

Review

Host Gene Regulation by Transposable Elements: The New, the Old and the Ugly

Rocio Enriquez-Gasca [†], Poppy A. Gould [†] and Helen M. Rowe ^{*†}

Centre for Immunobiology, Blizard Institute, Queen Mary University of London, London E1 2AT, UK; r.enriquez-gasca@qmul.ac.uk (R.E.-G.); p.gould@qmul.ac.uk (P.A.G.)

* Correspondence: h.rowe@qmul.ac.uk

† These authors contributed equally to this paper.

Received: 17 August 2020; Accepted: 23 September 2020; Published: 26 September 2020



Abstract: The human genome has been under selective pressure to evolve in response to emerging pathogens and other environmental challenges. Genome evolution includes the acquisition of new genes or new isoforms of genes and changes to gene expression patterns. One source of genome innovation is from transposable elements (TEs), which carry their own promoters, enhancers and open reading frames and can act as ‘controlling elements’ for our own genes. TEs include LINE-1 elements, which can retrotranspose intracellularly and endogenous retroviruses (ERVs) that represent remnants of past retroviral germline infections. Although once pathogens, ERVs also represent an enticing source of incoming genetic material that the host can then repurpose. ERVs and other TEs have coevolved with host genes for millions of years, which has allowed them to become embedded within essential gene expression programmes. Intriguingly, these host genes are often subject to the same epigenetic control mechanisms that evolved to combat the TEs that now regulate them. Here, we illustrate the breadth of host gene regulation through TEs by focusing on examples of young (The New), ancient (The Old), and disease-causing (The Ugly) TE integrants.

Keywords: gene regulation; transposable elements; endogenous retroviruses; epigenetic repression; Intracisternal A-type particle elements; position-effect variegation; KRAB-associated protein 1; X chromosome inactivation; genomic imprinting

1. Introduction

In contrast to their paramount functional importance, protein-coding genes constitute only a small fraction (~2–4%) of the total DNA sequence of the human genome. Exquisitely regulated control of coding genes in time and space is a defining feature of development of multi-cellular organisms. For example, transcription can be regulated by the generation of multiple isoforms of the same gene by alternative splicing, alternative promoter/enhancer usage, non-coding RNAs and epigenetic modifications, which control chromatin structure and function (reviewed in [1]). On the other hand transposable elements (TEs) constitute an estimated two thirds of the human genome [2,3], and contribute to the regulation of protein-coding genes through their regulatory elements. TEs exercise a complex dialog with their host genomes that is distinct from a conventional virus-host arms race because they are not only potential parasites, but also a vital source of genome innovation [4–8]. TEs are subject to epigenetic silencing by histone modifications and DNA methylation [9–11] and become mutated and inactive over the course of evolution. A fraction, however, are co-opted and preserved under purifying selection.

In this review, we illustrate how TEs have been co-opted to regulate host genes by focusing on TEs that can alter their surrounding epigenetic context, with our goal to highlight TEs as a normal feature of host gene regulation. The fact that TEs are so ubiquitous in the genome, contain their own

regulatory sequences and have become hotbeds of epigenetic regulatory marks, due to their initial transcriptional silencing, means that they are ideally placed to re-shape host gene expression profiles. We will journey back in time to explore first how young or ‘new’ TEs, followed by ‘old’ TEs regulate mammalian genes. New TEs are here defined as specific to the primate or murine lineage, whereas old TEs correspond to those which predate the split between mouse and human ancestral lineages (see Figure 1 for examples selected in this review). This distinction allows us to emphasize that, while gene regulatory mechanisms involving TEs are generally conserved across species, the precise TEs that rewire genes are often species-specific. This is due to the different TE invasions that each species has encountered. We finally review instances whereby ‘ugly’ TEs have been retained by the host genome, likely due to them being beneficial, as well as potentially detrimental and therefore discuss the risk that TE co-option poses.

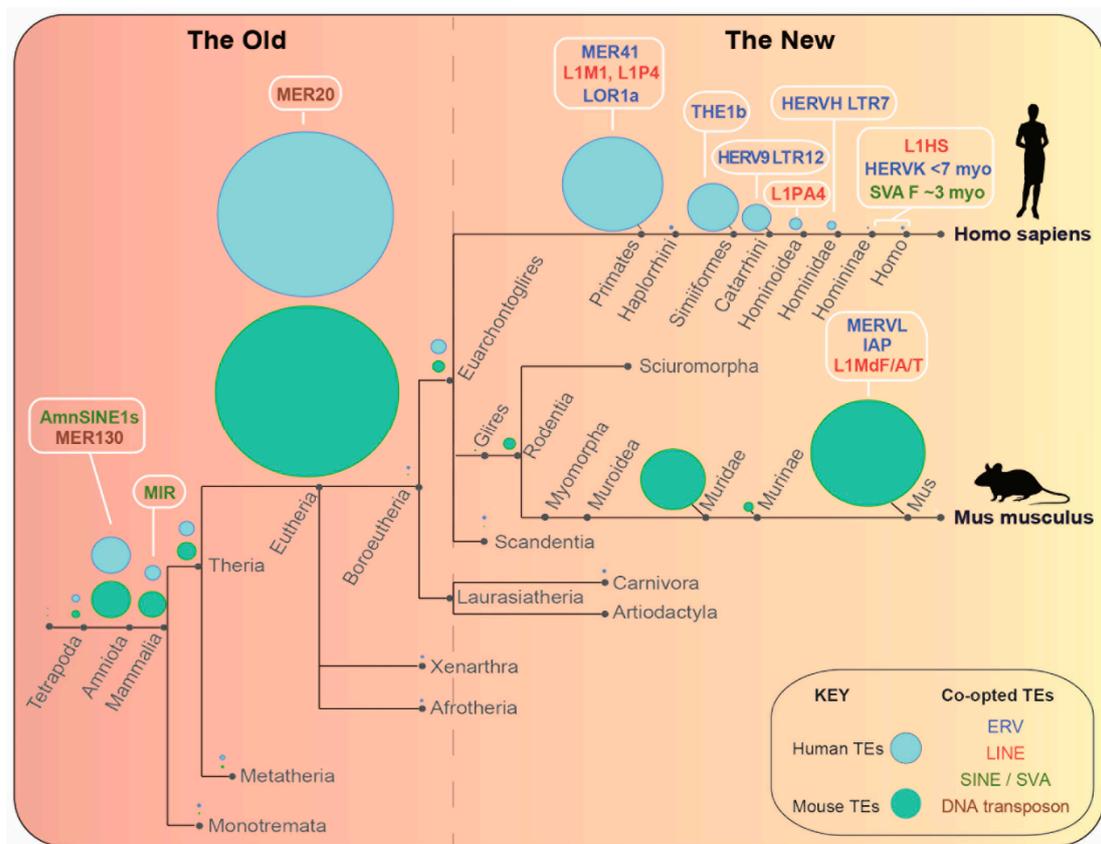


Figure 1. Evolutionary map of example co-opted transposable elements (TEs) for gene regulation. Clade diagram of human and mouse evolutionary trajectories overlaid with bubble plots showing the relative prevalence of all TEs in the human (blue) and mouse (green) genomes, according to their taxonomic specificity (from the *Dfam* database of repetitive DNA families). The dotted line represents an arbitrary evolutionary time cutoff to classify TEs as ‘Old’ vs. ‘New’ in this review. TEs discussed in this review are annotated within their respective taxa and colour coded according to their TE class (see the key). ERV; endogenous retrovirus, LINE; long interspersed elements, SINE; short interspersed elements, SVA; SINE/VNTR/Alu elements.

2. Gene Regulation by Transposable Elements: The New

Since the human and mouse lineages diverged from a common ancestor around 80 million years ago, their genomes have been subject to different selective pressures, innovations and invasions. The present-day human genome has been found to contain no endogenous retroviruses (ERVs) capable of replication/transposition [12], but to host around 100 retrotransposition competent Long Interspersed Element 1s (LINE-1s or L1s) [13]. The human genome is also home to SVA elements, a newly evolved

composite TE derived from SINEs and an ERV (HERV-K10). SVAs harbour a variable number of tandem repeats (VNTRs) and hijack L1 retrotransposition machinery for their mobilisation. With the youngest SVA family (SVA_F), around three million years old (myo), SVA elements represent the youngest TE in the human genome [14]. In contrast, the mouse genome appears to contain cohorts of ERVs and L1s capable of retrotransposition [15,16]. Here, we inspect examples of epigenetic control of host genes through regulatory sequences embedded in young species-specific TEs. We draw on mouse and human examples and discuss how this can shape our understanding of how TEs underpin human adaptation and genetic variation. We include scenarios whereby parallel TEs have been independently co-opted for the same purpose in both organisms. Future studies on actively transposing TEs may allow us to observe how genome invaders become co-opted into gene-regulatory networks in real time.

2.1. New Transposable Elements in Mouse

Mouse-specific endogenous retroviruses include the Intracisternal A-type particles (IAPs), which adapted to retrotranspose intracellularly following loss of their envelope gene [17] and murine endogenous retrovirus L (MERVL). A small fraction of IAP elements can still retrotranspose and this subfamily has been a source of polymorphisms that have been actively studied for their effects on the expression of nearby genes [18–20]. In this section, we will focus on examples of gene regulation through specific MERVL and IAP-derived regulatory elements, which provides insight into how ERVs may directly influence gene expression and host fitness. Mechanisms by which ERVs that are discussed in this review regulate host genes are summarized in Figure 2.

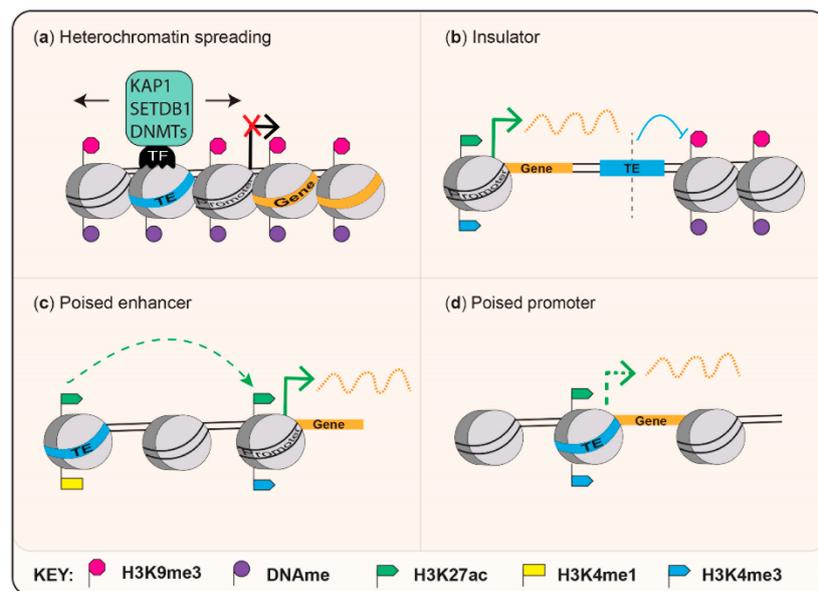


Figure 2. Mechanisms by which transposable elements (TEs) regulate host genes that are discussed in this review. Heterochromatin spreading into genes is represented in (a), which can occur stochastically but often radiates from silencers bound by sequence-specific transcription factor (TF) repressors. These can recruit heterochromatin-related proteins, some of which are highlighted here. Heterochromatin spreading has been described for IAP ERVs [20,21]. An insulator function is portrayed in (b), whereby a TE can protect a host gene from heterochromatin spreading and an example is MIR elements [22]. ‘Poised’ or cryptic TE-derived enhancers and promoters are shown in (c, d), which can be poised due to their dynamic epigenetic repression as discussed in this review. TE promoters can function as alternative promoters for protein-coding genes as is the case for MERVL long terminal repeats (LTRs), which can generate chimeric transcripts with host genes expressed at the 2-cell stage of mouse development [23]. MER41 has been shown to act as a poised enhancer [24].

2.1.1. Co-Option of TEs to Regulate Gene Networks

Perhaps one of the best examples of a co-opted TE regulating a network of genes is MERVL LTR (MT2) promoters driving expression of genes specific to the totipotent 2-cell (2C) stage of development [25,26]. It is possible that this TE-gene network evolved due to MERVL invasions into genes actively expressed during this stage of development, which represents a window of opportunity for escape from ERV repression due to epigenetic reprogramming. Alternatively, insertion prior to differentiation of the germline would also represent a selective advantage to the TE [27]. Originally identified as TEs, which generate chimeric transcripts with host genes in cleavage-stage mouse embryos [28], MERVL expression was later shown to be associated with enhanced developmental potency in *in vitro* and *in vivo* assays [23]. Further work has shown that the MERVL-2C gene network is activated following depletion of the chromatin assembly factor-1, CAF-1, in mouse embryonic stem cells (mESCs), through increased chromatin mobility [29] and is associated with genome-wide DNA demethylation, potentially through upregulation of the translation inhibitor *Eif1a-like* [30].

Elegant work has identified that the transcription factors DUX, ZSCAN4, DPPA2 and DPPA4 activate 2C specific genes by binding to MT2 LTR promoters [30–37]. Mechanistically, MERVL LTRs, therefore, regulate genes by acting as poised promoters (Figure 2). Intriguingly, overexpression of the zinc finger protein, ZSCAN4 has also been described to protect cleavage embryos from DNA damage [38]. In Figure 3, we illustrate the MERVL-2C gene network by depicting DUX-targeted MT2 LTRs within 10 kb of 2C-associated genes [23,30]. The expression profile of *Zscan4c* mRNA as defined in [39] is also displayed (Figure 3, and Table S1 (Supplementary Materials) for raw data). Of note, L1 elements also exert co-opted roles in totipotency and early developmental transitions [6,40,41]. Overexpression of the human DUX orthologue, DUX4 in human ESCs leads to an induction of ERVL promoters [35], which are usually expressed at the cleavage stage of human development [42]. Therefore, there are obvious parallels between regulation of mouse and human totipotency [26] and MERVL and ERVL are derived from the same retrovirus superfamily. ERV regulation of the 2C stage gene network serves as a striking example of convergent evolution [26] or convergent co-option, and further examples of convergent co-option are discussed below.

2.1.2. Sequence-Specific Epigenetic Silencing

IAP elements (restricted to the ‘mus’ lineage, see Figure 1) are subject to sequence-specific epigenetic silencing through KRAB-zinc finger proteins (KZFPs), of which there are around 700 in the mouse genome [43]. KZFPs recruit KAP1 and SETDB1, which create heterochromatin foci that can spread into and repress neighbouring genes [21,44–47]. This concept is illustrated by the KZFP, ZFP932, which binds to a subfamily of IAP elements through a sequence within the proviral 3′ polypurine tract, which is a determinant of retroviral replication [48]. This regulatory sequence now serves to regulate expression of an IAP-proximal gene, *Bgalp3*. Inactivation of *Zfp932* results in loss of local silent chromatin marks and a gain of the enhancer marks H3K27ac and H3K4me1 and Pol II accumulation. Similarly, depletion of KAP1 or SETDB1 in mESCs or neural progenitor cells (NPCs) leads to multiple instances of increased expression of ERVs and their proximal genes [21,44,49]. This is accompanied by an epigenetic switch from a dual H3K9me3 and H4K20me3 repressed configuration to an enhancer signature, characterised by H3K27ac and H3K4me1 [21]. A causative role for IAP-embedded enhancers in regulating proximal genes upon KAP1-depletion was recently demonstrated, by employing strain-specific IAP-integrants [50].

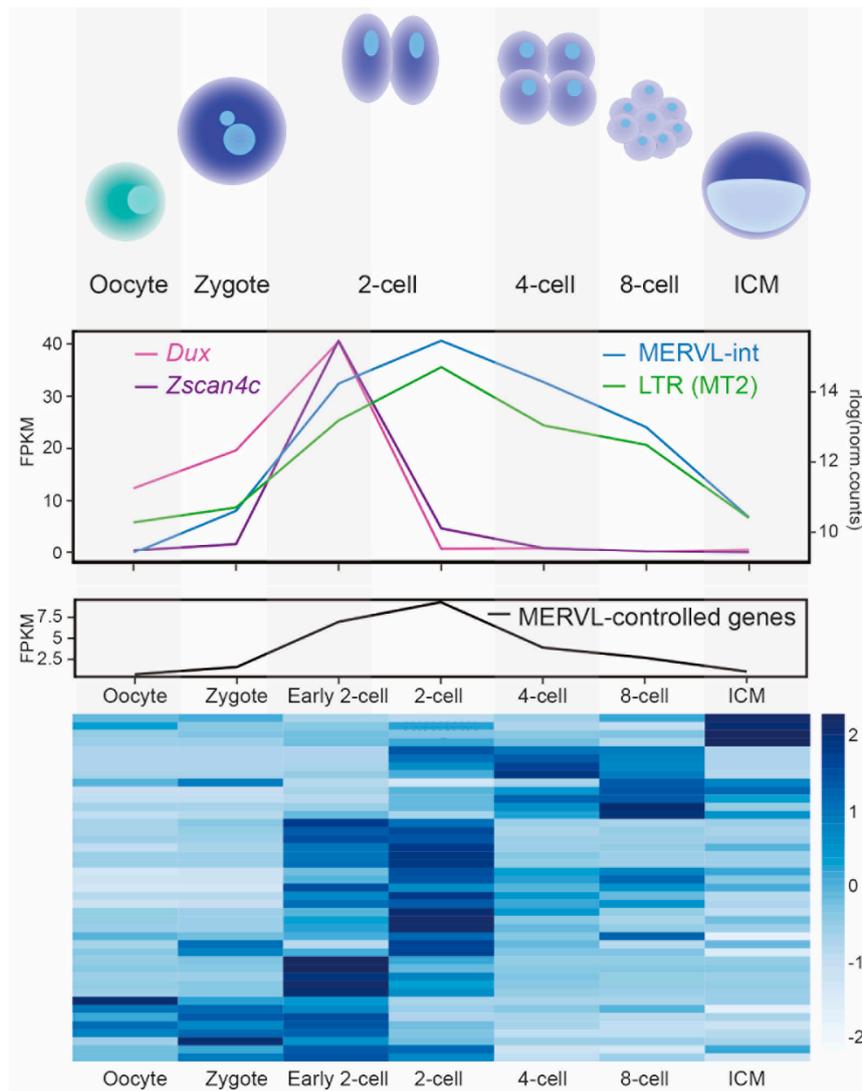


Figure 3. Temporal gene regulation through MERVL-derived LTR promoters. Top: Diagram of early mouse development. Middle: Median expression values (FPKM) of the transcription factors, *Dux* and *Zscan4c* (left axis) and of MERVL and its LTR promoter (MT2) as defined by ‘TEcounts’ software (right axis) are shown through development using data from [39]. Bottom: Expression of 43 2C stage-specific genes, controlled through MERVL LTRs, depicted both as the median expression in FPKM (‘MERVL-controlled genes’ line graph) and as expression values of individual genes in a heatmap (values are scaled by row). A list of 2C stage-expressed genes was compiled by including previously published lists from both the following studies: [23,30]. FPKM expression values through development were extracted from [39]. Only genes that are within 10 kb downstream of MT2 LTRs that overlap a DUX binding peak (peaks extracted from [35]) and have detectable expression levels in [39] were considered here. See Table S1 (Supplementary Materials) for raw data.

Figure 4 provides an illustration of how IAP elements can exhibit genome-wide effects on gene expression through their epigenetic repression, using published data [21]: IAP elements are depicted that exhibit KAP1-dependent H3K9me3 peaks proximal to KAP1-regulated genes. It is not known if IAP elements exert a natural role in the tissue-specific regulation of these genes [21]. Mechanistically, IAP elements regulate genes through spreading heterochromatin from silencers and act as poised enhancers (see Figure 2). Recently, the KZFP, *Zfp708* has been discovered to exert transgenerational maintenance of DNA methylation at LTR retrotransposons [51]. Many KZFPs still have unknown roles,

although the functions and binding profiles of clusters of young KZFPs have been assessed in a new study using knockout mice, coupled with chromatin-immunoprecipitation assays in mESCs [43].

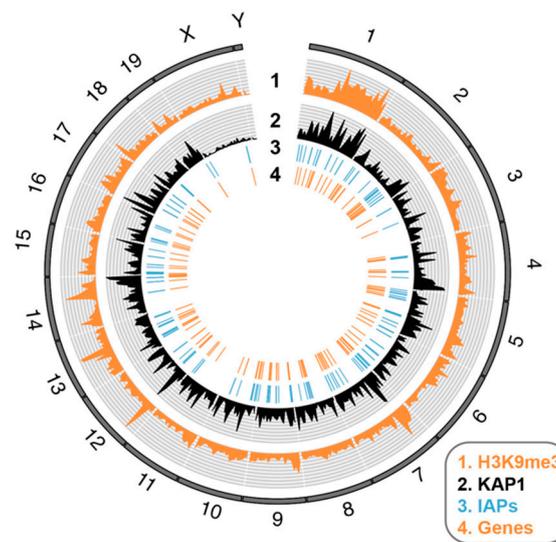


Figure 4. Heterochromatin spreading from silencers embedded in IAP-type ERV elements. Circos plot showing circular visualization of the mouse genome depicting IAP elements coated with silent chromatin marks and proximal genes that they potentially regulate. Data are from C57BL/6J-derived mESCs and extracted from [21]. Track 1 corresponds to a histogram of KAP1-dependent H3K9me3 coverage across 5-Mb windows. This represents regions of H3K9me3-enrichment that are present in KAP1 control mESCs but lost in KAP1 KO mESCs. Track 2 depicts a histogram of KAP1 peak coverage in 5-Mb windows. Track 3 shows individual occurrences of IAP elements (290) that overlap H3K9me3 peaks shown in track 1, and which are less than 10 kb from genes upregulated (>2 fold) in KAP1 KO mESCs. Note that KAP1-dependent H3K9me3 peaks are used rather than KAP1 peaks, due to their increased mappability at highly repetitive (young) IAPs, resulting from their spreading into genes. Track 4 depicts the coordinates of 145 unique genes <10 kb from IAPs in track 3 and upregulated in KAP1 KO mESCs. See Table S2 (Supplementary Materials) for raw data.

2.1.3. Metastable Epialleles

The best characterised examples of ERV integrants differentially regulating a gene between individual mice may be the *Agouti viable yellow* (*Avy*) [52,53] and *Axin-fused* (*Axin^{Fu}*) [54] alleles, which arose due to insertions of IAP elements either upstream or within an intron of the *Agouti* or *Axin* genes, respectively. These IAPs are variably silenced by DNA methylation between individuals that are genetically identical at these loci, resulting in variable expression of *Agouti* and a range of coat colours, or expression of a truncated version of *Axin*, which causes a kinked tail phenotype. These and similar events, the majority of which are evolutionarily young, have been termed metastable epialleles and have been recently catalogued in a genome-wide screen, although only a few have been shown to alter gene expression thus far [18]. How these metastable epialleles arose and to what degree they function as regulators of gene expression remains unclear [55,56]. Importantly, many IAP copies are conserved across mouse strains and subject to KAP1/KZFP-mediated stable epigenetic repression (see above, Section 2.1.2) and the latter copies may play a more prominent role in repressing host genes than the polymorphic metastable epialleles, which may have arisen through mutation of cis-acting silencers, as has been documented to occur for L1 [57]. Future work on variably methylated IAP elements may further our understanding of genetic variation between individuals as well as providing insight into epigenetic silencing mechanisms of actively transposing ERVs.

2.1.4. Position-Effect Variegation

Young IAP insertions can contribute to position-effect variegation (PEV). This is a phenomenon whereby genes or transgenes exhibit variegated expression in some cells but not others, due to their position nearby heterochromatin, which can spread. For example, it was shown that a strain-specific IAP insertion approximately 300 bp upstream of the *B3galtl* gene (beta 1,3-galactosyltransferase-like) could repress gene transcription through spreading of H3K9me3, H4K20me3 and DNA methylation, in mESCs [20]. Of note, the human silencing hub (HUSH) complex was identified as a novel epigenetic complex involved in PEV [58] and will be discussed below. Further examples of IAP elements regulating genes are being discovered regularly [59], suggesting that data thus far represent only the tip of the iceberg.

2.2. New TEs in Humans and Convergent Co-Option

The capacity for evolutionarily young intact TEs to repress proximal genes in development has also been documented in human models. For example, loss of the maintenance DNA methyltransferase, DNMT1 in human NPCs results in demethylation and transcription from young, hominoid-specific L1 antisense promoters, which can give rise to chimeric transcripts with proximal genes (<fifty kb away) [60]. This work builds on the previous discovery that the L1 5'UTR has an antisense promoter and kozak sequence and produces ORF0, which can form fusion proteins with proximal exons [61]. There are some similarities between DNMT1-depletion and KAP1-depletion but in the case of KAP1, upregulated genes were mainly proximal to HERVKs (HML2) and SVAs (of < seven myo and ~three myo, respectively) [62–64]. Of note, KAP1 can target the primer binding site (PBS) of HERVKs, in a similar way to its targeting the PBS of MLV [65–67], and this mechanism is known to silence an adjacent reporter promoter [63,68,69]. It is not known if the above TEs can naturally activate adjacent genes, since the above studies involve knockout of epigenetic modifiers.

An important example of a TE co-opted to activate genes is the hominoid-specific HERVH LTR7 [70] (Figure 1), which functions as an enhancer in pluripotent cells and is hypomethylated and expressed in differentiation-defective hiPSCs (human-induced pluripotent stem cells), reviewed in [7,71]. Similarly, upregulation of HERVK in pluripotent cells has been reported [64,72]. Notable studies have documented how TEs may regulate human genes in adult tissues, for example in CD4 + T cells and macrophages [68,73–76]. Still relatively little is known, however, about how the ever-evolving TE burden contributes to present-day human gene regulatory networks. By comparing gene expression and histone marks associated with functional and poised enhancers (H3K27ac and H3K4me1) across primate lineages [77,78], it has recently been shown that many regulatory regions are derived from new TE insertions (including SVA_B,C,D and F; LTR12 and Alu). For example, a human-specific SVA_F insertion located in the intron of the gene *Jarid2*, was identified to function as a silencer in the liver and nervous system [79].

Below, we will discuss several examples of convergent co-option. This will highlight how TEs exert parallel roles in the human and murine lineage, despite being species-specific, and emphasizes the need for more comparative genomics in future work.

2.2.1. The HUSH Complex

The human silencing hub, or HUSH complex was identified in a screen for mediators of PEV in human cells [58,80] and is comprised of TASOR (also known as FAM208A), MPP8 (encoded by MPHOSPH8) and periphilin-1 (PPHLN1). HUSH is recruited to H3K9me3-dense genomic loci and partners with the chromatin remodeler, MORC2 [80–82] or MORC2A in mice [11]. The HUSH complex also partakes in the restriction of incoming exogenous retroviruses to which it is recruited through a novel DNA binding protein, NP220 (*ZNF638*) that is attracted to clusters of cytidines [83]. Although identified as a complex together with MPP8 and periphilin-1 in human cells, FAM208A was earlier identified as a novel epigenetic modifier in an ENU mutagenesis screen in mice [84]. FAM208A plays

a critical role in development because homozygous mutant mice are not viable beyond gastrulation. We and others have shown that the HUSH complex is required to regulate expression of young (<five myo) transcriptionally active L1 elements (L1Md_F/A/T, see Figure 1) in mESCs [11,82,85]. It also represses genes that have accumulated these TEs upstream or within their introns, a trait shared by KAP1 [85]. Some of these genes are mouse-specific, suggesting they are recently evolved. The HUSH complex exerts a parallel role in human cells in regulating full-length L1 elements (L1PA4 and L1HS, see Figure 1) in the hominoidea lineage [82,86]. Similarly to mESCs, HUSH-regulated L1s are often located within introns of active genes, where they attract local H3K9me3, resulting in a slight downregulation of the genes in which they are positioned [82]. Genes repressed by the HUSH complex include KZFPs [58], which regulate TEs themselves. New data suggest that the HUSH complex targets RNA, revealing how it could be recruited to and exert epigenetic repression on transcriptionally active L1 elements [87].

2.2.2. A TE Origin to Genomic Imprinting

Genomic imprinting refers to the differential DNA methylation at imprinting control regions (ICRs), established in the germline, which determines parental allele-specific expression of a set of imprinted genes [88]. This epigenetic mechanism occurs in eutherian mammals as well as, less frequently, in marsupials and is essential for the regulation of development. Historically postulated to be a phenomenon exemplifying “the battle of the sexes” [89], the evolutionary origin of genomic imprinting remains enigmatic. However, one prominent theory is that it arose from a DNA methylation-based defence mechanism against exogenous DNA [90]. There are various lines of evidence to support this theory reviewed in [91], including the fact that some, but not all, imprinted genes resemble TEs while others originate from retrotransposition events [92–94].

Two recent studies have shed light on the link between TEs and imprinting by demonstrating that species-specific ERVs epigenetically regulate mouse- and human-specific imprinted genes [95], as well as non-canonically imprinted genes in mouse extra-embryonic tissues [96]. Another link between imprinted genes and TEs relates to the role of KZFPs in the maintenance of genomic imprints: ZFP57 and the more recently identified ZNF445/ZFP445 [97] are both critical to this process in mice and humans and, interestingly, have both been shown to also bind to TEs [98,99], suggesting that the binding motif of these proteins may have derived from a TE. Likewise, a recent study into the stochastic loss of imprinting (LOI) between mouse ESC strains mapped this instability to a region of chromosome 13 that overlaps a cluster of KZFPs [100], including some which have been suggested to regulate sex-specific gene expression [101], further implicating KZFPs in the regulation of imprinted genes. Thus, while genomic imprinting is conserved in eutherian mammals, the contribution of TEs to the regulation of genomic imprinting is evolving in a species-specific manner.

2.2.3. Fighting Fire with Fire: TEs as Effectors of Immunity

The innate immune system, while conserved among mammals, displays marked species-specificity in the transcriptional response to interferon signalling, consistent with its role in adaptation against pathogens [102–104]. MER41 is a primate-specific ERV, which has been shown to act as a poised enhancer (see Figures 1 and 2) for a number of interferon- γ (IFN γ)-stimulated genes through its recruitment of the transcription factor, STAT1 [24]. While the regulation of innate immunity genes by MER41 elements may be largely species-specific, MER41-like elements with the ability to act as IFN-dependent enhancers are common to many mammalian lineages [105]. Interestingly, in the mouse, which lacks any MER41-like elements, a murid-specific ERV RLTR30B is enriched for STAT1 binding and exhibits IFN-inducible enhancer activity in reporter assays [24]. The mammalian innate immune system, therefore, appears to have been recurrently but independently shaped in individual lineages by ERV-derived IFN-inducible enhancers.

3. Gene Regulation by Transposable Elements: The Old

The genomes of present-day species are rife with remnants of their previous bombardment with TE insertions throughout evolution. Such ancient elements have had ample time for mutations to render them incapable of mobilizing in the genome, while simultaneously evolving potentially beneficial roles. It has therefore been hypothesized that TEs that are conserved across species are more likely to have been co-opted [106]. Of note, the contribution of ancient TEs to gene regulatory networks is underestimated due to the erasure of ancient TEs through genetic drift and their loss over evolutionary time except for the transcription factor binding sites, which can be preserved under purifying selection.

3.1. When X-Chromosome Inactivation Is on the LINE

Perhaps the best illustration of the importance of gene regulation by TEs at a higher order chromatin level is X chromosome inactivation (XCI). In contrast to autosomal chromosomes, where cells inherit two copies of each chromosome, the sex chromosomes pose a problem of unequal dosage, whereby females receive two copies of the X chromosome, while males have only one. To resolve this problem an entire X chromosome is subject to transcriptional silencing in females during development. The mechanism behind this involves coating of the inactive X by the long non-coding RNA *Xist*, which is expressed from the inactive X chromosome [107], reviewed in [108].

Mary Lyon originally hypothesized that the high density of L1 elements across the X chromosome may facilitate XCI through heterochromatin spreading of H3K9me3 marks [109] and see Figure 2 for a diagram of heterochromatin spreading. Indeed, L1s occupy roughly twice as much of the X chromosome than they do of autosomal chromosomes [110] with an enrichment of L1M1 and L1P4 subclasses (see Figure 1). These elements were active during the transition between eutherians to prosimians, 60 to 100 million years ago. Interestingly, 10% of genes on the human X chromosome that escape inactivation are located in segments with significantly fewer L1s [110]. It was later discovered that the density of L1s is an important factor in enabling efficient XCI [111]. It remains to be determined if phase separation, which has been recently defined as a feature of heterochromatin spreading [112,113] applies to XCI. Intriguingly, there appears to be preferential invasion of the X chromosome by LINES of any age [110], suggesting that these elements are continuously co-opted to play a role in XCI.

3.2. Ancient TEs Shaping the Brain of Mammals

Two remarkable examples that substantiate the Britten and Davidson hypothesis [114], wherein the emergence of novel structures or functions could be aided by the co-option of ancient TEs, involve structures specific to mammalian brains: the corpus callosum and the neocortex.

The corpus callosum: Tashiro and colleagues identified a Short INterspersed Element (SINE) locus to exert an enhancer function specifically in the corpus callosum [115]. This SINE locus (AS021) belonged to an ancient family of SINE elements conserved among amniotes with some copies over three hundred myo [116]. Using a *lacZ* reporter assay, SINE AS021 was shown to drive reporter expression specifically in mouse cortical neurons that project axons into the corpus callosum [117] and served as a natural enhancer for *Satb2*, a transcription factor (TF) involved in corpus callosum formation. Of note, the authors could identify several other conserved 'Amniota SINE1s' (AmnSINE1s, see Figure 1), with evidence of their co-option in regulation of the corpus callosum, which is interestingly only present in placental mammals [118]. The corpus callosum connects the two hemispheres of the brain, facilitating their communication.

The neocortex: another brain-specific enhancer derived from a TE is MER130 (MEdium Reiteration frequency), which is also conserved amongst amniotes (Figure 1). This enhancer was identified through mapping the genome-wide binding sites of the co-activator p300 in the developing mouse embryo and shown to be enriched in the neocortex of E14.5 mouse brains. Importantly, a number of MER130 elements were found to be marked with H3K27ac and to contain binding motifs for several TFs important for brain development. These MER130 loci were verified to function as enhancers

using a *luciferase* assay and were located next to genes annotated as being associated with abnormal telencephalon morphology. Figure 5a shows the distribution of MER130 elements in the mouse genome and highlights those which are bound by p300 in the E14.5 neocortex as well as loci where this element overlaps a H3K27ac mark in the E14.5 whole brain. Here, we also reveal that (1) 54 of the 107 mouse MER130 elements described in this study are conserved in the human genome and (2) the 12 genes associated with abnormal telencephalon development in the mouse genome are also located in proximity to MER130s in the human genome. One example of this, *Zfp432/ZNF432*, is depicted in the figure. Furthermore, this subset of MER130s appear to overlap with DNase hypersensitivity sites in day 85 human brain (the equivalent timepoint to E14.5 in the mouse). These results illustrate that TEs have been co-opted as enhancers in the neocortex, another mammalian specific structure [119], and suggest that this function is conserved between mammalian lineages. The neocortex is involved in conscious thought and reasoning, and in humans it is involved in language. Notably, the neocortex has evolved a significantly different structure in primates compared to rodents with abundant folds to increase its overall surface area. Little is known about the origin of MER130 from which this enhancer is derived, except that it was likely once a DNA transposon.

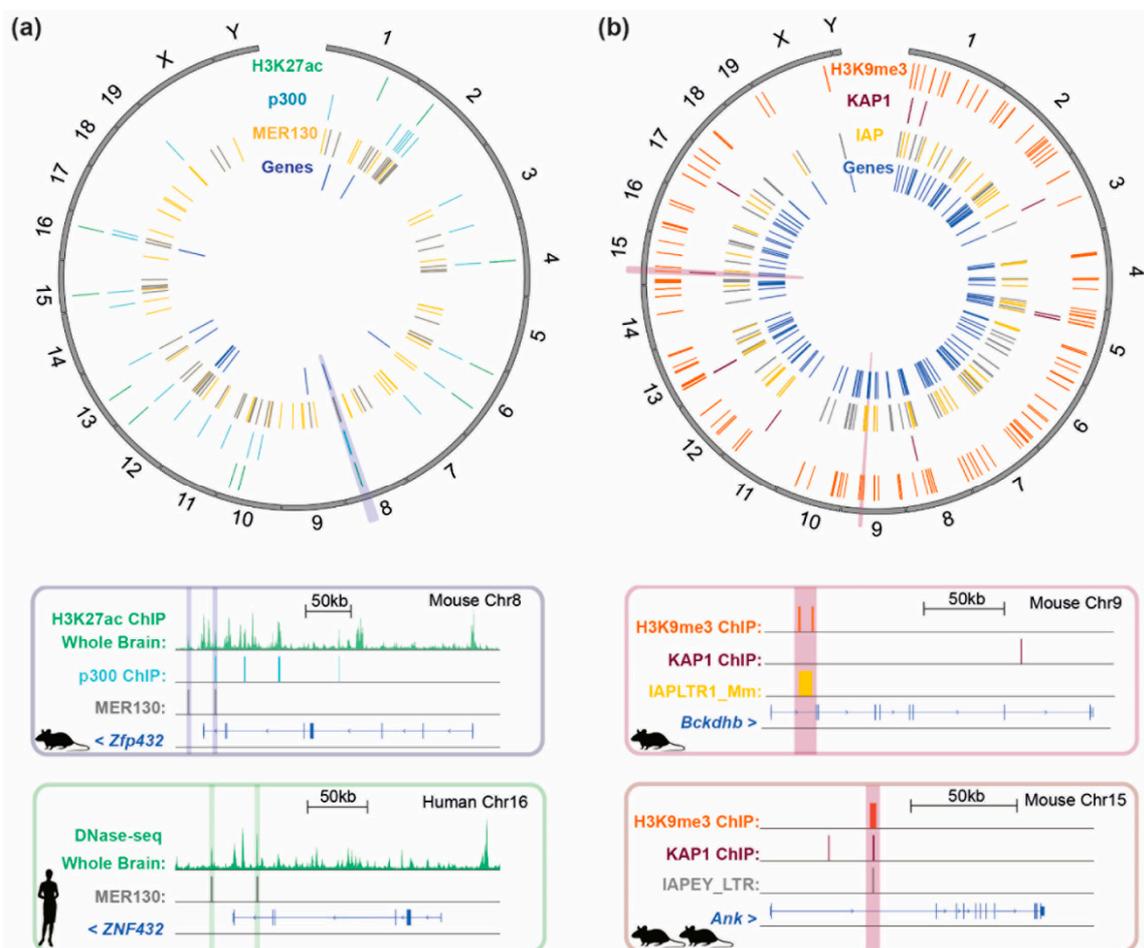


Figure 5. Chromatin landscape of past co-option events vs. co-option ‘in action’ of new TEs (a) Circos plot of MER130 TEs in the mouse genome according to [119]. (1) H3K27ac peaks (ENCODE) in E14.5 whole brain which overlap MER130 elements. (2) p300 binding sites in E14.5 dorsal cerebral wall, which overlap MER130 elements. (3) MER130 elements in the mouse genome depicted in yellow, except for those MER130 instances that are conserved in 60 placental mammals, including in the human genome, which are depicted in grey. (4) Genes located near to MER130 cortical enhancers which, when perturbed, result in abnormal telencephalon morphology [119]. (b) Circos plot of IAP elements in the mouse genome under KAP1 regulation according to [21]. (1) Coordinates of KAP1-dependent

H3K9me3 peaks. (2) KAP1 peaks that overlap IAPs proximal (<10 kb) to genes upregulated upon KAP1 KO. (3) IAP elements which intersect H3K9me3 peaks in (1) and overlap or are proximal (<10 kb) to genes upregulated upon KAP1 KO. IAPs conserved between 14 inbred laboratory mouse strains are depicted in grey, whereas IAPs for which there is evidence of their absence in more than one strain (exhibiting strain variability) are highlighted in yellow. Genome browser windows in the lower half of the figure depict the highlighted TE/gene intersections from the circos plot in (a) showing conservation in the mouse and human genomes or (b) an IAP integrant that is only present in C57BL/6J-strain mice vs. an IAP integrant conserved across all 14 inbred mouse genomes. See Table S3 (Supplementary Materials) for raw data.

3.3. Ancient Mammalian-Conserved TEs as Insulators

Co-opted repeats have also been documented from the ancient MIRs (mammalian-wide interspersed repeats), which are amongst the most ancient TE families in the human genome and are classified as tRNA-derived SINEs (see Figure 1). This repeat family was originally found to be relatively enriched in the proximity of transcriptional start sites and to correlate with tissue-specific gene expression [120]. Later analyses employing CD4+ T cells identified a small subset of MIRs (0.36%) that function as insulators, characterised by a presence of a B-box and their capacity to recruit Pol III. Their chromatin barrier activity appears to be cell-type specific for CD4+ T cells, where the authors could also identify differences in the expression levels of genes found on either side of the MIR-associated insulators [22]. Remarkably, a MIR integrant was proposed to have enabled differentiation of regulatory CD4+ T cells in placental mammals to reduce inflammatory immune responses that would otherwise target the foetus during pregnancy [121]. A similar example is the eutherian DNA transposase MER20 (see Figure 1), which possesses an insulator function suggested to have contributed to the process of differentiation of endometrial stromal cells by limiting the spread of heterochromatin [122]. See Figure 2 for a diagram of a TE function as an insulator.

4. Gene Regulation by Transposable Elements: The Ugly

Given the ability of TEs to regulate genes, including generating chimeric transcripts, it is not surprising that TE insertions can cause disease. The phenomenon whereby cancer progression imitates exaptation events occurring during evolution by employing TEs has been termed onco-exaptation [4]. In this section, we highlight examples whereby the epigenetic dysregulation of TEs has been shown to underlie human diseases. Examples of actively transposing TEs causing disease, although important will not be discussed here, since they are beyond the scope of this review.

4.1. Gene Dysregulation by TEs in Cancer

Like the aberrant activation of TEs following experimental disruption of epigenetic factors discussed above, in B-cell derived Hodgkin's lymphoma, the de-repression of the simian MaLR-type LTR, THE1B (Figure 1) results in the expression of a non-canonical transcript of the proto-oncogene CSF1R. The derepressed LTR is silenced by DNA methylation in non-Hodgkin cell lines, whereas cells from these patients exhibit a loss of ETO2 (CBFA2T3), a transcriptional repressor interacting with histone deacetylases (HDACs). Additional transcripts originating from the THE subfamily of LTRs were also observed in Hodgkin's lymphoma cells [123], illustrating that multiple derepressed LTRs can have further-reaching consequences in lymphoma development. In fact, a later study identified an additional derepressed LTR to result in the activation of interferon regulatory factor 5 (IRF5), a transcription factor previously shown to be necessary for the survival of Hodgkin Lymphoma cells [124]. Expression of IRF5 was driven by the primate-specific LTR, LOR1a, which was hypomethylated in samples showing aberrant expression of IRF5 [125]. Notably, it has also been reported that infection by Epstein-Barr Virus can lead to the activation of LTRs, such as ERV1 and ERVL, in the context of primary B-cells and lymphoblastoid cell lines [126]. New examples of LTRs functioning as enhancers to disease-relevant genes have recently been discovered for acute myeloid leukaemia [127]. Cancers are not only associated

with upregulation of ERV regulatory elements but also the production of exapted TE-derived proteins that regulate the immune response [128].

4.2. Gene Dysregulation by TEs in Autoimmune Disease

As well as cancers, multiple instances of autoimmune disease, such as rheumatoid arthritis or multiple sclerosis involve some degree of aberrant TE activation [129]. In some such instances, natural or chimeric proteins produced by TEs when they are derepressed are thought to contribute to disease [130]. One example is in the case of Systemic Lupus Erythematosus (SLE), in which several TEs have been found to be hypomethylated in neutrophils from diseased patients, compared to healthy controls, particularly at L1 elements [131]. Importantly, genes upregulated in the neutrophils of SLE patients were found to be associated with the presence of L1s in the antisense orientation with respect to the gene. Upregulated genes with L1s were enriched for biological processes involving apoptosis and programmed cell death [132]. Remarkably, KAP1 and KZFPs have been linked to the pathogenesis of lupus, in which disease was associated with the expression of a retroviral envelope protein produced from an ERV that was under epigenetic regulation in healthy individuals through binding of several KZFPs to the ERV LTR [133]. Of note, TEs are emerging more broadly as key players in driving inflammation and autoimmunity, including neuroinflammation in the context of various brain disorders, as reviewed in [134,135].

The above examples represent a mere snapshot of the potential consequences of loss of epigenetic control of TEs. Despite little causative data in this area thus far, given the direct impact of the described TEs on genes with critical roles in these pathologies, it is possible that these events are causative rather than passenger events.

5. Conclusions

In this review, we have sought to illustrate how epigenetic repression of invading TEs has led to the evolution of epigenetic regulation of gene networks, into which invading TEs have become embedded. We have focused on evolutionary young (new) TEs, a fraction of which are still active, as well as old conserved TEs, drawing on mouse and human models, in order to capture some of the breadth of TE co-option into gene regulatory pathways. Figure 5 provides a summary of ancient (MER130) TE-driven conservation throughout placental mammals (Figure 5a), compared to young (IAP) TE-driven gene regulation that is not necessarily fixed across mouse strains (Figure 5b) and represents a unique snapshot of ongoing TE co-option in real time. We speculate that new TEs continually fulfil and are co-opted into the same roles as old TEs, as long as they outperform their predecessors in terms of acting as effective gene regulatory elements. We note that while work to date has identified many instances of individual TEs contributing to the regulation of host genes, very little is known about TE co-option more broadly, particularly into the potential regulation of whole gene expression programmes and biological systems. Since TEs can drive genome innovation and adaptation in response to pathogen challenge, we hypothesize that TEs will be uncovered to play a more prominent role in the evolution of the human immune system. We also highlight the benefit of cross-species comparative studies in future work on TE co-option.

Notable examples we have focused on of TEs regulating genes in this review are MERVL LTRs driving expression of early embryonic genes, IAP ERV silencers repressing proximal genes in mESCs, and TEs repeatedly co-opted to mediate genomic imprinting. Note that although we have limited our scope to mouse and human studies, exciting examples of gene regulation by TEs are widespread and extend to the control of fruit colour in the tomato plant, for example [136]. Co-option of TEs is not limited to discrete genes or networks and here we also discuss how TEs have contributed to the evolution of X chromosome inactivation, innate immunity and even to new brain structures such as the corpus callosum, which is involved in communication between left and right hemispheres of the brain. Finally, we discuss how the advantage of the host harnessing TEs as elegant tools for genome

innovation also comes at a cost: TE integrants can also cause disease as exemplified by increased obesity and diabetes observed in yellow *Avy* mice, or cancers and autoimmune disease in humans.

6. Methods

6.1. Expression Analysis in Mouse Pre-Implantation Development

The annotation of MERVL LTRs (MT2_Mm) was extracted from the RepeatMasker track for mm10. Dux binding was obtained from [35], the intersection of peaks called in the two individual replicates was used to identify Dux-bound MT2s using [137]. A list of 2-cell expressed genes was compiled from [23,30] which was refined to only include genes that are within 10 kb downstream of Dux-bound LTRs. Fragments Per Kilobase of transcript per Million mapped reads (FPKMs) were calculated using data from [39] using [138–140]. Read counts for repeat families were calculated using [141] and normalised for library size. Dux expression was calculated by aligning reads to reference AM398147 with [139]. Mean values across replicates are reported.

6.2. Genome-Wide Visualization of Features

Data for KAP1 binding, location of H3K9me3 and gene expression values following KAP1-KO were taken from [21]. The annotation of IAP elements was extracted from the RepeatMasker track for mm9. Variability of repeat integrants was derived from [142]. IAP elements were considered variable when they overlapped a deletion in more than one of the inbred laboratory strains. Overlapping and closest features were calculated with [137] and visualized with [143]. MER130 annotation was obtained from [119] and converted to version mm10 of the mouse genome. p300 coordinates from E14.5 dorsal cerebral wall taken from [144], and converted to mm10. Broadpeaks for H3K27ac in E14.5 mouse whole brains were downloaded from the UCSC ENCODE [145] tracks. PhastCons [146] conservation scores calculated for a multiple alignment across 60 placental mammals species were retrieved from UCSC Table Browser [147] specifically for MER130 coordinates. Instances of MER130 were termed conserved when the conservation score was > 0.7 for more than half of the repeat's length.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4915/12/10/1089/s1>, Table S1: Genes regulated through MERVL LTRs, Table S2: Genes regulated through KAP1-repressed IAP elements, Table S3: Genes regulated through ancient vs. species-specific ERVs.

Author Contributions: Conceptualization, methodology, data presentation and writing: R.E.-G., P.A.G. and H.M.R. All authors have read and agreed to the published version of the manuscript.

Funding: R.E.G., P.A.-G. and H.M.R. are grateful to receive funding through the European Research Council (TransposonsReprogram: 678350) and a Barts' Charity Lectureship award (MMBG1R).

Acknowledgments: We regret that we were unable to cite all relevant publications in this field due to space constraints, and refer readers to other reviews in this Special Issue for further insight into the regulation of host genomes through endogenous retroviruses and other transposable elements.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Theunissen, T.W.; Jaenisch, R. Mechanisms of gene regulation in human embryos and pluripotent stem cells. *Development* **2017**, *144*, 4496–4509. [[CrossRef](#)] [[PubMed](#)]
2. de Koning, A.P.; Gu, W.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* **2011**, *7*, e1002384. [[CrossRef](#)] [[PubMed](#)]
3. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [[CrossRef](#)] [[PubMed](#)]
4. Babaian, A.; Mager, D.L. Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* **2016**, *7*, 24. [[CrossRef](#)]

5. Friedli, M.; Trono, D. The developmental control of transposable elements and the evolution of higher species. *Annu. Rev. Cell Dev. Biol.* **2015**, *31*, 429–451. [[CrossRef](#)]
6. Percharde, M.; Sultana, T.; Ramalho-Santos, M. What Doesn't Kill You Makes You Stronger: Transposons as Dual Players in Chromatin Regulation and Genomic Variation. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **2020**, *42*, e1900232. [[CrossRef](#)]
7. Robbez-Masson, L.; Rowe, H.M. Retrotransposons shape species-specific embryonic stem cell gene expression. *Retrovirology* **2015**, *12*, 45. [[CrossRef](#)]
8. Thompson, P.J.; Macfarlan, T.S.; Lorincz, M.C. Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. *Mol. Cell* **2016**, *62*, 766–776. [[CrossRef](#)]
9. Bruno, M.; Mahgoub, M.; Macfarlan, T.S. The Arms Race Between KRAB-Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *Annu. Rev. Genet.* **2019**, *53*, 393–416. [[CrossRef](#)]
10. Deniz, O.; Frost, J.M.; Branco, M.R. Regulation of transposable elements by DNA modifications. *Nat. Rev. Genet.* **2019**, *20*, 417–431. [[CrossRef](#)]
11. Fukuda, K.; Okuda, A.; Yusa, K.; Shinkai, Y. A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells. *Genome Res.* **2018**, *28*, 846–858. [[CrossRef](#)] [[PubMed](#)]
12. Magiorkinis, G.; Blanco-Melo, D.; Belshaw, R. The decline of human endogenous retroviruses: Extinction and survival. *Retrovirology* **2015**, *12*, 8. [[CrossRef](#)]
13. Brouha, B.; Schustak, J.; Badge, R.M.; Lutz-Prigge, S.; Farley, A.H.; Moran, J.V.; Kazazian, H.H., Jr. Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5280–5285. [[CrossRef](#)] [[PubMed](#)]
14. Gianfrancesco, O.; Geary, B.; Savage, A.L.; Billingsley, K.J.; Bubb, V.J.; Quinn, J.P. The Role of SINE-VNTR-Alu (SVA) Retrotransposons in Shaping the Human Genome. *Int. J. Mol. Sci.* **2019**, *20*, 5977. [[CrossRef](#)] [[PubMed](#)]
15. Campos-Sanchez, R.; Cremona, M.A.; Pini, A.; Chiaromonte, F.; Makova, K.D. Integration and Fixation Preferences of Human and Mouse Endogenous Retroviruses Uncovered with Functional Data Analysis. *PLoS Comput. Biol.* **2016**, *12*, e1004956. [[CrossRef](#)] [[PubMed](#)]
16. Goodier, J.L.; Ostertag, E.M.; Du, K.; Kazazian, H.H., Jr. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* **2001**, *11*, 1677–1685. [[CrossRef](#)] [[PubMed](#)]
17. Ribet, D.; Harper, F.; Dupressoir, A.; Dewannieux, M.; Pierron, G.; Heidmann, T. An infectious progenitor for the murine IAP retrotransposon: Emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome Res.* **2008**, *18*, 597–609. [[CrossRef](#)]
18. Kazachenka, A.; Bertozzi, T.M.; Sjoberg-Herrera, M.K.; Walker, N.; Gardner, J.; Gunning, R.; Pahita, E.; Adams, S.; Adams, D.; Ferguson-Smith, A.C. Identification, Characterization, and Heritability of Murine Metastable Epialleles: Implications for Non-genetic Inheritance. *Cell* **2018**, *175*, 1717. [[CrossRef](#)]
19. Lilue, J.; Doran, A.G.; Fiddes, I.T.; Abrudan, M.; Armstrong, J.; Bennett, R.; Chow, W.; Collins, J.; Collins, S.; Czechanski, A.; et al. Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. *Nat. Genet.* **2018**, *50*, 1574–1583. [[CrossRef](#)]
20. Rebollo, R.; Karimi, M.M.; Bilenky, M.; Gagnier, L.; Miceli-Royer, K.; Zhang, Y.; Goyal, P.; Keane, T.M.; Jones, S.; Hirst, M.; et al. Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet.* **2011**, *7*, e1002301. [[CrossRef](#)]
21. Rowe, H.M.; Kapopoulou, A.; Corsinotti, A.; Fasching, L.; Macfarlan, T.S.; Tarabay, Y.; Viville, S.; Jakobsson, J.; Pfaff, S.L.; Trono, D. TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. *Genome Res.* **2013**, *23*, 452–461. [[CrossRef](#)] [[PubMed](#)]
22. Wang, J.; Vicente-Garcia, C.; Seruggia, D.; Molto, E.; Fernandez-Minan, A.; Neto, A.; Lee, E.; Gomez-Skarmeta, J.L.; Montoliu, L.; Lunyak, V.V.; et al. MIR retrotransposon sequences provide insulators to the human genome. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E4428–E4437. [[CrossRef](#)]
23. Macfarlan, T.S.; Gifford, W.D.; Driscoll, S.; Lettieri, K.; Rowe, H.M.; Bonanomi, D.; Firth, A.; Singer, O.; Trono, D.; Pfaff, S.L. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **2012**, *487*, 57–63. [[CrossRef](#)] [[PubMed](#)]
24. Chuong, E.B.; Elde, N.C.; Feschotte, C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **2016**, *351*, 1083–1087. [[CrossRef](#)]
25. Eckersley-Maslin, M.A.; Alda-Catalinas, C.; Reik, W. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 436–450. [[CrossRef](#)] [[PubMed](#)]

26. Torres-Padilla, M.E. On transposons and totipotency. *Philos. Trans. R. Soc. B Biol. Sci.* **2020**, *375*, 20190339. [[CrossRef](#)]
27. Izsvak, Z.; Wang, J.; Singh, M.; Mager, D.L.; Hurst, L.D. Pluripotency and the endogenous retrovirus HERVH: Conflict or serendipity? *BioEssays News Rev. Mol. Cell. Dev. Biol.* **2016**, *38*, 109–117. [[CrossRef](#)]
28. Peaston, A.E.; Evsikov, A.V.; Graber, J.H.; de Vries, W.N.; Holbrook, A.E.; Solter, D.; Knowles, B.B. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **2004**, *7*, 597–606. [[CrossRef](#)]
29. Ishiuchi, T.; Enriquez-Gasca, R.; Mizutani, E.; Boskovic, A.; Ziegler-Birling, C.; Rodriguez-Terrones, D.; Wakayama, T.; Vaquerizas, J.M.; Torres-Padilla, M.E. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **2015**, *22*, 662–671. [[CrossRef](#)]
30. Eckersley-Maslin, M.A.; Svensson, V.; Krueger, C.; Stubbs, T.M.; Giehr, P.; Krueger, F.; Miragaia, R.J.; Kyriakopoulos, C.; Berrens, R.V.; Milagre, I.; et al. MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs. *Cell Rep.* **2016**, *17*, 179–192. [[CrossRef](#)]
31. Alda-Catalinas, C.; Bredikhin, D.; Hernando-Herraez, I.; Santos, F.; Kubinyecz, O.; Eckersley-Maslin, M.A.; Stegle, O.; Reik, W. A Single-Cell Transcriptomics CRISPR-Activation Screen Identifies Epigenetic Regulators of the Zygotic Genome Activation Program. *Cell Syst.* **2020**. [[CrossRef](#)] [[PubMed](#)]
32. De Iaco, A.; Coudray, A.; Duc, J.; Trono, D. DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells. *EMBO Rep.* **2019**, *20*. [[CrossRef](#)] [[PubMed](#)]
33. De Iaco, A.; Planet, E.; Coluccio, A.; Verp, S.; Duc, J.; Trono, D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **2017**, *49*, 941–945. [[CrossRef](#)]
34. Eckersley-Maslin, M.; Alda-Catalinas, C.; Blotenburg, M.; Kreibich, E.; Krueger, C.; Reik, W. Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev.* **2019**, *33*, 194–208. [[CrossRef](#)] [[PubMed](#)]
35. Hendrickson, P.G.; Dorais, J.A.; Grow, E.J.; Whiddon, J.L.; Lim, J.W.; Wike, C.L.; Weaver, B.D.; Pflueger, C.; Emery, B.R.; Wilcox, A.L.; et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **2017**, *49*, 925–934. [[CrossRef](#)]
36. Young, J.M.; Whiddon, J.L.; Yao, Z.; Kasinathan, B.; Snider, L.; Geng, L.N.; Balog, J.; Tawil, R.; van der Maarel, S.M.; Tapscott, S.J. DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet.* **2013**, *9*, e1003947. [[CrossRef](#)]
37. Zhang, W.; Chen, F.; Chen, R.; Xie, D.; Yang, J.; Zhao, X.; Guo, R.; Zhang, Y.; Shen, Y.; Goke, J.; et al. Zscan4c activates endogenous retrovirus MERVL and cleavage embryo genes. *Nucleic Acids Res.* **2019**, *47*, 8485–8501. [[CrossRef](#)]
38. Srinivasan, R.; Nady, N.; Arora, N.; Hsieh, L.J.; Swigut, T.; Narlikar, G.J.; Wossidlo, M.; Wysocka, J. Zscan4 binds nucleosomal microsatellite DNA and protects mouse two-cell embryos from DNA damage. *Sci. Adv.* **2020**, *6*, eaaz9115. [[CrossRef](#)]
39. Wu, J.; Huang, B.; Chen, H.; Yin, Q.; Liu, Y.; Xiang, Y.; Zhang, B.; Liu, B.; Wang, Q.; Xia, W.; et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **2016**, *534*, 652–657. [[CrossRef](#)]
40. Jachowicz, J.W.; Bing, X.; Pontabry, J.; Boskovic, A.; Rando, O.J.; Torres-Padilla, M.E. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* **2017**, *49*, 1502–1510. [[CrossRef](#)]
41. Percharde, M.; Lin, C.J.; Yin, Y.; Guan, J.; Peixoto, G.A.; Bulut-Karslioglu, A.; Biechele, S.; Huang, B.; Shen, X.; Ramalho-Santos, M. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* **2018**, *174*, 391–405.e319. [[CrossRef](#)] [[PubMed](#)]
42. Liu, L.; Leng, L.; Liu, C.; Lu, C.; Yuan, Y.; Wu, L.; Gong, F.; Zhang, S.; Wei, X.; Wang, M.; et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nat. Commun.* **2019**, *10*, 364. [[CrossRef](#)] [[PubMed](#)]
43. Wolf, G.; de Iaco, A.; Sun, M.A.; Bruno, M.; Tinkham, M.; Hoang, D.; Mitra, A.; Ralls, S.; Trono, D.; Macfarlan, T.S. KRAB-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *eLife* **2020**, *9*. [[CrossRef](#)]
44. Karimi, M.M.; Goyal, P.; Maksakova, I.A.; Bilenky, M.; Leung, D.; Tang, J.X.; Shinkai, Y.; Mager, D.L.; Jones, S.; Hirst, M.; et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **2011**, *8*, 676–687. [[CrossRef](#)]

45. Matsui, T.; Leung, D.; Miyashita, H.; Maksakova, I.A.; Miyachi, H.; Kimura, H.; Tachibana, M.; Lorincz, M.C.; Shinkai, Y. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **2010**, *464*, 927–931. [[CrossRef](#)] [[PubMed](#)]
46. Rowe, H.M.; Friedli, M.; Offner, S.; Verp, S.; Mesnard, D.; Marquis, J.; Aktas, T.; Trono, D. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development* **2013**, *140*, 519–529. [[CrossRef](#)]
47. Rowe, H.M.; Jakobsson, J.; Mesnard, D.; Rougemont, J.; Reynard, S.; Aktas, T.; Maillard, P.V.; Layard-Liesching, H.; Verp, S.; Marquis, J.; et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **2010**, *463*, 237–240. [[CrossRef](#)] [[PubMed](#)]
48. Ecco, G.; Cassano, M.; Kauzlaric, A.; Duc, J.; Coluccio, A.; Offner, S.; Imbeault, M.; Rowe, H.M.; Turelli, P.; Trono, D. Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Dev. Cell* **2016**, *36*, 611–623. [[CrossRef](#)]
49. Fasching, L.; Kapopoulou, A.; Sachdeva, R.; Petri, R.; Jonsson, M.E.; Manne, C.; Turelli, P.; Jern, P.; Cammas, F.; Trono, D.; et al. TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells. *Cell Rep.* **2015**, *10*, 20–28. [[CrossRef](#)]
50. Hummel, B.; Hansen, E.C.; Yoveva, A.; Aprile-Garcia, F.; Hussong, R.; Sawarkar, R. The evolutionary capacitor HSP90 buffers the regulatory effects of mammalian endogenous retroviruses. *Nat. Struct. Mol. Biol.* **2017**, *24*, 234–242. [[CrossRef](#)]
51. Seah, M.K.Y.; Wang, Y.; Goy, P.A.; Loh, H.M.; Peh, W.J.; Low, D.H.P.; Han, B.Y.; Wong, E.; Leong, E.L.; Wolf, G.; et al. The KRAB-zinc-finger protein ZFP708 mediates epigenetic repression at RMER19B retrotransposons. *Development* **2019**, *146*. [[CrossRef](#)] [[PubMed](#)]
52. Duhl, D.M.; Vrieling, H.; Miller, K.A.; Wolff, G.L.; Barsh, G.S. Neomorphic agouti mutations in obese yellow mice. *Nat. Genet.* **1994**, *8*, 59–65. [[CrossRef](#)] [[PubMed](#)]
53. Morgan, H.D.; Sutherland, H.G.; Martin, D.I.; Whitelaw, E. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **1999**, *23*, 314–318. [[CrossRef](#)] [[PubMed](#)]
54. Rakyan, V.K.; Chong, S.; Champ, M.E.; Cuthbert, P.C.; Morgan, H.D.; Luu, K.V.; 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*, 2538–2543. [[CrossRef](#)]
55. Elmer, J.L.; Ferguson-Smith, A.C. Strain-Specific Epigenetic Regulation of Endogenous Retroviruses: The Role of Trans-Acting Modifiers. *Viruses* **2020**, *12*, 810. [[CrossRef](#)]
56. Rebollo, R.; Galvao-Ferrarini, M.; Gagnier, L.; Zhang, Y.; Ferraj, A.; Beck, C.R.; Lorincz, M.C.; Mager, D.L. Inter-Strain Epigenomic Profiling Reveals a Candidate IAP Master Copy in C3H Mice. *Viruses* **2020**, *12*, 783. [[CrossRef](#)]
57. Sanchez-Luque, F.J.; Kempen, M.H.C.; Gerdes, P.; Vargas-Landin, D.B.; Richardson, S.R.; Troskie, R.L.; Jesuadian, J.S.; Cheetham, S.W.; Carreira, P.E.; Salvador-Palomeque, C.; et al. LINE-1 Evasion of Epigenetic Repression in Humans. *Mol. Cell* **2019**, *75*, 590–604.e512. [[CrossRef](#)]
58. Tchasovnikarova, I.A.; Timms, R.T.; Matheson, N.J.; Wals, K.; Antrobus, R.; Gottgens, B.; Dougan, G.; Dawson, M.A.; Lehner, P.J. GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **2015**, *348*, 1481–1485. [[CrossRef](#)]
59. Maeda-Smithies, N.; Hiller, S.; Dong, S.; Kim, H.S.; Bennett, B.J.; Kayashima, Y. Ectopic expression of the Stabilin2 gene triggered by an intracisternal A particle (IAP) element in DBA/2J strain of mice. *Mamm. Genome* **2020**, *31*, 2–16. [[CrossRef](#)]
60. Jonsson, M.E.; Ludvik Brattas, P.; Gustafsson, C.; Petri, R.; Yudovich, D.; Piracs, K.; Verschuere, S.; Madsen, S.; Hansson, J.; Larsson, J.; et al. Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. *Nat. Commun.* **2019**, *10*, 3182. [[CrossRef](#)]
61. Denli, A.M.; Narvaiza, I.; Kerman, B.E.; Pena, M.; Benner, C.; Marchetto, M.C.; Diedrich, J.K.; Aslanian, A.; Ma, J.; Moresco, J.J.; et al. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell* **2015**, *163*, 583–593. [[CrossRef](#)] [[PubMed](#)]
62. Jacobs, F.M.; Greenberg, D.; Nguyen, N.; Haeussler, M.; Ewing, A.D.; Katzman, S.; Paten, B.; Salama, S.R.; Haussler, D. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* **2014**, *516*, 242–245. [[CrossRef](#)] [[PubMed](#)]

63. Turelli, P.; Castro-Diaz, N.; Marzetta, F.; Kapopoulou, A.; Raclot, C.; Duc, J.; Tieng, V.; Quenneville, S.; Trono, D. Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. *Genome Res.* **2014**, *24*, 1260–1270. [[CrossRef](#)] [[PubMed](#)]
64. Friedli, M.; Turelli, P.; Kapopoulou, A.; Rauwel, B.; Castro-Diaz, N.; Rowe, H.M.; Ecco, G.; Unzu, C.; Planet, E.; Lombardo, A.; et al. Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. *Genome Res.* **2014**, *24*, 1251–1259. [[CrossRef](#)] [[PubMed](#)]
65. Wolf, D.; Goff, S.P. TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* **2007**, *131*, 46–57. [[CrossRef](#)]
66. Wolf, D.; Goff, S.P. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **2009**, *458*, 1201–1204. [[CrossRef](#)]
67. Wolf, G.; Yang, P.; Fuchtbauer, A.C.; Fuchtbauer, E.M.; Silva, A.M.; Park, C.; Wu, W.; Nielsen, A.L.; Pedersen, F.S.; Macfarlan, T.S. The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. *Genes Dev.* **2015**, *29*, 538–554. [[CrossRef](#)]
68. Tie, C.H.; Fernandes, L.; Conde, L.; Robbez-Masson, L.; Sumner, R.P.; Peacock, T.; Rodriguez-Plata, M.T.; Mickute, G.; Gifford, R.; Towers, G.J.; et al. KAP1 regulates endogenous retroviruses in adult human cells and contributes to innate immune control. *EMBO Rep.* **2018**, *19*. [[CrossRef](#)]
69. Wolf, D.; Hug, K.; Goff, S.P. TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12521–12526. [[CrossRef](#)]
70. Wang, J.; Xie, G.; Singh, M.; Ghanbarian, A.T.; Rasko, T.; Szvetnik, A.; Cai, H.; Besser, D.; Prigione, A.; Fuchs, N.V.; et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **2014**, *516*, 405–409. [[CrossRef](#)]
71. Romer, C.; Singh, M.; Hurst, L.D.; Izsvak, Z. How to tame an endogenous retrovirus: HERVH and the evolution of human pluripotency. *Curr. Opin. Virol.* **2017**, *25*, 49–58. [[CrossRef](#)]
72. Grow, E.J.; Flynn, R.A.; Chavez, S.L.; Bayless, N.L.; Wossidlo, M.; Wesche, D.J.; Martin, L.; Ware, C.B.; Blish, C.A.; Chang, H.Y.; et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **2015**, *522*, 221–225. [[CrossRef](#)]
73. Collins, P.L.; Kyle, K.E.; Egawa, T.; Shinkai, Y.; Oltz, E.M. The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8367–8372. [[CrossRef](#)] [[PubMed](#)]
74. Faulkner, G.J.; Kimura, Y.; Daub, C.O.; Wani, S.; Plessy, C.; Irvine, K.M.; Schroder, K.; Cloonan, N.; Steptoe, A.L.; Lassmann, T.; et al. The regulated retrotransposon transcriptome of mammalian cells. *Nat. Genet.* **2009**, *41*, 563–571. [[CrossRef](#)] [[PubMed](#)]
75. Kato, M.; Takemoto, K.; Shinkai, Y. A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. *Nat. Commun.* **2018**, *9*, 1683. [[CrossRef](#)] [[PubMed](#)]
76. Pehrsson, E.C.; Choudhary, M.N.K.; Sundaram, V.; Wang, T. The epigenomic landscape of transposable elements across normal human development and anatomy. *Nat. Commun.* **2019**, *10*, 5640. [[CrossRef](#)] [[PubMed](#)]
77. Trizzino, M.; Park, Y.; Holsbach-Beltrame, M.; Aracena, K.; Mika, K.; Caliskan, M.; Perry, G.H.; Lynch, V.J.; Brown, C.D. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Res.* **2017**, *27*, 1623–1633. [[CrossRef](#)]
78. Villar, D.; Berthelot, C.; Aldridge, S.; Rayner, T.F.; Lukk, M.; Pignatelli, M.; Park, T.J.; Deaville, R.; Erichsen, J.T.; Jasinska, A.J.; et al. Enhancer Evolution across 20 Mammalian Species. *Cell* **2015**, *160*, 554–566. [[CrossRef](#)]
79. Kaneko, S.; Bonasio, R.; Saldana-Meyer, R.; Yoshida, T.; Son, J.; Nishino, K.; Umezawa, A.; Reinberg, D. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol. Cell* **2014**, *53*, 290–300. [[CrossRef](#)]
80. Tchasovnikarova, I.A.; Timms, R.T.; Douse, C.H.; Roberts, R.C.; Dougan, G.; Kingston, R.E.; Modis, Y.; Lehner, P.J. Hyperactivation of HUSH complex function by Charcot-Marie-Tooth disease mutation in MORC2. *Nat. Genet.* **2017**, *49*, 1035–1044. [[CrossRef](#)]
81. Douse, C.H.; Bloor, S.; Liu, Y.; Shamin, M.; Tchasovnikarova, I.A.; Timms, R.T.; Lehner, P.J.; Modis, Y. Neuropathic MORC2 mutations perturb GHKL ATPase dimerization dynamics and epigenetic silencing by multiple structural mechanisms. *Nat. Commun.* **2018**, *9*, 651. [[CrossRef](#)]
82. Liu, N.; Lee, C.H.; Swigut, T.; Grow, E.; Gu, B.; Bassik, M.C.; Wysocka, J. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* **2018**, *553*, 228–232. [[CrossRef](#)] [[PubMed](#)]

83. Zhu, Y.; Wang, G.Z.; Cingoz, O.; Goff, S.P. NP220 mediates silencing of unintegrated retroviral DNA. *Nature* **2018**, *564*, 278–282. [[CrossRef](#)] [[PubMed](#)]
84. Harten, S.K.; Bruxner, T.J.; Bharti, V.; Blewitt, M.; Nguyen, T.M.; Whitelaw, E.; Epp, T. The first mouse mutants of D14Abb1e (Fam208a) show that it is critical for early development. *Mamm. Genome* **2014**, *25*, 293–303. [[CrossRef](#)] [[PubMed](#)]
85. Robbez-Masson, L.; Tie, C.H.C.; Conde, L.; Tunbak, H.; Husovsky, C.; Tchasovnikarova, I.A.; Timms, R.T.; Herrero, J.; Lehner, P.J.; Rowe, H.M. The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res.* **2018**. [[CrossRef](#)] [[PubMed](#)]
86. Ardeljan, D.; Steranka, J.P.; Liu, C.; Li, Z.; Taylor, M.S.; Payer, L.M.; Gorbounov, M.; Sarnecki, J.S.; Deshpande, V.; Hruban, R.H.; et al. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. *Nat. Struct. Mol. Biol.* **2020**, *27*, 168–178. [[CrossRef](#)] [[PubMed](#)]
87. Douse, C.H.; Tchasovnikarova, I.A.; Timms, R.T.; Protasio, A.V.; Seczynska, M.; Prigozhin, D.M.; Albecka, A.; Wagstaff, J.; Williamson, J.C.; Freund, S.M.V.; et al. TASOR Is a pseudo-PARP that Directs HUSH Complex Assembly and Epigenetic Transposon Control. *bioRxiv.* **2020**. Available online: <https://www.biorxiv.org/content/10.1101/2020.03.09.974832v1> (accessed on 3 August 2020).
88. Ferguson-Smith, A.C. Genomic imprinting: The emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **2011**, *12*, 565–575. [[CrossRef](#)]
89. Reik, W.; Walter, J. Genomic imprinting: Parental influence on the genome. *Nat. Rev. Genet.* **2001**, *2*, 21–32. [[CrossRef](#)]
90. Barlow, D.P. Methylation and imprinting: From host defense to gene regulation? *Science* **1993**, *260*, 309–310. [[CrossRef](#)]
91. Ondicova, M.; Oakey, R.J.; Walsh, C.P. Is imprinting the result of “friendly fire” by the host defense system? *PLoS Genet.* **2020**, *16*, e1008599. [[CrossRef](#)]
92. Wood, A.J.; Bourc’his, D.; Bestor, T.H.; Oakey, R.J. Allele-specific demethylation at an imprinted mammalian promoter. *Nucleic Acids Res.* **2007**, *35*, 7031–7039. [[CrossRef](#)]
93. Wood, A.J.; Roberts, R.G.; Monk, D.; Moore, G.E.; Schulz, R.; Oakey, R.J. A screen for retrotransposed imprinted genes reveals an association between X chromosome homology and maternal germ-line methylation. *PLoS Genet.* **2007**, *3*, e20. [[CrossRef](#)]
94. Youngson, N.A.; Kocialkowski, S.; Peel, N.; Ferguson-Smith, A.C. A small family of sushi-class retrotransposon-derived genes in mammals and their relation to genomic imprinting. *J. Mol. Evol.* **2005**, *61*, 481–490. [[CrossRef](#)] [[PubMed](#)]
95. Bogutz, A.B.; Brind’Amour, J.; Kobayashi, H.; Jensen, K.N.; Nakabayashi, K.; Imai, H.; Lorincz, M.C.; Lefebvre, L. Evolution of imprinting via lineage-specific insertion of retroviral promoters. *Nat. Commun.* **2019**, *10*, 5674. [[CrossRef](#)]
96. Hanna, C.W.; Perez-Palacios, R.; Gahurova, L.; Schubert, M.; Krueger, F.; Biggins, L.; Andrews, S.; Colome-Tatche, M.; Bourc’his, D.; Dean, W.; et al. Endogenous retroviral insertions drive non-canonical imprinting in extra-embryonic tissues. *Genome Biol.* **2019**, *20*, 225. [[CrossRef](#)] [[PubMed](#)]
97. Takahashi, N.; Coluccio, A.; Thorball, C.W.; Planet, E.; Shi, H.; Offner, S.; Turelli, P.; Imbeault, M.; Ferguson-Smith, A.C.; Trono, D. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* **2019**, *33*, 49–54. [[CrossRef](#)] [[PubMed](#)]
98. Imbeault, M.; Helleboid, P.Y.; Trono, D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **2017**, *543*, 550–554. [[CrossRef](#)] [[PubMed](#)]
99. Strogantsev, R.; Krueger, F.; Yamazawa, K.; Shi, H.; Gould, P.; Goldman-Roberts, M.; McEwen, K.; Sun, B.; Pedersen, R.; Ferguson-Smith, A.C. Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. *Genome Biol.* **2015**, *16*, 112. [[CrossRef](#)]
100. Swanzey, E.; McNamara, T.F.; Apostolou, E.; Tahilian, M.; Stadtfeld, M. A Susceptibility Locus on Chromosome 13 Profoundly Impacts the Stability of Genomic Imprinting in Mouse Pluripotent Stem Cells. *Cell Rep.* **2020**, *30*, 3597–3604.e3593. [[CrossRef](#)]
101. Krebs, C.J.; Larkins, L.K.; Khan, S.M.; Robins, D.M. Expansion and diversification of KRAB zinc-finger genes within a cluster including Regulator of sex-limitation 1 and 2. *Genomics* **2005**, *85*, 752–761. [[CrossRef](#)]
102. Barreiro, L.B.; Marionni, J.C.; Blekhman, R.; Stephens, M.; Gilad, Y. Functional comparison of innate immune signaling pathways in primates. *PLoS Genet.* **2010**, *6*, e1001249. [[CrossRef](#)] [[PubMed](#)]

103. Plataniias, L.C. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat. Rev. Immunol.* **2005**, *5*, 375–386. [[CrossRef](#)] [[PubMed](#)]
104. Schroder, K.; Irvine, K.M.; Taylor, M.S.; Bokil, N.J.; Le Cao, K.A.; Masterman, K.A.; Labzin, L.I.; Semple, C.A.; Kapetanovic, R.; Fairbairn, L.; et al. Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E944–E953. [[CrossRef](#)] [[PubMed](#)]
105. Bao, W.; Kojima, K.K.; Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **2015**, *6*, 11. [[CrossRef](#)]
106. Silva, J.C.; Shabalina, S.A.; Harris, D.G.; Spouge, J.L.; Kondrashovi, A.S. 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. [[CrossRef](#)]
107. Brockdorff, N.; Ashworth, A.; Kay, G.F.; Cooper, P.; Smith, S.; McCabe, V.M.; Norris, D.P.; Penny, G.D.; Patel, D.; Rastan, S. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **1991**, *351*, 329–331. [[CrossRef](#)] [[PubMed](#)]
108. Loda, A.; Heard, E. Xist RNA in action: Past, present, and future. *PLoS Genet.* **2019**, *15*, e1008333. [[CrossRef](#)] [[PubMed](#)]
109. Lyon, M.F. X-chromosome inactivation: A repeat hypothesis. *Cytogenet. Cell Genet.* **1998**, *80*, 133–137. [[CrossRef](#)]
110. Bailey, J.A.; Carrel, L.; Chakravarti, A.; Eichler, E.E. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: The Lyon repeat hypothesis. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6634–6639. [[CrossRef](#)] [[PubMed](#)]
111. Chow, J.C.; Ciaudo, C.; Fazzari, M.J.; Mise, N.; Servant, N.; Glass, J.L.; Attreed, M.; Avner, P.; Wutz, A.; Barillot, E.; et al. LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* **2010**, *141*, 956–969. [[CrossRef](#)] [[PubMed](#)]
112. Larson, A.G.; Elnatan, D.; Keenen, M.M.; Trnka, M.J.; Johnston, J.B.; Burlingame, A.L.; Agard, D.A.; Redding, S.; Narlikar, G.J. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **2017**, *547*, 236–240. [[CrossRef](#)] [[PubMed](#)]
113. Strom, A.R.; Emelyanov, A.V.; Mir, M.; Fyodorov, D.V.; Darzacq, X.; Karpen, G.H. Phase separation drives heterochromatin domain formation. *Nature* **2017**, *547*, 241–245. [[CrossRef](#)] [[PubMed](#)]
114. Britten, R.J.; Davidson, E.H. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q. Rev. Biol.* **1971**, *46*, 111–138. [[CrossRef](#)]
115. Tashiro, K.; Teissier, A.; Kobayashi, N.; Nakanishi, A.; Sasaki, T.; Yan, K.; Tarabykin, V.; Vigier, L.; Sumiyama, K.; Hirakawa, M.; et al. A mammalian conserved element derived from SINE displays enhancer properties recapitulating Satb2 expression in early-born callosal projection neurons. *PLoS ONE* **2011**, *6*, e28497. [[CrossRef](#)] [[PubMed](#)]
116. Nishihara, H.; Smit, A.F.; Okada, N. Functional noncoding sequences derived from SINEs in the mammalian genome. *Genome Res.* **2006**, *16*, 864–874. [[CrossRef](#)]
117. Alcamo, E.A.; Chirivella, L.; Dautzenberg, M.; Dobрева, G.; Farinas, I.; Grosschedl, R.; McConnell, S.K. Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. *Neuron* **2008**, *57*, 364–377. [[CrossRef](#)] [[PubMed](#)]
118. Flower, W.H. On the Commissures of the Cerebral Hemispheres of the Marsupialia and Monotremata, as Compared with Those of the Placental Mammals. *Proc. R. Soc. Lond.* **1865**, *14*, 71–74.
119. Notwell, J.H.; Chung, T.; Heavner, W.; Bejerano, G. A family of transposable elements co-opted into developmental enhancers in the mouse neocortex. *Nat. Commun.* **2015**, *6*, 6644. [[CrossRef](#)] [[PubMed](#)]
120. Jjingo, D.; Huda, A.; Gundapuneni, M.; Marino-Ramirez, L.; Jordan, I.K. Effect of the transposable element environment of human genes on gene length and expression. *Genome Biol. Evol.* **2011**, *3*, 259–271. [[CrossRef](#)]
121. Samstein, R.M.; Josefowicz, S.Z.; Arvey, A.; Treuting, P.M.; Rudensky, A.Y. Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict. *Cell* **2012**, *150*, 29–38. [[CrossRef](#)]
122. Lynch, V.J.; Leclerc, R.D.; May, G.; Wagner, G.P. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* **2011**, *43*, 1154–1159. [[CrossRef](#)] [[PubMed](#)]
123. Lamprecht, B.; Walter, K.; Kreher, S.; Kumar, R.; Hummel, M.; Lenze, D.; Kochert, K.; Bouhlel, M.A.; Richter, J.; Soler, E.; et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat. Med.* **2010**, *16*, 571–579. [[CrossRef](#)] [[PubMed](#)]

124. Kreher, S.; Bouhleb, M.A.; Cauchy, P.; Lamprecht, B.; Li, S.; Grau, M.; Hummel, F.; Kochert, K.; Anagnostopoulos, I.; Johrens, K.; et al. Mapping of transcription factor motifs in active chromatin identifies IRF5 as key regulator in classical Hodgkin lymphoma. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E4513–E4522. [[CrossRef](#)]
125. Babaian, A.; Romanish, M.T.; Gagnier, L.; Kuo, L.Y.; Karimi, M.M.; Steidl, C.; Mager, D.L. Onco-exaptation of an endogenous retroviral LTR drives IRF5 expression in Hodgkin lymphoma. *Oncogene* **2016**, *35*, 2542–2546. [[CrossRef](#)] [[PubMed](#)]
126. Leung, A.; Trac, C.; Kato, H.; Costello, K.R.; Chen, Z.; Natarajan, R.; Schones, D.E. LTRs activated by Epstein-Barr virus-induced transformation of B cells alter the transcriptome. *Genome Res.* **2018**, *28*, 1791–1798. [[CrossRef](#)]
127. Deniz, O.; Ahmed, M.; Todd, C.D.; Rio-Machin, A.; Dawson, M.A.; Branco, M.R. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat. Commun.* **2020**, *11*, 3506. [[CrossRef](#)]
128. Ng, K.W.; Attig, J.; Young, G.R.; Ottina, E.; Papamichos, S.I.; Kotsianidis, I.; Kassiotis, G. Soluble PD-L1 generated by endogenous retroelement exaptation is a receptor antagonist. *eLife* **2019**, *8*. [[CrossRef](#)]
129. Groger, V.; Cynis, H. Human Endogenous Retroviruses and Their Putative Role in the Development of Autoimmune Disorders such as Multiple Sclerosis. *Front. Microbiol.* **2018**, *9*, 265. [[CrossRef](#)]
130. Rolland, A.; Jouvin-Marche, E.; Viret, C.; Faure, M.; Perron, H.; Marche, P.N. The envelope protein of a human endogenous retrovirus-W family activates innate immunity through CD14/TLR4 and promotes Th1-like responses. *J. Immunol.* **2006**, *176*, 7636–7644. [[CrossRef](#)]
131. Wu, Z.; Mei, X.; Zhao, D.; Sun, Y.; Song, J.; Pan, W.; Shi, W. DNA methylation modulates HERV-E expression in CD4+ T cells from systemic lupus erythematosus patients. *J. Dermatol. Sci.* **2015**, *77*, 110–116. [[CrossRef](#)]
132. Sukapan, P.; Promnarate, P.; Avihingsanon, Y.; Mutirangura, A.; Hirankarn, N. Types of DNA methylation status of the interspersed repetitive sequences for LINE-1, Alu, HERV-E and HERV-K in the neutrophils from systemic lupus erythematosus patients and healthy controls. *J. Hum. Genet.* **2014**, *59*, 178–188. [[CrossRef](#)] [[PubMed](#)]
133. Treger, R.S.; Pope, S.D.; Kong, Y.; Tokuyama, M.; Taura, M.; Iwasaki, A. The Lupus Susceptibility Locus Sgp3 Encodes the Suppressor of Endogenous Retrovirus Expression SNERV. *Immunity* **2019**, *50*, 334–347.e9. [[CrossRef](#)] [[PubMed](#)]
134. Jonsson, M.E.; Garza, R.; Johansson, P.A.; Jakobsson, J. Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet.* **2020**, *36*, 610–623. [[CrossRef](#)] [[PubMed](#)]
135. Tam, O.H.; Ostrow, L.W.; Gale Hammell, M. Diseases of the nERVous system: Retrotransposon activity in neurodegenerative disease. *Mob. DNA* **2019**, *10*, 32. [[CrossRef](#)]
136. Benoit, M.; Drost, H.G.; Catoni, M.; Gouil, Q.; Lopez-Gomollon, S.; Baulcombe, D.; Paszkowski, J. Environmental and epigenetic regulation of Rider retrotransposons in tomato. *PLoS Genet.* **2019**, *15*, e1008370. [[CrossRef](#)]
137. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)]
138. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
139. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
140. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
141. Jin, Y.; Tam, O.H.; Paniagua, E.; Hammell, M. TETranscripts: A package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **2015**, *31*, 3593–3599. [[CrossRef](#)]
142. Nellaker, C.; Keane, T.M.; Yalcin, B.; Wong, K.; Agam, A.; Belgard, T.G.; Flint, J.; Adams, D.J.; Frankel, W.N.; Ponting, C.P. The genomic landscape shaped by selection on transposable elements across 18 mouse strains. *Genome Biol.* **2012**, *13*, R45. [[CrossRef](#)]
143. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)]

144. Wenger, A.M.; Clarke, S.L.; Notwell, J.H.; Chung, T.; Tuteja, G.; Guturu, H.; Schaar, B.T.; Bejerano, G. The enhancer landscape during early neocortical development reveals patterns of dense regulation and co-option. *PLoS Genet.* **2013**, *9*, e1003728. [[CrossRef](#)]
145. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74. [[CrossRef](#)]
146. Pollard, K.S.; Hubisz, M.J.; Rosenbloom, K.R.; Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **2010**, *20*, 110–121. [[CrossRef](#)]
147. Karolchik, D.; Hinrichs, A.S.; Furey, T.S.; Roskin, K.M.; Sugnet, C.W.; Haussler, D.; Kent, W.J. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **2004**, *32*, D493–D496. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).