Review

Reuben M. Buckley and David L. Adelson*

Mammalian genome evolution as a result of epigenetic regulation of transposable elements

Abstract: Transposable elements (TEs) make up a large proportion of mammalian genomes and are a strong evolutionary force capable of rewiring regulatory networks and causing genome rearrangements. Additionally, there are many eukaryotic epigenetic defense mechanisms able to transcriptionally silence TEs. Furthermore, small RNA molecules that target TE DNA sequences often mediate these epigenetic defense mechanisms. As a result, epigenetic marks associated with TE silencing can be reestablished after epigenetic reprogramming – an event during the mammalian life cycle that results in widespread loss of parental epigenetic marks. Furthermore, targeted epigenetic marks associated with TE silencing may have an impact on nearby gene expression. Therefore, TEs may have driven species evolution *via* their ability to heritably alter the epigenetic regulation of gene expression in mammals.

Keywords: epigenetics; genome evolution; mammals; transposable elements.

DOI 10.1515/bmc-2014-0013
Received April 7, 2014; accepted May 27, 2014

List of abbreviations: AGO2, Argonaute 2; APOBEC3, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3; ChIP, chromatin immunoprecipitation; ChIP-seq, ChIP sequencing; CTCF, CCCTC binding factor; DNMT1, DNA methyltransferase 1; DNMT3a, DNA methyltransferase 3a; DNMT3b, DNA methyltransferase 3b; DNMT3L, DNA methyltransferase 3-like; dsRNA, double-stranded RNA; ERVs, endogenous retroviruses; ESC, embryonic stem cell; ESR1, estrogen receptor 1; G9a, euchromatic histone-lysine *N*-methyltransferase 2; H3K27ac, histone H3 lysine 27 acetylation; H3K27me3, histone H3 lysine 27 methylation 3; H3K4me1, histone H3 lysine 4 methylation 1; H3K4me3, histone H3 lysine 4 methylation 3; H3K9, histone H3 lysine 9; H3K9me2/3, histone H3 lysine 9 methylation 2/3; IAP, intracisternal A particle; kb, kilobases; LINEs, long interspersed nuclear elements; LTR, long terminal repeat; MILI, piwi-like RNA-mediated gene silencing 2; miRNA, micro-RNA; MIWI2, piwi-like RNA-mediated gene silencing 4; MOV10, Moloney leukemia virus 10, homologue; MOV10L1, MOV10-like 1; ORFs, open reading frames; PGCs, primordial germ cells; piRNAs, PIWI-interacting RNAs; Pld6, phospholipase D family, member 6; priRNAs, primary RNAs; RdDM, RNA-directed DNA methylation; RNAi, RNA interference; RNP, ribonucleoprotein; SETDB1, SET domain bifurcated 1; SINEs, short interspersed nuclear elements; siRNA, small interfering RNA; sRNA, small RNA; Suv39, suppressor of variegation 3–9; TEs, transposable elements; TFBSs, transcription factor binding sites; TFs, transcription factors; UHRF1, ubiquitin-like with PHD and ring finger domains 1.

Introduction

Transposable elements (TEs) are mobile DNA segments that have had an extensive effect on mammalian genome evolution (1). As much as two thirds of the human genome may be composed of repetitive sequences, of which TE-derived sequences are a major component (2). Because of their ability to replicate themselves and their potential to cause mutation *via* insertional mutagenesis or ectopic recombination resulting in large genomic rearrangements, TEs have long been thought of as selfish genetic elements (3–5). This view is also consistent with the significant role TEs have been shown to play in various diseases (6, 7). Genome defense from TEs is largely mediated by transcriptional silencing. This is achieved by epigenetic modifications that disrupt the accessibility of the necessary transcriptional machinery. Both DNA methylation and

***Corresponding author: David L. Adelson,** School of Molecular and Biomedical Science, University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia,
e-mail: david.adelson@adelaide.edu.au
Reuben M. Buckley: School of Molecular and Biomedical Science, University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia

histone modifications are involved in these processes and are mediated by RNA intermediates (8–11).

However, evidence is emerging that suggest TEs are more than just genomic parasites. TE insertions have been shown to affect nearby gene expression in a variety of ways. Examples include TEs providing alternative splice sites, transcription factor binding sites (TFBSs), and alternative promoters for genes [reviewed in (1)]. Interestingly, epigenetic silencing mechanisms associated with TEs also affect gene expression. A well-studied example of this phenomenon is epigenetic inheritance at the *axin-fused* allele in which a kinky tail phenotype associates with differential methylation of the long terminal repeat (LTR) at the 3′ end of an intracisternal A particle (IAP) element in mice. Hypermethylation of the 3′ LTR of the IAP element was shown to suppress the kinked tail phenotype by silencing a cryptic promoter. Crosses between penetrant and silent *axin-fused* mice with null mice showed that the penetrance of the allele was inherited. This implied that the epigenetic methylation state of the IAP element remained stable as it passed from one generation to the next. Furthermore, the epigenetic state of an individual's sperm cells reflected the epigenetic state of that individual's somatic cells, thereby providing a mechanism for inheritance (12).

In the above example, the epigenetic state associated with the IAP element was inherited by the next generation and also affected expression of the *axin-fused* allele. Transgenerational inheritance of stable patterns, such as DNA sequences, provides the foundations on which evolutionary processes such as natural selection act. Although IAP elements appear to be an exception to most TEs, due to their ability to avoid epigenetic reprograming, the RNA intermediates that target TEs for transcriptional silencing through epigenetic modifications provide a mechanism by which epigenetic patterns associated with TEs can be inherited (13, 14). In this context, TEs can be thought of as providing a unique epigenetic environment. Throughout this review, we explore the role TEs have played in altering the epigenetic landscape, which in turn, may have altered gene expression patterns and regulatory networks and thereby driven evolution in different species.

The mammalian TE landscape

To understand the potential evolutionary impact of TEs, we must take into account the various types and families of TEs with different ages, mechanisms of action, distributions across species, and distributions within genomes. In this section, we discuss how each of these factors shapes the mammalian TE landscape.

According to the Repbase classification system, there are two main types of TEs: type 1 and type 2. Type 1 TEs consist of LTR and non-LTR retrotransposons. Many LTR retrotransposons in mammals are endogenous retroviruses (ERVs). ERVs have been grouped into several different classes based on such criteria as structural features and phylogeny [reviewed in (15)]. Non-LTR retrotransposons are made up of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (16). LINEs are usually several kilobases long and contain two open reading frames (ORFs), one of which encodes a ribonucleoprotein (RNP) that reverse transcribes the element and inserts the DNA copy into the genome. The copy and paste process of L1 retrotransposition is shown in Figure 1. However, SINEs are only approximately 300 base pairs (bp) long, contain no ORFs, and require the retrotransposition machinery encoded by LINE elements for retrotransposition. SINEs are derived from the 3′ end of LINEs, and these 3′ sequences bind the LINE-encoded RNP required for replication. LINEs and their derived SINEs are referred to as LINE-SINE pairs. In humans, LINE L1 and SINE Alu are an active LINE-SINE pair. In the mouse, a similar pairing also exists, where mouse LINE L1s form a LINE-SINE pair with SINE B1 elements (17). However, the majority of TE sequences in mammalian genomes are inactive. Type 2 TEs, also known as DNA transposons, are able to excise themselves from the genome and reinsert themselves elsewhere in the genome using a transposase encoded in their single ORF. Because of this cut-and-paste mobilization that does not generate additional copies, type 2 TEs are found in much lower numbers than type 1

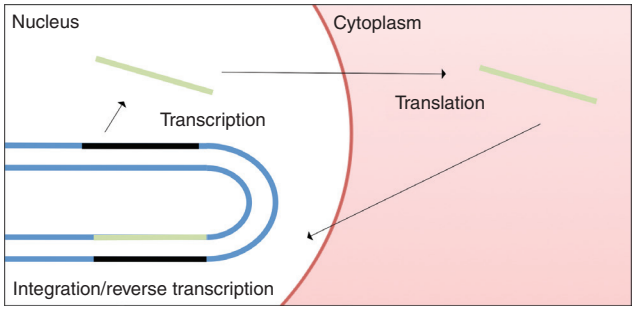

Figure 1 Retrotransposition.
Class 1 non-LTR TEs (black) are transcribed (green) and transported to the cytoplasm. Within the cytoplasm, non-autonomous TEs undergo translation and produce an RNP. TE transcripts are transported back into the nucleus where they are reverse transcribed and integrated into the genome. The above process has resulted in large portions of the genome comprising of repeated DNA sequences.

TEs in mammalian genomes. The above findings have previously been reviewed by Jurka et al. (18).

The mammalian TE landscape is very complex, and every species of mammal has both shared and unique TEs that can be traced back to various lineages within the mammalian radiation. Initial genome-wide studies of TE distribution based on the human genome concluded that LINE L1s were more prevalent in AT-rich regions; SINE MIRs and SINE Alus showed a preference for GC-rich regions; and LINE L2s were distributed independent of GC content (19, 20). LINE L1s, LINE L2s, and SINE Alus also all had a preference for antisense insertions within genes; this was most pronounced for LINE L1s. SINE MIR sequences, however, showed no such insertion preference. This observation was interpreted as the result of selection against LINE L1 sense insertions because the LINE L1 element's poly A signal/tail may cause shortened gene transcripts (19). Therefore, the observed TE distribution is the product of both TE insertion preference and selection against specific types of TE insertions (19). In the mouse, the TE landscape is very different. For example, young SINE B1 and SINE B2 elements insert into SINE-rich GC areas, whereas young SINE Alus in human insert into SINE Alu-poor AT-rich areas. Human and mouse also differ in retrotransposon content. For example, the human genome has fewer LTR/ERVs compared with the mouse genome, and the mouse genome has far fewer SINE MIRs and LINE L2s than the human genome (21).

Throughout the mammalian lineage, older TEs show signs of being retained, which result from selective pressures. For example, SINE MIRs and LINE L2s and TE-free regions are often found in conserved orthologous segments between human and mouse (22, 23). Moreover, subsequent analyses of repeat families in different species have adopted a more global approach to identify associations of repetitive elements in different families across species. This led to the identification of regions enriched for ancestral repeats (SINE MIR and LINE L2) in human, horse, and cow. Therefore, ancestral mammalian TEs show signs of both positional and sequence conservation in a number of species (24, 25). This conservation suggests a role for these repeats in the genome structure associated with the regulation of gene expression.

Although distantly related species have been used to compare the distribution of inactive and ancestral repeats, comparisons between closely related species have been used to compare distributions of young, active TEs. Deep sequencing of 17 strains of mouse revealed over 100,000 TE variants, each of which had survived selection over the past 2 million years. The ERV family of repeats underwent the largest expansion, and deleterious ERVs were rapidly purged from the mouse genome. Deleterious LINE L1s were also purged but not quite as rapidly as ERVs. ERV insertions were also shown to be most highly associated with changes in gene expression between the mouse strains (26). It is clear that TEs are a source of variation among species and can cause large genomic changes. However, most such changes are detrimental, and it is therefore advantageous to be able to reduce the probability of potentially detrimental changes.

Silencing of TEs *via* targeted epigenetic mechanisms

TE silencing through DNA methylation and chromatin modification keeps retrotransposition in check by suppressing TE transcription. However, during germ cell early embryonic development, DNA and histone methylation patterns are transiently erased, allowing TEs to mobilize. However, mobilizing TEs are quickly inhibited by targeted small RNA (sRNA) TE-silencing mechanisms (27, 28). After this transient demethylation, DNA methyltransferase 3a (DNMT3a) and DNA methyltransferase 3-like (DNMT3L) specifically methylate TEs, thereby suppressing TE transcription (8, 29, 30).

The most well-characterized sRNA-targeting mechanism for TE transcriptional silencing in mammals is the PIWI-based recognition system (27). sRNA molecules approximately 26–31 nucleotides long direct DNA methylation at TE promoters in a general process known as RNA-directed DNA methylation (RdDM), and these RNA molecules are known as PIWI-interacting RNAs (piRNAs) (11). Primary piRNAs are generated from piRNA clusters during widespread TE transcription during epigenetic reprogramming. Primary piRNAs then bind to piwi-like RNA-mediated gene silencing 2 (MILI) to form complexes that then bind to and cleave the TE transcripts. The cleaved transcript product is a secondary piRNA that forms a complex with piwi-like RNA-mediated gene silencing 4 (MIWI2) and targets the primary piRNA cluster transcript, thereby leading to increased production of primary piRNAs. This process is known as a 'ping-pong' amplification cycle and is very effective in dealing with large numbers of TE transcripts (Figure 2). The PIWI-based recognition system occurs before DNMT3L-guided methylation and is believed to be the causal factor in TE methylation specificity (11). RNA intermediates may also be involved in directing chromatin modifications that can silence TE transcription. This idea is well established in plants and supported in *Drosophila* but has not

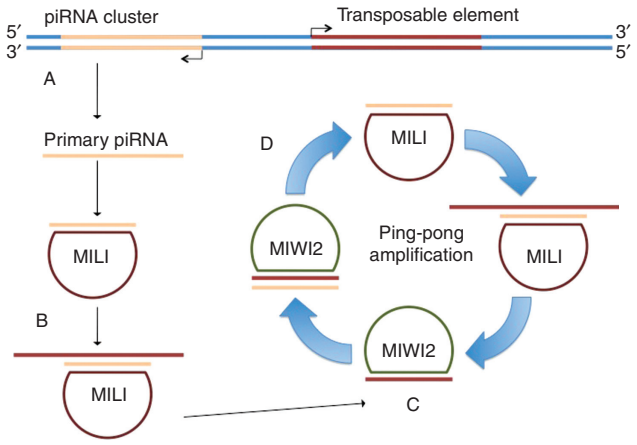

Figure 2 piRNAs and ping-pong amplification.
(A) piRNAs targeting TEs are transcribed from piRNA clusters. (B) Primary piRNAs are processed and loaded into MILI where they are able to guide MILI to TE transcripts, initiating the formation of the ping-pong amplification pathway. (C) Cleaved TE transcripts become secondary piRNAs and are loaded into MIWI2. (D) The complex then targets and cleaves primary piRNA clusters, resulting in the generation of more piRNAs loaded into MILI. piRNAs generated by this process are believed to drive RdDM by an unknown mechanism.

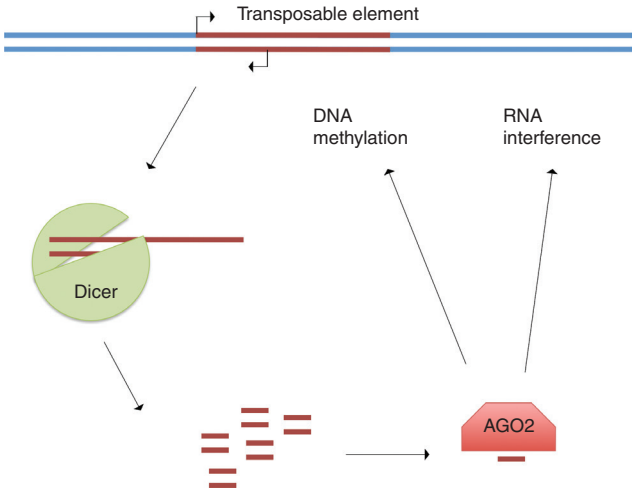

Figure 3 TE silencing *via* sRNA molecules.
Double-stranded TE RNA is produced as a result of transcription from bidirectional TE promoters. This dsRNA is then targeted and cleaved by DICER. The resulting sRNAs are then loaded into AGO2 and direct DNA methylation or RNA interference.

been confirmed in mammals (31, 32). piRNA sequences are found in clusters throughout the genome and share sequence similarity with TEs. The production of piRNAs corresponding to a specific TE is likely the result of a TE insertion into a piRNA cluster (33, 34). Although piRNAs play a large role in silencing TEs, some TE families are effectively silenced even in their absence. For example, SINE B1 elements in mouse have a locus-to-locus variation in their methylation patterns. Knockout of phospholipase D family, member 6 (*Pld6*), and *MILI* genes, which are involved in piRNA biogenesis, results in disrupted piRNA-mediated silencing of LINE L1 elements, whereas methylated SINE B1 elements remain methylated in spermatogonia (35, 36). The knockouts also show no increase in SINE B1 expression, indicating that SINE B1 silencing occurs independently of piRNA activity (36).

Recent work on LINE L1 silencing shows the potential involvement of other RdDM mechanisms in mammals. Mammalian micro-RNAs (miRNAs) associated with repeats are 22 nucleotides long and are products of double-stranded RNA (dsRNA) cleaved by DICER and loaded into Argonaute 2 (AGO2) (37, 38) (Figure 3). Mouse embryonic stem cell (ESC) *DICER* knockouts showed that mammalian repeat-associated miRNAs were depleted, LINE L1 promoter elements were hypomethylated, and that LINE L1 transcription, translation, and copy number had increased. Therefore, components of miRNA biogenesis in mammals are linked to transcriptional silencing of TEs (39, 40). Furthermore, sRNAs in mammals are involved with RNA interference (RNAi) or post-transcriptional silencing of TEs. dsRNA processed by DICER caused sRNAs in mammals to form perfect small interfering RNA (siRNA) duplexes with two-nucleotide 3′ overhangs, a characteristic associated with RNAi in other systems (40).

Post-transcriptional processing of TE transcripts involves several other regulators. The microprocessor, a multiprotein complex able to recognize and cleave primary RNAs (priRNAs), plays an important role in miRNA biogenesis (41, 42). The microprocessor recognizes structures within LINE L1 elements and promotes their degradation (43). Another regulator of TE activity through RNAi is the human RNA helicase, Moloney leukemia virus 10, homologue (MOV10) (44, 45). MOV10 is part of a multiprotein complex with other components involved in RNA-induced silencing (46). MOV10 prevents retrotransposition of non-autonomous TEs by interacting with the LINE L1 RNP (45). Furthermore, a protein related to MOV10 known as MOV10-like 1 (MOV10L1), interacts with piRNA proteins in male mouse germ cells and is involved in transcriptional silencing of certain TE families (47, 48). MOV10 is also found in complexes with members of the apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3 (APOBEC3) family (APOBEC3G and APOBEC3F) and is associated with defense against retroviruses, which have replication mechanisms similar to retrotransposons (49, 50). Therefore, APOBEC3 proteins may also be involved in various processes that protect host genomes against TEs (51). The APOBEC3 family of proteins is a family of

cytidine deaminases that convert cytidine to uridine to edit retrotransposon DNA and cDNA as a defense against retrotransposition (51).

Histone modifications provide another mechanism to regulate TE expression. Most active TE sequences are associated with histone H3 lysine 9 (H3K9) methylation and are therefore transcriptionally repressed (52). For example, mutations in methyltransferases that are associated with repressive histone modifications lead to increased TE activity (53). In mouse early embryogenesis, a methyltransferase known as SET domain bifurcated 1 (SETDB1) targets specific promoter-proximal class I and class II ERVs. Embryonic cells lacking SETDB1 show transcription of the promoter-proximal ERVs in the form of aberrant gene transcripts that would otherwise be silenced. Therefore, SETDB1 is involved with transcriptional control of TEs independent of DNA methylation (10). Another methyltransferase found in mouse involved in TE silencing is euchromatic histone-lysine *N*-methyltransferase 2 (G9a) (54). G9a does not appear to be involved with silencing class I and II ERVs but is necessary for silencing class III ERVs (55). Suppressor of variegation 3–9 (Suv39) is another H3K9 methyltransferase also linked to TE silencing; deletions of *Suv39* result in a moderate increase in TE activity (52). Furthermore, the above mechanisms may hold for a variety of mammals including the pig. For example, porcine ERVs are silenced by similar chromatin modifications as seen in mouse (56). However, heterochromatic silencing during embryogenesis may not be an active driver of TE silencing. Moreover, TE silencing usually occurs after loss of an active histone mark and before gain of a repressive histone mark (57).

Interestingly, heterochromatin modifications associated with TE sequences may be selected to play a dual role, resulting in further downstream implications of TE accumulation. Generally, heterochromatin causes a transcriptionally repressive environment-reducing TE activity. However, heterochromatic regions are also unable to undergo recombination (3). This is important because unregulated TE genomic-enriched regions are particularly prone to hazardous recombination events that have been linked to disease in humans (5, 7). Therefore, prevention of non-homologous recombination may be a driving force behind heterochromatic repression of TEs. In addition, recombination also often results in TE deletion, implying that heterochromatic silencing may be the cause of TE accumulation (4). Furthermore, simulations have shown that under an ectopic recombination model, TEs accumulate in regions of low recombination (58).

This section shows how epigenetic mechanisms are involved in the regulation of TEs. Epigenetics are now known to contribute to many regulatory processes, especially throughout development. The following section aims to show the breadth of developmental regulation under the control of epigenetic processes in the context of the regulatory impact of retrotransposition.

Epigenetic regulatory mechanisms are essential for mammalian development

Epigenetic mechanisms are well characterized in terms of the roles they play in development. Mammalian development is highly complex and requires extensive regulation of intricate cellular processes. During development, the differentiation potential of cells is gradually reduced at each stage until cells terminally differentiate. This reduction of differentiation potential is largely regulated by epigenetics.

DNA methylation is the critical modification of DNA with respect to the epigenetic regulation of transcription. Specifically, DNA methylation refers to the methylation of cytosine and occurs mostly in CpG sequences (59). In mammals, approximately 60%–80% of CpGs are methylated. However, approximately 10% of CpGs are resistant to methylation and are found in GC-rich regions of the genome. These CpG sites are known as CpG islands and are found in gene and retrotransposon promoters (28, 60). The DNA methyltransferases, DNA methyltransferase 1 (DNMT1), DNMT3a, and DNA methyltransferase 3b (DNMT3b) all have roles in maintaining DNA methylation throughout mammalian development (61, 62). The deletion of *DNMT1* in ESCs results in apoptosis, whereas simultaneous deletion of *DNMT3a* and *DNMT3b* did not affect survivability yet resulted in ESCs unable to differentiate (63). Once established, methylation patterns are able to persist through multiple rounds of mitosis. During the S phase, DNMT1 directly interacts with proliferating cell nuclear antigen and ubiquitin-like with PHD and ring finger domains 1 (UHRF1); this complex is recruited to sites of DNA replication and binds hemi-methylated DNA *via* a SET- and RING-associated domain (64–66). UHRF1 binds to parental methylated DNA and thereby directs DNMT1 to the daughter strand (28, 66). Therefore, DNA methylation is a stable process for transmitting epigenetic regulatory information from parent to daughter cell, unlike transmission of epigenetic regulatory information from parent to offspring at the level of multicellular organisms.

Inherited information from a parent to offspring is largely mediated by one cell, a single gamete. Gametes from each parent fuse to form a zygote, which then develops into an organism made up of a large variety of tissues and differentiated cell types. For this process to occur properly, there are two stages during development where cells undergo epigenetic reprogramming resulting in global hypomethylation. The processes governing how epigenetic patterns are reestablished during development are complex and remain an area of intense research. In ESCs and primordial germ cells (PGCs), epigenetic states are reset requiring that methylation patterns are reestablished in a targeted manner for differentiation to occur. Various DNA methylation target sites have been identified. These include promoters, pericentromic repeats, TEs, and imprint control regions (28).

Another form of epigenetic regulation during mammalian development is through histone modifications. Histone proteins form a complex with DNA called a nucleosome, in which approximately 147 nucleotides of DNA are wrapped around the nucleosomal histones H2A, H2B, H3, and H4. Two copies of each histone make up the nucleosome, and a collection of nucleosomes results in the formation of chromatin. Each one of these histones can also be chemically modified, usually by a methyltransferase or an acetylase. Chemical modifications of histone proteins regulate the accessibility of surrounding DNA (67). For example, repressive histone modifications cause nucleosomes to tightly associate, resulting in the surrounding DNA being made inaccessible and transcriptionally silent. Known repressive histone modifications include histone H3 lysine 9 methylation 2/3 (H3K9me2/3) and histone H3 lysine 27 methylation 3 (H3K27me3) (68–70). Meanwhile, active histone modifications can cause the nucleosomes to dissociate, resulting in the surrounding DNA becoming accessible to transcriptional machinery. Active histone modifications at promoters include histone H3 lysine 4 methylation 3 (H3K4me3), histone H3 lysine 27 acetylation (H3K27ac), and known modifications at enhancers include H3K27ac and histone H3 lysine 4 methylation 1 (H3K4me1) (70–72). These histone modifications cause a change in chromatin status at particular loci; however, they are not as stable as DNA methylation (67). Therefore, the DNA loci associated with histone modifications that are analogous to DNA methylation are described as being repressed rather than silenced (28, 73). Interestingly, histone modifications in PGCs and ESCs contribute to the transcriptionally permissive environment characteristic of these cell types and the reductions in DNA methylation they experience. For example, global loss of repressive H3K9 methylation marks are an essential step in epigenetic reprogramming in PGCs and induced pluripotent stem cells (74, 75).

Throughout development, most histone modifications remain dynamic as various genes are switched on and off. However, some loci, including some TE loci, appear to have a stable repertoire of histone modifications (73, 76, 77). These modifications, like DNA methylation, may be due to targeted mechanisms. As a result, TEs located next to the promoter of a gene can affect the epigenetic regulation of that gene's promoter. Therefore, new TE insertions are able to change the regulation of a gene.

TE DNA sequences are more than just repressors

A large body of evidence shows that TEs can cause large changes to gene regulatory networks. However, not all of these involve transcriptional silencing.

One of the ways TEs alter regulatory networks is through the binding of transcription factors (TFs). Using chromatin immunoprecipitation (ChIP), Bourque et al. (78) showed that several TFs had binding sites within specific TE families. Additionally, TEs with a particular TFBS were more likely to be adjacent to genes regulated by that TF than genes not regulated by that TF. One of the TFs analyzed was estrogen receptor 1 (ESR1) and was bound to MIR elements and ERV-like elements. Moreover, subsequent analyses showed these elements also harbored TFBS motifs for ESR1 co-regulators (79), thereby strengthening the idea that TFBSs from TEs affect gene expression, as the control of gene expression usually requires binding of multiple TFs (80). Further implications of combinatorial TF binding patterns found in TEs have also been linked to the evolution of particular traits. For example, MER20 is a eutherian-specific TE and is located within 200 bp of 13% of the genes associated with pregnancy in mammals (81). Of 21 randomly selected MER20s, 14 were shown to bind combinations of TFs associated with insulator activity and four were shown to bind combinations of TFs with repressor functions.

Recently, species comparisons have yielded even further insight into how TEs are able to alter regulatory networks through changes in TFBS repertoire. Schmidt et al. (82) showed that expansion of CCCTC binding factor (CTCF) binding sites in various mammalian lineages was likely due to TE expansion. CTCF is a zinc-finger protein that is able to bind DNA at a highly conserved DNA binding motif and is involved in looping DNA in long-range interactions (82–84). ChIP sequencing (ChIP-seq) characterization of CTCF binding sites in five mammalian species: human, macaque, mouse, rat, and dog showed a shared core of approximately 5000 CTCF binding sites. However, there

were also large numbers of species-specific binding sites, and many of the species-specific binding sites in mouse, rat, and dog mapped to lineage-specific TEs (both shared and unique SINE B2 elements in mouse and rat and SINEC cf elements in dog) (82). Like many of the combinations of TFs that bind to MER20s, CTCF is also a known insulator protein. Insulator proteins cause changes in gene regulation by creating chromatin boundaries. CTCF is also sensitive to methylation, and this raises questions about the extent to which TEs are transcriptionally silenced and their ability to potentially escape transcriptional silencing (85, 86). It is clear that TF binding of TEs supports a role for TEs as a potent evolutionary force in mammals. However, it is likely that binding of TFs to binding sites embedded within TEs also alters the epigenetic landscape at the TE locus.

Instrumental in the discovery of the regulatory potential of MER20s was that MER20s were enriched for chromatin marks associated with insulator activity (81). This approach has also been used in identifying the regulatory potential of other TEs in human. For example, Xie et al. (87) analyzed genome-wide methylation patterns and found that LFSINE and LTR77 TE families were differentially methylated in various tissues. Both TE families were also associated with gene expression in a tissue-specific manner and had histone modifications representative of enhancers. These findings show that epigenetic regulation is not only involved in silencing the activity of TEs but also allows them to function as enhancers or insulator elements. We can therefore say that some epigenetic regulatory mechanisms override TE-silencing mechanisms, making it more likely that TEs that contain certain TFBS are able to effectively replicate within the genome.

Transcriptional epigenetic silencing may be a powerful driver of evolution

TEs have contributed significantly to mammalian evolution in a variety of ways. The silencing of newly inserted TEs may result in an epigenetic change at a particular locus, which could therefore result in large changes in nearby gene expression, thereby altering phenotypes subject to selection.

In plant systems, Hollister et al. (88) have established that TEs contribute to an epigenetic variation that results in differences in gene expression. However, although this has not been validated in mammals, many of the components that silence TEs in plants are conserved in mammals. Comparisons between *Arabidopsis thaliana* and *Arabidopsis lyrata* revealed that sRNA-targeted TEs were associated with reduced gene expression in both species and differences in gene expression between orthologues. In addition, it was reported that changes in gene expression due to TE silencing had deleterious effects resulting in the accumulation of silenced TEs in gene-poor regions (89). This result illustrates the degree by which gene expression can be altered through silencing of TE insertions. It is important to note that some eukaryotic mechanisms responsible for silencing TEs *via* sRNAs consist of largely conserved components. Plants, fungi, and animals all use sRNAs that are cleaved by DICER-type proteins and are then bound to Argonaut proteins that then either target transcripts for RNAi or target the appropriate DNA for DNA methylation (90) (Figure 3). It is therefore likely that the observations in plants will be replicated in mammals.

Although it has not yet been shown on a genome-wide scale how TE-associated epigenetic silencing mechanisms affect gene expression in mammals, epigenomes in several mammals have been mapped. Xiao et al. compared the epigenomes of pig, mouse, and human to gain further insight into the evolution of genome-wide epigenetic regulation. Results showed that the correlations between epigenetic and gene expression conservation were higher than the correlations between sequence and gene expression conservation (91). This approach reveals that the main driver of mammalian transcriptome evolution may in fact be changes to the epigenome rather than changes in DNA sequence. It is worth noting that while patterns of epigenetic chromatin modifications may differ between mammalian species, the mechanisms that regulate them are conserved (92). For instance, the level of conservation associated with the stability of histone modifications indicates regulation of histone modifications by conserved mechanisms (93). This means that species-specific TE families can cause the same kinds of epigenetic changes in different species. Epigenetic changes resulting from heritable TE insertions can alter gene expression and hence phenotype. Therefore, TEs and their associated silencing mechanisms may have exerted significant influence on the evolution of the mammalian transcriptome.

Expert opinion

The total impact of epigenetic regulation of TEs on mammalian evolution remains largely unexplored. TEs are a major component of genome architecture, and the extent of their impact can be vast. Comparative studies involving

TEs remain a challenge due to the complexity of analyzing many closely related sequences and the economic cost of generating transcriptome and epigenome data. However, by analyzing the genomic distribution of particular TE families and developing new techniques that can compare these distributions across different species, we may be able to better understand the impact of TEs on mammalian evolution. This kind of analysis merged with transcriptome, and epigenome data will help develop a deeper understanding of the evolutionary outcomes of mammalian genomes and TE families in regard to the epigenetic mechanisms that control TE mobilization.

Highlights

- TE families are classified based on a number of criteria and have discernible features. However, an understanding of TE insertion preferences remains elusive due to divergent genomic landscapes.
- Most hypotheses aimed at reconciling the distribution of TEs use a negative selection viewpoint, that is to say, TE insertions accumulate in areas where they would be least harmful.
- The role of epigenetics in regard to TEs is largely believed to be one of defending the genome against TEs. However, some findings show that epigenetic regulation of TEs may contribute to the control of gene expression.
- Instances in which TEs provide a binding site for a DNA methylation-sensitive TF may provide TEs with an opportunity to escape transcriptional silencing.
- Several mechanisms are believed to be involved in TE silencing. However, some of these mechanisms appear to only target specific TE families.
- sRNA-mediated silencing of TEs can provide a mechanism by which TEs can alter the epigenome and pass those alterations on to the next generation. However, this has not yet been confirmed in mammals.
- Comparative studies involving epigenetics are extremely scarce because of their expense. Despite this, strong correlations have been observed between conservation in epigenomes and conservation in transcriptomes.

Outlook

AS more and more genome data become available as a result of better sequencing technologies, our understanding of the nature of genetic regulation and genome architecture will improve. One of the current bottlenecks is that both DNA and RNA sequencing analysis require assembly of short reads, usually between <100 bp and 1 kb. Because TEs are often longer than reads, it is often impossible to assemble reads from TEs accurately. Fortunately, this problem will be eliminated with the advent of sequencing technologies that use longer reads (94) such as nanopores that have the ability to read single molecules and produce reads longer than 10 kb (95).

Acknowledgments: Thanks to Dan Kortschak for helpful discussions, advice, and criticism.

References

1. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet 2009; 10: 691–703.
2. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet 2011; 7: e1002384.
3. George CM, Alani E. Multiple cellular mechanisms prevent chromosomal rearrangements involving repetitive DNA. Crit Rev Biochem Mol Biol 2012; 47: 297–313.
4. Fedoroff NV. Presidential address. Transposable elements, epigenetics, and genome evolution. Science 2012; 338: 758–67.
5. Robberecht C, Voet T, Esteki MZ, Nowakowska BA, Vermeesch JR. Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. Genome Res 2013; 23: 411–8.
6. Hancks DC, Kazazian HH Jr. Active human retrotransposons: variation and disease. Curr Opin Genet Dev 2012; 22: 191–203.
7. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 elements in structural variation and disease. Annu Rev Genom Hum G 2011; 12: 187–215.
8. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 2004; 431: 96–9.
9. Peng JC, Karpen GH. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol 2007; 9: 25–35.
10. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M, Lorincz MC. DNA ethylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell 2011; 8: 676–87.
11. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 2008; 31: 785–99.
12. Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KVK, Whitelaw E. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci USA 2003; 100: 2538–43.

13. Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 2003; 35: 88–93.
14. Daxinger L, Whitelaw E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat Rev Genet 2012; 13: 153–62.
15. Blomberg J, Benachenhou F, Blikstad V, Sperber G, Mayer J. Classification and nomenclature of endogenous retroviral sequences (ERVs): problems and recommendations. Gene 2009; 448: 115–23.
16. Kapitonov VV, Jurka J. A universal classification of eukaryotic transposable elements implemented in Repbase. Nat Rev Genet 2008; 9: 411–2; author reply 4.
17. Ohshima K, Okada N. SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. Cytogenet Genome Res 2005; 110: 475–90.
18. Jurka J, Kapitonov VV, Kohany O, Jurka MV. Repetitive sequences in complex genomes: structure and evolution. Annu Rev Genomics Hum Genet 2007; 8: 241–59.
19. Medstrand P, van de Lagemaat LN, Mager DL. Retroelement distributions in the human genome: variations associated with age and proximity to genes. Genome Res 2002; 12: 1483–95.
20. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860–921.
21. Jurka J, Kohany O, Pavlicek A, Kapitonov VV, Jurka MV. Clustering, duplication and chromosomal distribution of mouse SINE retrotransposons. Cytogenet Genome Res 2005; 110: 117–23.
22. Silva JC, Shabalina SA, Harris DG, Spouge JL, Kondrashovi AS. Conserved fragments of transposable elements in intergenic regions: evidence for widespread recruitment of MIR- and L2-derived sequences within the mouse and human genomes. Genet Res 2003; 82: 1–18.
23. Simons C, Pheasant M, Makunin IV, Mattick JS. Transposon-free regions in mammalian genomes. Genome Res 2006; 16: 164–72.
24. Adelson DL, Raison JM, Edgar RC. Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. Proc Natl Acad Sci USA 2009; 106: 12855–60.
25. Adelson DL, Raison JM, Garber M, Edgar RC. Interspersed repeats in the horse (Equus caballus); spatial correlations highlight conserved chromosomal domains. Anim Genet 2010; 41: 91–9.
26. Nellaker C, Keane TM, Yalcin B, Wong K, Agam A, Belgard TG, Flint J, Adams DJ, Frankel WN, Ponting CP. The genomic landscape shaped by selection on transposable elements across 18 mouse strains. Genome Biol 2012; 13: R45.
27. Castaneda J, Genzor P, Bortvin A. piRNAs, transposon silencing, and germline genome integrity. Mutat Res 2011; 714: 95–104.
28. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet 2013; 14: 204–20.
29. Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 2002; 129: 1983–93.
30. Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum Mol Genet 2007; 16: 2272–80.
31. Olovnikov I, Aravin AA, Fejes Toth K. Small RNA in the nucleus: the RNA-chromatin ping-pong. Curr Opin Genet Dev 2012; 22: 164–71.
32. Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 2013; 27: 390–9.
33. Kelleher ES, Barbash DA. Analysis of piRNA-mediated silencing of active TEs in Drosophila melanogaster suggests limits on the evolution of host genome defense. Mol Biol Evol 2013; 30: 1816–29.
34. Yamamoto Y, Watanabe T, Hoki Y, Shirane K, Li YF, Ichiiyanagi K, Kuramochi-Miyagawa S, Toyoda A, Fujiyama A, Oginuma M, Suzuki H, Sado T, Nakano T, Sasaki H. Targeted gene silencing in mouse germ cells by insertion of a homologous DNA into a piRNA generating locus. Genome Res 2013; 23: 292–9.

35. Hu Q, Rosenfeld MG. Epigenetic regulation of human embryonic stem cells. Front Genet 2012; 3: 238.
36. Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA. Human embryonic stem cells as a model for studying epigenetic regulation during early development. Cell Cycle 2005; 4: 1323–6.
37. Maniataki E, Mourelatos Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes Dev 2005; 19: 2979–90.
38. Hammond SM. Dicing and slicing: the core machinery of the RNA interference pathway. FEBS Lett 2005; 579: 5822–9.
39. Faulkner GJ. Retrotransposon silencing during embryogenesis: dicer cuts in LINE. PLoS Genet 2013; 9: e1003944.
40. Ciaudo C, Jay F, Okamoto I, Chen CJ, Sarazin A, Servant N, Barillot E, Heard E, Voinnet O. RNAi-dependent and independent control of LINE1 accumulation and mobility in mouse embryonic stem cells. PLoS Genet 2013; 9: e1003791.
41. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature 2004; 432: 231–5.
42. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The Microprocessor complex mediates the genesis of microRNAs. Nature 2004; 432: 235–40.
43. Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyras E, Garcia-Perez JL, Cáceres JF. The Microprocessor controls the activity of mammalian retrotransposons. Nat Struct Mol Biol 2013; 20: 1173–81.
44. Meister G, Landthaler M, Peters L, Chen PY, Urlaub H, Luhrmann R, Tuschl T. Identification of novel argonaute-associated proteins. Current Biology 2005; 15: 2149–55.
45. Goodier JL, Cheung LE, Kazazian HH Jr. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. PLoS Genet 2012; 8: e1002941.
46. Chendrimada TP, Finn KJ, Ji XJ, Baillat D, Gregory RI, Liebhaber SA, Pasquinelli AE, Shiekhattar R. MicroRNA silencing through RISC recruitment of eIF6. Nature 2007; 447: 823–U1.
47. Zheng K, Xiol J, Reuter M, Eckardt S, Leu NA, McLaughlin KJ, Stark A, Sachidanandam R, Pillai RS, Wang PJ. Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. Proc Natl Acad Sci USA 2010; 107: 11841–6.
48. Frost RJA, Hamra FK, Richardson JA, Qi XX, Bassel-Duby R, Olson EN. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. Proc Natl Acad Sci USA 2010; 107: 11847–52.
49. Luo K, Wang T, Liu BD, Tian CJ, Xiao ZX, Kappes J, Yu XF. Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. J Virol 2007; 81: 7238–48.
50. Izumi T, Burdick R, Shigemi M, Plisov S, Hu WS, Pathak VK. Mov10 and APOBEC3G localization to processing bodies is not required for virion incorporation and antiviral activity. J Virol 2013; 87: 11047–62.
51. Chiu YL, Greene WC. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. Annu Rev Immunol 2008; 26: 317–53.
52. Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J 2005; 24: 800–12.
53. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007; 8: 272–85.
54. Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. EMBO J 2008; 27: 2681–90.
55. Maksakova IA, Thompson PJ, Goyal P, Jones SJ, Singh PB, Karimi MM, Lorincz MC. Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. Epigenet Chromatin 2013; 6: 15.
56. Wolf G, Nielsen AL, Mikkelsen JG, Pedersen FS. Epigenetic marking and repression of porcine endogenous retroviruses. J Gen Virol 2013; 94: 960–70.
57. Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, Carninci P, Torres-Padilla ME. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat Struct Mol Biol 2013; 20: 332–8.
58. Dolgin ES, Charlesworth B. The effects of recombination rate on the distribution and abundance of transposable elements. Genetics 2008; 178: 2169–77.
59. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci USA 2000; 97: 5237–42.
60. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev 2011; 25: 1010–22.
61. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999; 99: 247–57.
62. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992; 69: 915–26.
63. Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol Cell Biol 2004; 24: 8862–71.
64. Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 2007; 317: 1760–4.
65. Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature 2008; 455: 818–21.
66. Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH, Dhe-Paganon S. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. Nature 2008; 455: 822–5.
67. Felsenfeld G, Groudine M. Controlling the double helix. Nature 2003; 421: 448–53.
68. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 2002; 298: 1039–43.
69. Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 2001; 292: 110–3.
70. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 2011; 473: 43–U52.

71. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009; 459: 108–12.
72. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 2010; 107: 21931–6.
73. Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, Zhang X, Coyne M, Fostel JL, Holmes L, Meldrim J, Guttman M, Epstein C, Park H, Kohlbacher O, Rinn J, Gnirke A, Lander ES, Bernstein BE, Meissner A. Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell 2013; 153: 1149–63.
74. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 2008; 452: 877–81.
75. Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, Liang H, Chen Y, Chen J, Wu Y, Guo L, Zhu J, Zhao X, Peng T, Zhang Y, Chen S, Li X, Li D, Wang T, Pei D. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nat Genet 2013; 45: 34–42.
76. Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, Yang H, Wang T, Lee AY, Swanson SA, Zhang J, Zhu Y, Kim A, Nery JR, Urich MA, Kuan S, Yen CA, Klugman S, Yu P, Suknuntha K, Propson NE, Chen H, Edsall LE, Wagner U, Li Y, Ye Z, Kulkarni A, Xuan Z, Chung WY, Chi NC, Antosiewicz-Bourget JE, Slukvin I, Stewart R, Zhang MQ, Wang W, Thomson JA, Ecker JR, Ren B. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 2013; 153: 1134–48.
77. Yu P, Xiao S, Xin X, Song CX, Huang W, McDee D, Tanaka T, Wang T, He C, Zhong S. Spatiotemporal clustering of the epigenome reveals rules of dynamic gene regulation. Genome Res 2013; 23: 352–64.
78. Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew JL, Ruan Y, Wei CL, Ng HH, Liu ET. Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res 2008; 18: 1752–62.
79. Testori A, Caizzi L, Cutrupi S, Friard O, De Bortoli M, Cora' D, Caselle M. The role of transposable elements in shaping the combinatorial interaction of transcription factors. BMC Genomics 2012; 13: 400.
80. Stefflova K, Thybert D, Wilson MD, Streeter I, Aleksic J, Karagianni P, Brazma A, Adams DJ, Talianidis I, Marioni JC, Flicek P, Odom DT. Cooperativity and rapid evolution of cobound transcription factors in closely related mammals. Cell 2013; 154: 530–40.
81. Lynch VJ, Leclerc RD, May G, Wagner GP. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet 2011; 43: 1154–9.
82. Schmidt D, Schwalie PC, Wilson MD, Ballester B, Goncalves A, Kutter C, Brown GD, Marshall A, Flicek P, Odom DT. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell 2012; 148: 335–48.
83. Merkenschlager M, Odom DT. CTCF and Cohesin: linking gene regulatory elements with their targets. Cell 2013; 152: 1285–97.
84. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006; 20: 2349–54.
85. Lee DH, Singh P, Tsai SY, Oates N, Spalla A, Spalla C, Brown L, Rivas G, Larson G, Rauch TA, Pfeifer GP, Szabó PE. CTCF-dependent chromatin bias constitutes transient epigenetic memory of the mother at the H19-Igf2 imprinting control region in prospermatogonia. PLoS Genet 2010; 6: e1001224.
86. Rand E, Ben-Porath I, Keshet I, Cedar H. CTCF elements direct allele-specific undermethylation at the imprinted H19 locus. Curr Biol 2004; 14: 1007–12.
87. Xie M, Hong C, Zhang B, Lowdon RF, Xing X, Li D, Zhou X, Lee HJ, Maire CL, Ligon KL, Gascard P, Sigaroudinia M, Tlsty TD, Kadlecek T, Weiss A, O'Geen H, Farnham PJ, Madden PA, Mungall AJ, Tam A, Kamoh B, Cho S, Moore R, Hirst M, Marra MA, Costello JF, Wang T. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. Nat Genet 2013; 45: 836–41.
88. Hollister JD, Smith LM, Guo YL, Ott F, Weigel D, Gaut BS. Transposable elements and small RNAs contribute to gene expression divergence between Arabidopsis thaliana and Arabidopsis lyrata. Proc Natl Acad Sci USA 2011; 108: 2322–7.
89. Hollister JD, Gaut BS. Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res 2009; 19: 1419–28.
90. Girard A, Hannon GJ. Conserved themes in small-RNA-mediated transposon control. Trends Cell Biol 2008; 18: 136–48.
91. Xiao S, Xie D, Cao X, Yu P, Xing X, Chen CC, Musselman M, Xie M, West FD, Lewin HA, Wang T, Zhong S. Comparative epigenomic annotation of regulatory DNA. Cell 2012; 149: 1381–92.
92. Strahl BD, Allis CD. The language of covalent histone modifications. Nature 2000; 403: 41–5.
93. Woo YH, Li WH. Evolutionary conservation of histone modifications in mammals. Mol Biol Evol 2012; 29: 1757–67.
94. Spirov AV, Holloway DM. Modeling the evolution of gene regulatory networks for spatial patterning in embryo development. Procedia Comput Sci 2013; 18: 10.1016/j.procs.2013.05.303.
95. Spirov A, Holloway D. Using evolutionary computations to understand the design and evolution of gene and cell regulatory networks. Methods 2013; 62: 39–55.

Reuben Buckley is a PhD student in the laboratory of David Adelson. Reuben completed his First Class Honours degree in 2013 under the supervision of David Adelson. Reuben's Honours research project focussed on the characterisation of genomic regions based on their repeat content and evolutionary conservation, with the aim of identifying ancestral genomic territories.

David Adelson is Professor and Chair of Bioinformatics and Computational Genetics at the University of Adelaide. Dave has been a member of several livestock genome projects, including the cow, horse and sheep. Dave pioneered the de novo identification of repeats in mammalian genome sequences in the cow and horse genome projects. David Adelson's laboratory continues to work on interspersed DNA repeats, particularly with respect to retrotransposon evolution and horizontal transfer in higher organisms.