein translocation genes in maize, Genetics 157 (2001) 349–360.
[60] P. Chomet, D. Lisch, K.J. Hardeman, V.L. Chandler, M. Freeling, Identification of a regulatory transposon that controls the Mutator transposable element system in maize, Genetics 129 (1991) 261–270.
[61] M. Chatterjee, C. Martin, Tam3 produces a suppressible allele of the DAG locus of *Antirrhinum majus* similar to Mu-suppressible alleles of maize, Plant J. 11 (1997) 759–771.
[62] G. Hu, N. Yalpani, S.P. Briggs, G.S. Johal, A porphyrin pathway impairment is responsible for the phenotype of a dominant disease lesion mimic mutant of maize, Plant Cell 10 (1998) 1095–1105.
[63] X. Cui, A.P. Hsia, F. Liu, D.A. Ashlock, R.P. Wise, P.S. Schnable, Alternative transcription initiation sites and polyadenylation sites are recruited during Mu suppression at the rf2a locus of maize, Genetics 163 (2003) 685–698.
[64] B.P. May, H. Liu, E. Vollbrecht, L. Senior, P.D. Rabinowicz, D. Roh, X. Pan, L. Stein, M. Freeling, D. Alexander, R. Martienssen, Maize-targeted mutagenesis: A knockout resource for maize, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 11541–11546.
[65] L. Girard, M. Freeling, Mutator-suppressible alleles of rough sheath1 and liguleless3 in maize reveal multiple mechanisms for suppression, Genetics 154 (2000) 437–446.
[66] M. Cowley, R.J. Oakey, Transposable elements re-wire and fine-tune the transcriptome, PLoS Genet. 9 (2013), e1003234.
[67] D. Lisch, How important are transposons for plant evolution? Nat. Rev. Genet. 14 (2013) 49–61.
[68] R.A. Elbarbary, B.A. Lucas, L.E. Maquat, Retrotransposons as regulators of gene expression, Science 351 (2016) aac7247.
[69] L. Hu, N. Li, C. Xu, S. Zhong, X. Lin, J. Yang, T. Zhou, A. Yuliang, Y. Wu, Y.R. Chen, X. Cao, A. Zemach, S. Rustgi, D. von Wettstein, B. Liu, Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 10642–10647.
[70] B. Greene, R. Walko, S. Hake, Mutator insertions in an intron of the maize knotted1 gene result in dominant suppressible mutations, Genetics 138 (1994) 1275–1285.
[71] R.G. Schneeberger, P.W. Becraft, S. Hake, M. Freeling, Ectopic expression of the knox homeo box gene rough sheath1 alters cell fate in the maize leaf, Genes Dev. 9 (1995) 2292–2304.
[72] G.J. Muehlbauer, J.E. Fowler, L. Girard, R. Tyers, L. Harper, M. Freeling, Ectopic expression of the maize homeobox gene liguleless3 alters cell fates in the leaf, Plant Physiol. 119 (1999) 651–662.
[73] S. Kobayashi, N. Goto-Yamamoto, H. Hirochika, Retrotransposon-induced mutations in grape skin color, Science 304 (2004) 982.
[74] J. Doebley, A. Stec, C. Gustus, Teosinte Branched1 and the Origin of Maize: Evidence for Epistasis and the Evolution of Dominance, Genetics 141 (1995) 333–346.
[75] A. Studer, Q. Zhao, J. Ross-Ibarra, J. Doebley, Identification of a functional transposon insertion in the maize domestication gene tb1, Nat. Genet. 43 (2011) 1160–1163.
[76] Q. Yang, Z. Li, W. Li, L. Ku, C. Wang, J. Ye, K. Li, N. Yang, Y. Li, T. Zhong, J. Li, Y. Chen, J. Yan, X. Yang, M. Xu, CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 16969–16974.
[77] S. Salvi, G. Sponza, M. Morgante, D. Tomes, X. Niu, K.A. Fengler, R. Meeley, E.V. Ananiev, S. Svitashev, E. Bruggemann, B. Li, C.F. Hainey, S. Radovic, G. Zaina, J.A. Rafalski, S.V. Tingey, G.H. Miao, R.L. Phillips, R. Tuberosa, Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 11376–11381.
[78] S. Castelletti, R. Tuberosa, M. Pindo, S. Salvi, A MITE Transposon Insertion Is Associated with Differential Methylation at the Maize Flowering Time QTL Vgt1, G3 (Bethesda) (2014).
[79] J.D. Hollister, B.S. Gaut, Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression, Genome Res. 19 (2009) 1419–1428.
[80] F. Maumus, H. Quesneville, Ancestral repeats have shaped epigenome and genome composition for millions of years in *Arabidopsis thaliana*, Nat. Commun. 5 (2014) 4104.
[81] I. Ahmed, A. Sarazin, C. Bowler, V. Colot, H. Quesneville, Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in Arabidopsis, Nucleic Acids Res. 39 (2011) 6919–6931.
[82] H. Saze, T. Kakutani, Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1, EMBO J. 26 (2007) 3641–3652.
[83] A. Martin, C. Troadec, A. Boualem, M. Rajab, R. Fernandez, H. Morin, M. Pitrat, C. Dogimont, A. Bendahmane, A transposon-induced epigenetic change leads to sex determination in melon, Nature 461 (2009) 1135–1138.
[84] F. Zeng, B. Cheng, Transposable Element Insertion and Epigenetic Modification Cause the Multiallelic Variation in the Expression of FAE1 in *Sinapis alba*, Plant Cell 26 (2014) 2648–2659.
[85] B.P. Williams, D. Pignatta, S. Henikoff, M. Gehring, Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat, PLoS Genet. 11 (2015), e1005142.
[86] M. Lei, H. Zhang, R. Julian, K. Tang, S. Xie, J.K. Zhu, Regulatory link between DNA methylation and active demethylation in Arabidopsis, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 3553–3557.
[87] T. Kinoshita, A. Miura, Y. Choi, Y. Kinoshita, X. Cao, S.E. Jacobsen, R.L. Fischer, T. Kakutani, One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation, Science 303 (2004) 521–523.
[88] W.J. Soppe, S.E. Jacobsen, C. Alonso-Blanco, J.P. Jackson, T. Kakutani, M. Koornneef, A.J. Peeters, The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene, Mol. Cell 6 (2000) 791–802.
[89] Y. Kinoshita, H. Saze, T. Kinoshita, A. Miura, W.J. Soppe, M. Koornneef, T. Kakutani, Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats, Plant J. 49 (2007) 38–45.
[90] M. Gehring, K.L. Bubb, S. Henikoff, Extensive demethylation of repetitive elements during seed development underlies gene imprinting, Science 324 (2009) 1447–1451.
[91] T.F. Hsieh, C.A. Ibarra, P. Silva, A. Zemach, L. Eshed-Williams, R.L. Fischer, D. Zilberman, Genome-wide demethylation of Arabidopsis endosperm, Science 324 (2009) 1451–1454.
[92] S. Feng, S.E. Jacobsen, W. Reik, Epigenetic reprogramming in plant and animal development, Science 330 (2010) 622–627.
[93] M. Gehring, Genomic imprinting: insights from plants, Annu. Rev. Genet. 47 (2013) 187–208.
[94] M. Zhang, S. Xie, X. Dong, X. Zhao, B. Zeng, J. Chen, H. Li, W. Yang, H. Zhao, G. Wang, Z. Chen, S. Sun, A. Hauck, W. Jin, J. Lai, Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize, Genome Res. 24 (2014) 167–176.
[95] D. Pignatta, R.M. Erdmann, E. Scheer, C.L. Picard, G.W. Bell, M. Gehring, Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting, Elife 3 (2014), e03198.
[96] T.N. Le, U. Schumann, N.A. Smith, S. Tiwari, P.C. Au, Q.H. Zhu, J.M. Taylor, K. Kazan, D.J. Llewellyn, R. Zhang, E.S. Dennis, M.B. Wang, DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis, Genome Biol. 15 (2014) (458–014–0458-3).
[97] D. Secco, C. Wang, H. Shou, M.D. Schultz, S. Chiarenza, L. Nussaume, J.R. Ecker, J. Whelan, R. Lister, Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements, Elife 4 (2015), http://dx.doi.org/10.7554/eLife.09343.
[98] S.B. Carroll, Endless forms: the evolution of gene regulation and morphological diversity, Cell 101 (2000) 577–580.
[99] P. Capy, G. Gasperi, C. Biemont, C. Bazin, Stress and transposable elements: co-evolution or useful parasites? Heredity (Edinb.) 85 (Pt 2) (2000) 101–106.

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010

[100] M.A. Grandbastien, C. Audeon, E. Bonnivard, J.M. Casacuberta, B. Chalhoub, A.P. Costa, Q.H. Le, D. Melayah, M. Petit, C. Poncet, S.M. Tam, M.A. Van Sluys, C. Mhiri, Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae, Cytogenet. Genome Res. 110 (2005) 229–241.
[101] E. Bucher, J. Reinders, M. Mirouze, Epigenetic control of transposon transcription and mobility in Arabidopsis, Curr. Opin. Plant Biol. 15 (2012) 503–510.
[102] H. Ito, Small RNAs and regulation of transposons in plants, Genes Genet. Syst. 88 (2013) 3–7.
[103] H. Hirochika, Retrotransposons of rice: their regulation and use for genome analysis, Plant Mol. Biol. 35 (1997) 231–240.
[104] S.M. Kaeppler, H.F. Kaeppler, Y. Rhee, Epigenetic aspects of somaclonal variation in plants, Plant Mol. Biol. 43 (2000) 179–188.
[105] K. Naito, F. Zhang, T. Tsukiyama, H. Saito, C.N. Hancock, A.O. Richardson, Y. Okumoto, T. Tanisaka, S.R. Wessler, Unexpected consequences of a sudden and massive transposon amplification on rice gene expression, Nature 461 (2009) 1130–1134.
[106] K. Yasuda, M. Ito, T. Sugita, T. Tsukiyama, H. Saito, K. Naito, M. Teraishi, T. Tanisaka, Y. Okumoto, Utilization of transposable element as a novel genetic tool for modification of the stress response in rice, Mol. Breed. 32 (2013) 505–516.
[107] R.K. Slotkin, M. Vaughn, F. Borges, M. Tanurdzic, J.D. Becker, J.A. Feijo, R.A. Martienssen, Epigenetic reprogramming and small RNA silencing of transposable elements in pollen, Cell 136 (2009) 461–472.
[108] S. Bire, S. Casteret, B. Piegu, L. Beauclair, N. Moire, P. Arensbuger, Y. Bigot, Mariner Transposons Contain a Silencer: Possible Role of the Polycomb Repressive Complex 2, PLoS Genet. 12 (2016), e1005902.
[109] B. McClintock, Further Studies of Gene-Control Systems in Maize, Carnegie Inst. Wash. Yearb. 62 (1963) 486–493.
[110] B. McClintock, The significance of responses of the genome to challenge, Science 226 (1984) 792–801.
[111] B. McClintock, Controlling elements and the gene, Cold Spring Harb. Symp. Quant. Biol. 21 (1956) 197–216.
[112] H. Ito, H. Gaubert, E. Bucher, M. Mirouze, I. Vaillant, J. Paszkowski, An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress, Nature 472 (2011) 115–119.
[113] V.V. Cavrak, N. Lettner, S. Jamge, A. Kosarewicz, L.M. Bayer, O. Mittelsten Scheid, How a retrotransposon exploits the plant's heat stress response for its activation, PLoS Genet. 10 (2014), e1004115.
[114] I. Makarevitch, A.J. Waters, P.T. West, M. Stitzer, C.N. Hirsch, J. Ross-Ibarra, N.M. Springer, Transposable elements contribute to activation of maize genes in response to abiotic stress, PLoS Genet. 11 (2015), e1004915.
[115] P. Bundock, P. Hooykaas, An Arabidopsis hAT-like transposase is essential for plant development, Nature 436 (2005) 282–284.
[116] Z. Joly-Lopez, E. Forczek, D.R. Hoen, N. Juretic, T.E. Bureau, A gene family derived from transposable elements during early angiosperm evolution has reproductive fitness benefits in *Arabidopsis thaliana*, PLoS Genet. 8 (2012), e1002931.
[117] A.D. McCue, S. Nuthikattu, R.K. Slotkin, Genome-wide identification of genes regulated in trans by transposable element small interfering RNAs, RNA Biol. 10 (2013) 1379–1395.

Please cite this article as: C.D. Hirsch, N.M. Springer, Transposable element influences on gene expression in plants, Biochim. Biophys. Acta (2016), http://dx.doi.org/10.1016/j.bbagrm.2016.05.010