



Transposable Elements in Pluripotent Stem Cells and Human Disease

Gang Ma, Isaac A. Babarinde, Xuemeng Zhou and Andrew P. Hutchins*

Shenzhen Key Laboratory of Gene Regulation and Systems Biology, Department of Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen, China

Transposable elements (TEs) are mobile genetic elements that can randomly integrate into other genomic sites. They have successfully replicated and now occupy around 40% of the total DNA sequence in humans. TEs in the genome have a complex relationship with the host cell, being both potentially deleterious and advantageous at the same time. Only a tiny minority of TEs are still capable of transposition, yet their fossilized sequence fragments are thought to be involved in various molecular processes, such as gene transcriptional activity, RNA stability and subcellular localization, and chromosomal architecture. TEs have also been implicated in biological processes, although it is often hard to reveal cause from correlation due to formidable technical issues in analyzing TEs. In this review, we compare and contrast two views of TE activity: one in the pluripotent state, where TEs are broadly beneficial, or at least mechanistically useful, and a second state in human disease, where TEs are uniformly considered harmful.

OPEN ACCESS

Edited by:

Michelle S. Longworth,
Lerner Research Institute,
United States

Reviewed by:

Benoît Chénais,
Le Mans Université, France
Julie Secombe,
Albert Einstein College of Medicine,
United States

*Correspondence:

Andrew P. Hutchins
andrewh@sustech.edu.cn

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 23 March 2022

Accepted: 20 May 2022

Published: 02 June 2022

Citation:

Ma G, Babarinde IA, Zhou X and
Hutchins AP (2022) Transposable
Elements in Pluripotent Stem Cells and
Human Disease.
Front. Genet. 13:902541.
doi: 10.3389/fgene.2022.902541

Keywords: transposable element (TE), endogenous retrovirus (ERV), long terminal repeat (LTR), pluripotent stem cell (PSC), non-coding RNA (ncRNAs)

INTRODUCTION

Transposable elements (TEs) are mobile genetic elements that are found in multiple copies in the genome. TEs were first discovered as mutable loci in 1944, in the study of the corn kernel and leaf color variegation in maize (McClintock, 1950). Barbara McClintock proposed the concepts of genetic loci termed activator and dissociation that could influence gene activity by changing their positions on chromosomes (Guffanti et al., 2014). TEs are usually grouped into two main classes based on transposition mechanism and structural features: the retrotransposons, which transpose by “copy and paste” through an RNA intermediate, and the DNA transposons, which change their positions by a “cut and paste” mechanism (Finnegan, 1989). Retrotransposons are further subdivided into long interspersed elements (LINEs), short interspersed elements (SINEs), and long terminal repeats (LTRs) which are endogenous retroviruses (ERVs). In the human genome, around 40% of the genome is composed of TEs (Hutchins and Pei, 2015), especially retrotransposons. The LINEs are the single most frequent TE, followed by the SINEs, which are a “parasite on a parasite” as they rely on LINE-encoded proteins for their transposition. The LTRs take up third place with DNA transposons and other TEs in last place (Hutchins and Pei, 2015).

The functionality of TEs has always been under some debate. Back in 1972, Sozumu Ohno termed TEs as “junk” DNA in the Brookhaven Symposium in Biology journal (Ohno, 1972). Indeed, there is some sympathy for this view even today, as the vast majority of TEs inside the genome are molecular fossils that have lost their original transposition function, especially in the human genome (Callinan and Batzer, 2006; Hellen and Brookfield, 2013). Meanwhile, even with the development of sophisticated sequencing technologies and genome-wide screens, most TE

sequences remain uncharacterized due to difficulties in unambiguously identifying TEs from sequence reads. Thus, there is a persistent argument that TEs are functionally unimportant and are transcriptional or biological noise. Nevertheless, increasing research has suggested a role for at least some TEs in a wide range of biological processes, including genome evolution, gene organization, expression regulation, and numerous other aspects of cellular biology (Bourque et al., 2018). Consequently, there is an evolving view of TEs as both a potential source of genome innovation, and a potential danger to genome stability, and the development of disease (Payer and Burns, 2019). In this review, we will discuss two areas of TE action, in human pluripotent stem cells (hPSCs) and human disease.

Features and Polymorphism of Transposable Elements

TE abundance varies from several copies up to several thousands of copies of the same element. Although any individual element of the same type is different, due to mutation and truncations, they nonetheless retain some homology. Indeed, even though a TE is inactive due to truncations or mutations, high copy number TEs with similar sequences can provide a sufficient template for recombination and genome rearrangement (Gray, 2000; Bourque et al., 2018). Thus, even as inactive fossils, TEs can still contribute to genome rearrangements. Duplicating TEs in the genome can produce insertion, deletions, chromosomal fusions, and even more complex chromosome rearrangements (Bourque et al., 2018). TE insertions are thus a potential major source of harmful mutations that can cause DNA double-strand breaks, gene dysfunction, gene recombination, gene expression dysregulation, and other types of mutations. TEs are thus a potent source of genetic polymorphisms.

Newly inserted mobile elements could lead to structural variants including deletions, insertions, duplications, and inversions, which may exist as polymorphisms within the population. Strikingly, there are more than 16,000 polymorphic TEs in the human genome, accounting for ~24% of all known structural variants; many of these are common variants, with over 6,500 (36%) having a minor allele frequency >0.01 (Sudmant et al., 2015; Percharde et al., 2018). These insertions are generally located in hotspots with open chromatin. The alleles with novel TE insertions may differ from one another by poly-A tail length or nucleotide substitutions. These polymorphic elements reflect recent sequence insertions with few mutations. Furthermore, inserted TEs usually contain intrinsic functional sequences. Depending on the type of TE, insertional polymorphisms can include autonomous promoters, enhancers, and other regulatory sequences leading to heterochromatin formation, labeled secondary RNA or DNA structures, splicing regulators, and protein-coding sequences (Rishishwar et al., 2015; Bourque et al., 2018; Spirito et al., 2019). Therefore, it is critical to understand TE

polymorphism, which directly affects the genetic diversity and the function of genes in the host genome.

Molecular Functions and Co-Option of Protein-Coding Transposable Element Sequences

Over the past few decades, scientists have mainly focused on several areas in the study of TEs, particularly the annotation, classification, and evolution of TEs, and less emphasis has been paid to their functions. However, recent studies have increasingly shown that TEs play a vital role in a wide range of biological processes (Guffanti et al., 2014; Saleh et al., 2019). TEs can regulate gene transcription (Anwar et al., 2017), RNA subcellular distribution, RNA half-life, transcript abundance and splicing, and DNA methylation (Bourque et al., 2018). Strikingly, TE-encoded proteins have been co-opted for genuine biological function. For example, RAG1 and RAG2, which are essential for mediating sequence-specific DNA recognition in immunoglobulin and T-cell-receptor genes assembly, are derived from a transposase (Agrawal et al., 1998; Zhang et al., 2019) (**Figure 1A**). Syncytin-1 is another co-opted protein, in this case, derived from the viral envelope of a HERV-W (**Figure 1A**). Syncytin-1 has been proposed to have critical roles in normal human placental morphogenesis (Mi et al., 2000). Syncytin-1 is also high in the brains of schizophrenia patients, this correlates with the expression of the inflammation marker CRP (Wang et al., 2018). In addition to these examples, peptides derived from TE sequences have been detected in several cell types (Grow et al., 2015; Li et al., 2015; Babarinde et al., 2021) (**Figure 1B**), although their function (if any) remains unclear.

In addition to the direct co-option of protein-coding TE sequences, TE sequence fragments also play subtle and complex roles in other cellular processes. TE sequences can be transcribed and can influence RNA activity when TE-derived sequence fragments are embedded in RNA sequences (Thornburg et al., 2006) (**Figure 1C**). Indeed, noncoding RNAs are rich in TE-derived sequences (Kelley and Rinn, 2012; Kapusta et al., 2013). Many HERV- and LINE-derived long noncoding transcripts are enriched in stem cells and numerous disease models (Thomas et al., 2017; Babarinde et al., 2021). There is also evidence that the TE sequences inside RNAs influence RNA properties (**Figure 1D**). For example, the presence of TEs in both coding and non-coding transcripts tends to result in lower expression levels, and TEs can lead to retention of the RNA inside the nucleus (Faulkner et al., 2009; Carlevaro-Fita et al., 2019; Babarinde et al., 2021). The mechanism is unclear, but TEs form binding sites that are recognized by RNA binding proteins (**Figure 1E**) (Van Nostrand et al., 2020). On the other hand, epigenomic and transcriptomic studies have revealed that TE sequences contribute a significant fraction of species- and tissue-specific regulatory elements (Kunarso et al., 2010; Trizzino et al., 2017; Pehrsson et al., 2019). TEs often contain a promoter and novel insertions can thus generate new promoter regions of existing genes or

generate new non-coding RNAs (Figure 1F). Indeed, many transcripts start inside TE sequences, or TEs can provide alternative promoters or antisense transcripts (Faulkner et al., 2009). TEs can alter gene expression patterns by contributing cell-specific transcription factor binding sites (Thornburg et al., 2006), which function as enhancers to drive gene expression (Figure 1F), for example, several MER41 elements carry out enhancer functions of interferon response genes (Chuong et al., 2016). Overall, TEs can be utilized by the cell in several ways for legitimate biological functions.

TRANSPOSABLE ELEMENTS IN STEM CELLS

Activity of Fossil Transposable Elements in Pluripotent Stem Cells

TEs have been reported to be expressed in a highly tissue-specific manner (Ecco et al., 2016; Ashapkin et al., 2019; He et al., 2021). Although TEs are expressed in numerous cell and tissue types (Ecco et al., 2016; Ashapkin et al., 2019), they are especially active in early embryonic development (Goke et al., 2015; Barakat et al., 2018; Percharde et al., 2018; Wang et al., 2020). During early

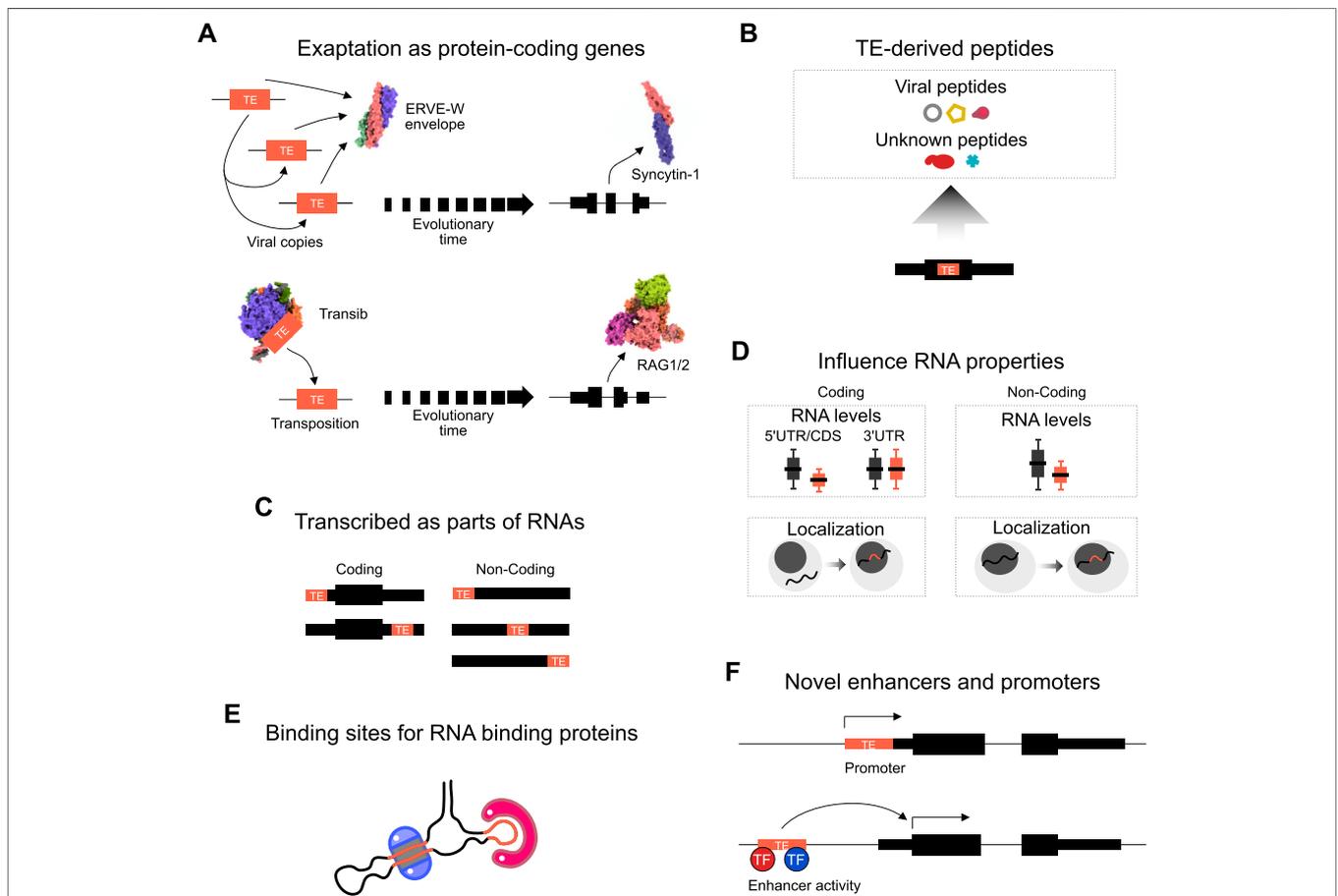


FIGURE 1 | Mechanisms of TE activity. **(A)** Schematic of two examples of the exaptation of transposable element coding sequences as protein-coding genes. TEs randomly duplicate themselves in the genome, and eventually over evolutionary time one TE copy is exapted for biological function. A HERVW envelope protein became Syncytin, and a Transib became RAG1/2 in the immune system. Structures are from 1qbz (SIV gp41) (Yang et al., 1999), 5ha6 (Syncytin-1), 6pr5 (HzTransib) (Liu et al., 2019), 4wwx (RAG1/2) (Kim et al., 2015). **(B)** TEs inserted into coding frames can also give rise to TE-derived peptide fragments. Some match to known TE peptide sequences, for example, LINE ORFs, or ERV gag, pol, env proteins, yet other proteins have no match. **(C)** TEs can be transcribed as parts of coding or non-coding RNAs. In coding transcripts, most TEs are part of the 5' or 3' untranslated regions (UTR). In non-coding transcripts, TEs can be embedded anywhere inside the transcript, and non-coding transcripts often contain multiple different types of TE. **(D)** The presence of TEs inside an RNA sequence impacts properties of the transcript. For example, TEs tend to lead to reduced expression in non-coding transcripts, but only reduce coding RNAs when the TE is present in the 5'UTR or coding sequence (CDS). The presence of a TE in both coding and non-coding transcripts uniformly leads to increased retention of the transcript in the nucleus. **(E)** TEs can form binding sites for the recognition of and binding by RNA binding proteins. RNAs fold up to form hairpins and complex structures which are recognized by sequence or structure-specific RBPs. **(F)** TEs, when inserted into the genome can lead to the evolution of novel enhancers or promoters. TEs contain promoters to promote their expression (and so enhance their ability to colonize the genome). However, TEs can also form novel promoter transcription start sites. TEs also contain transcription factor binding sites that can recruit endogenous transcription factors (TFs) to activate nearby gene expression.

embryonic development, specific types of TE are expressed in a stage-specific manner (Goke et al., 2015; Wang et al., 2020). Interestingly, TE activity in hPSCs coincides with the relaxed genome structure, global hypomethylation, and general genome activation (Zamudio and Bourc'his, 2010; Guo et al., 2014).

To explain the high activity of TEs in early embryo cells and hPSCs, two hypotheses have been proposed. The first suggests that the relatively high activity of the TEs in hPSCs is a consequence of the strategies used by TEs to duplicate themselves across generations (Gerdes et al., 2016). If a TE can duplicate itself in the embryonic or germline cells it will have a higher chance to pass to the next generation, hence TEs often contain embryonic-specific transcription factor binding sites to promote their expression. The second hypothesis takes the opposite approach: that TE activity in hPSCs is an evolutionary innovation of the host cell, and instead of excessively silencing TEs, the host cells intentionally permit the activity of TEs in hPSCs (Brennecke et al., 2008; Sidorenko et al., 2017). The decision of the host cells to permit TE activity might have conferred at least two advantages. First, TE sequence diversity is a potential resource for evolutionary innovation (Heng et al., 2010; Rodriguez-Terrones and Torres-Padilla, 2018). Second, it has been proposed that the host cells transiently relax TE repression during embryogenesis as a way to recognize active TEs and so later activate efficient TE repressive mechanisms during development (Zamudio and Bourc'his, 2010). Indeed, this latter mechanism is employed in the germline cells where TEs are transiently activated during genome reprogramming and are silenced by the PIWI Interacting RNA (piRNA) pathway (Castaneda et al., 2011; Yang and Wang, 2016; Hurst and Magiorkinis, 2017). This class of hypotheses suggests that host cells are the main determinant of TE activities, which makes sense considering the vast majority of TEs are mutated and no longer functional.

Only a handful of TEs (<0.05%) are still capable of transposition in the human genome (Hormozdiari et al., 2011), the majority are molecular fossils. Hence, TE activity doesn't just mark the pluripotent state; they have also been proposed to perform functions in normal embryonic genome activation and development (Mi et al., 2000; Haig, 2016; Izsvak et al., 2016). Some evidence for this comes from the highly stage-specific expression of TEs, for example, SINE Alu is active at the 4-8 cell stage in human (Gerdes et al., 2016), and LINE-1 is highly expressed in hPSCs (Wang et al., 2020). While LTR5-HERVK is mainly activated from the 8-cell stage to blastocyst (Goke et al., 2015). There is also some evidence that the TE sequences themselves are involved in embryogenesis, for example, knocking down HERVHs in hPSCs leads to differentiation (Lu et al., 2014). Nonetheless, there remains a large number of TE sequence fragments expressed in the early embryo and in hPSCs that have no assigned function.

Transposable Elements Are Components of Embryonic Regulatory Networks

TE activity is tightly regulated in biological systems. In somatic cells, the regulation of TE activities is controlled by DNA methylation (Zamudio and Bourc'his, 2010; Guo et al., 2014).

However, in early embryogenesis DNA is demethylated and in naïve hPSCs, the majority of human TEs are hypomethylated (Gkoutela et al., 2015). Instead of repression by DNA methylation, TEs are suppressed by a range of mechanisms, mainly histone modifications, such as Histone 3 lysine 9 trimethylation (H3K9me3), mediated by KRAB-ZNFs through their cofactor TRIM28 which recruits the H3K9me3 methyltransferase SETDB1 (Castro-Diaz et al., 2014; Turelli et al., 2014; Pontis et al., 2019). Specifically, depletion of TRIM28 led to the removal of the repressive chromatin marks, thereby activating ERVs (Rowe et al., 2013; Turelli et al., 2014). Indeed, ZNF93 and TRIM28 mediate the deposition of H3K9me3 on a subset of LINE-1 elements to repress their transcription (Castro-Diaz et al., 2014), while the transcription factor YY1 is involved in indirectly silencing younger LINE-1 subfamilies by mediating DNA methylation in hPSCs (Sanchez-Luque et al., 2019). However, the regulation of TEs is more complex than just repression, as TEs in mouse PSCs harbor a wide range of histone patterns, including histone marks for active transcription and enhancers (He et al., 2019; Zhou et al., 2021). Taken together, TEs are regulated by a large array of biological processes such as DNA methylation, histone modification, RNA degradation, and translational control.

TEs play important roles in the regulatory networks of ESC (Barakat et al., 2018; Wang et al., 2020; Bakoulis et al., 2022). They contain binding sites of key hPSC-specific regulatory factors, such as POU5F1, SOX2, and NANOG (Wang et al., 2020). For example, more than 70% of the SOX2 binding sites in hPSCs overlapped a TE (**Figure 2A**). The relationship between hPSC-specific factors and TEs is also mechanistic, as disruption of naïve pluripotent stem cells transcription factor LBP9 downregulated the transcription of HERVH-derived transcripts and compromised the self-renewal of hPSCs (Wang et al., 2014). Importantly, TEs have been incorporated into normal regulatory networks to enable species-specific innovations (Heng et al., 2010; Rodriguez-Terrones and Torres-Padilla, 2018). For instance, a comparison of the genome-wide binding locations of two pluripotency transcription factors POU5F1 and NANOG in human and mouse PSCs found that the binding locations of POU5F1 and NANOG were highly divergent between the two species, and around a quarter of the transcription factor binding sites were contained inside TE sequences (Kunarsow et al., 2010). Ultimately, the study suggested that the activity of TEs moves transcription factor binding sites throughout the genome over evolutionary time. In this way, transcription factors can regulate the same genes in the same cell type, but the binding sites themselves need not be well conserved.

The transcription of TEs in human pluripotent stem cells (hPSCs) has been extensively investigated. Several studies have reported various proportions of TEs in human transcripts, it is widely agreed that TEs contribute more to lncRNAs than protein-coding transcripts, suggesting they are a major component of lncRNAs (Babarinde et al., 2021). For example, 64% of lncRNAs in hPSCs contain a TE-derived sequence, compared to 27% of coding transcripts

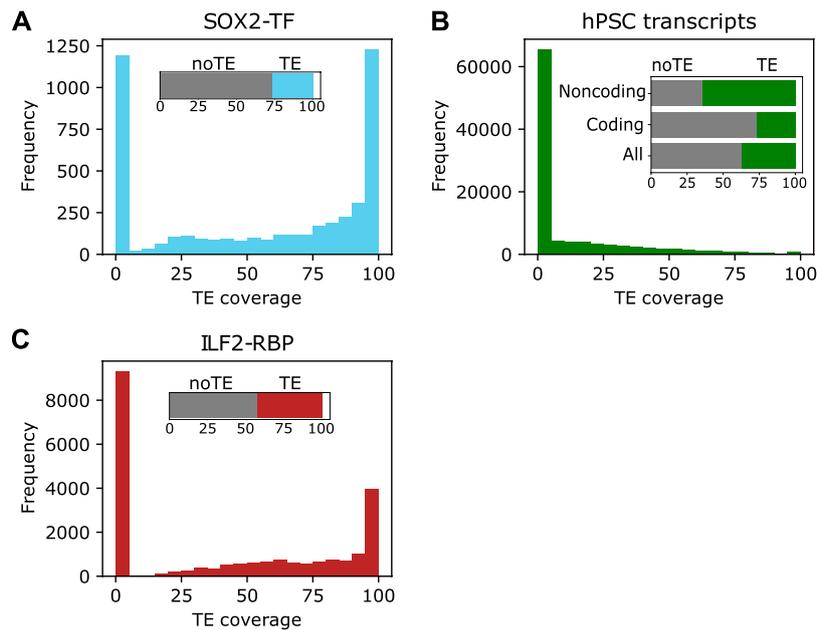


FIGURE 2 | TEs are incorporated into different elements in stem cells. **(A)** The distribution of TE coverage in SOX2 binding sites in human naïve PSCs. The data were retrieved from (Huang et al., 2021). **(B)** The distribution of TEs in the hPSC transcriptome assembly. The insert shows the distribution of TE-sequences in all, coding and noncoding transcripts. **(C)** The distribution of TEs in ILF2 RNA binding sites. The insert is the total classification based on any TE overlap.

(Babarinde et al., 2021). Interestingly, TE-sequence-containing transcripts also tended to contain varying proportions of unique sequences (Figure 2B), demonstrating that TEs have been incorporated into different hPSC transcripts to various degrees (Babarinde et al., 2021). Additionally, TE-containing hPSC transcripts tend to be less stable, more localized to the nucleus, and with less coding ability (Babarinde et al., 2021). One well-studied example of a TE-derived lncRNA in human ESCs is HERVH transcripts. Functionally, the disruption of HERVH and HERVH-derived transcripts negatively affects the self-renewal of hPSCs by reducing the recruitment of pluripotent transcription factors to particular binding sites (Lu et al., 2014; Wang et al., 2014). These studies highlight the roles of TE-derived transcripts in hPSCs.

Transposable Elements Are Targets of RNA-Binding Proteins in Human Pluripotent Stem Cells

Transcript stability and subcellular localizations are largely controlled by RNA binding proteins (Kelley et al., 2014). Mechanistically, TEs may lead to changes in RNA activities through binding to RNA binding proteins (RBPs). By mapping the RNA binding sites of 51 human proteins, Kelley et al. (2014) found that the RNA binding proteins (RBPs) were differentially bound to specific TEs. A large-scale genome-wide study also found RBPs bound to antisense sequences of LINEs and Alu SINE TEs (Van Nostrand et al., 2020). This suggests a possible mechanism for TEs to alter RNA properties,

such as transcript abundance and splicing. Indeed, the TEs themselves may form binding sites for RBPs, for example in how STAUFEN binds to Alu SINE sequences (Gong and Maquat, 2011).

A similar pattern is seen in hPSCs, and TE-sequences in RNAs are bound by RBPs. Analysis of the RBPs DDX6, ILF2, FUS, and DCP1B in hPSCs revealed that TE-containing transcripts have unique RBP interaction (Babarinde et al., 2021). For example, 58% of the RNA-binding sites for ILF2 identified are derived from TEs (Figure 2C). Functionally, DDX6 which has widespread binding sites with no TE preference in hPSC transcripts (Babarinde et al., 2021) is associated with cell plasticity and parental RNA decay for cellular reprogramming to pluripotency (Kami et al., 2018; Di Stefano et al., 2019). Despite limited studies reporting RBP interactions with TEs or TE-derived transcripts in hPSCs, the analyses of RBPs in other cell types (Kelley et al., 2014; Attig et al., 2018; Van Nostrand et al., 2020) as well as the abundance and properties of TEs incorporated into transcripts, point to the importance of the interplay between RBPs and TEs. TEs inside RNAs, and especially in lncRNAs, may form binding platforms or regulatory domains for RBPs to bind to and execute biological functions (Johnson and Guigo, 2014).

TRANSPOSABLE ELEMENTS IN DISEASE

The involvement of TEs in harmful mutations, gene dysfunction, DNA double-strand breaks, gene recombination, gene expression dysregulation, and other

TABLE 1 | Transposable elements implicated in human disease.

Condition type	TEs	Mechanism	References
Cancer			
Colon cancer	LINE-1	LINE-1 promoter hypomethylation	(Ogino et al., 2008; Cruickshanks and Tufarelli, 2009; Weber et al., 2010)
Esophageal squamous cell carcinoma	LINE-1	LINE-1 promoter hypomethylation	Iwagami et al. (2013)
Breast cancer	LINE-1	LINE-1 promoter hypomethylation	(Cruickshanks and Tufarelli, 2009; van Hoesel et al., 2012)
Hepatocellular carcinomas	LINE-1	LINE-1 promoter hypomethylation	Harada et al. (2015)
Ovarian cancer	LINE-1	LINE-1 promoter hypomethylation	Pattamadilok et al. (2008)
Chronic myeloid leukemia	LINE-1	LINE-1 promoter hypomethylation	(Roman-Gomez et al., 2005; Weber et al., 2010)
Bladder tumors	LINE-1	LINE-1 promoter hypomethylation	Wolff et al. (2010)
Colorectal cancer	LINE-1	LINE-1 promoter hypomethylation	Hur et al. (2014)
Colon cancers	LINE-1	Forms dsRNA and suppresses <i>TFPI2</i>	Cruickshanks et al. (2013)
Colon cancer	LINE-1 ORF1p	LINE-1 ORF1p overexpression	Rodic et al. (2014)
Ovarian cancers	LINE-1 ORF1p	LINE-1 ORF1p overexpression	Rodic et al. (2014)
Lung cancers	LINE-1 ORF1p	LINE-1 ORF1p overexpression	Ardejan et al. (2017)
Colon cancer	LINE-1	LINE-1 insertion into tumor suppressor APC	(Miki et al., 1992; Scott et al., 2016)
Colorectal cancer	LINE-1	Insertion causes gene mutation	(Lee et al., 2012; Solyom et al., 2012)
NSCLCs	LINE-1	Insertion causes gene mutation	(Iskov et al., 2010; Tubio et al., 2014)
Head and neck cancers	LINE-1	Insertion causes gene mutation	(Helman et al., 2014; Tubio et al., 2014)
Ovarian cancers	LINE-1	Insertion causes gene mutation	(Lee et al., 2012; Tubio et al., 2014; Tang et al., 2017)
Gastric cancer	LINE-1	LINE-1 hypomethylation	Shigaki et al. (2013)
Ovarian cancer	HERV-K	Increased expression	Rycay et al. (2015)
Melanoma	HERV-K	Increased expression	(Schivavetti et al., 2002; Singh et al., 2020)
Pancreatic cancer	HERV-K	Increased expression	Li et al. (2017)
Psychiatric disorders, neurofibromatosis, Alzheimer's disease			
Multiple sclerosis	HERV-W	LINE-1 expression	Perron et al. (1993)
Aicardi-goutières syndrome	LINE-1	Re-activates LINE-1	Thomas et al. (2017)
Rett syndrome	LINE-1 ORF2	MECP2 loss of function increases susceptibility to LINE-1 insertions	Muotri et al. (2010)
Systemic lupus erythematosus	HERV	Increased expression correlates with SLE	Wu et al. (2015)
Sporadic amyotrophic lateral sclerosis	HERV-K	Increased expression correlates with SALS	Li et al. (2015)
Autism spectrum disorders	LINE-1	An increase in LINE-1 expression correlates with autism	(Shpyleva et al., 2018; Tangsuwansri et al., 2018)
Amyotrophic lateral sclerosis	HERV-K	Aberrant expression	Curzio et al. (2020)
Multiple sclerosis	HERV-W	Increased expression	(Perron et al., 1997; Antony et al., 2007)
Immune system			
Fibromyalgia	HERVs	Increased expression correlates with fibromyalgia	Ovejero et al. (2020)
Autoimmunity	HERV envelope	Expression triggers both innate and adaptive immunity	Grandi and Tramontano, (2018)
Aging			
Age-associated inflammation	LINE-1	Derepresses LINE-1 and activates a type I interferon (IFN-I) response	De Cecco et al. (2019)
Senescence	Alu, SVA, and L1	More accessible for Alu, SVA, and L1 transcription	De Cecco et al. (2013a)
Aging	LINE-1	SIRT6 fails to repress LINE-1 activity	Van Meter et al. (2014)

types of mutations implies that TEs might contribute to human disease. There is growing evidence of a link between TE sequence fragments, PSCs, and cancer. TEs can act as oncogene-specific enhancers, promoters, and exons for pluripotency-specific genes, this drives their expression and converts them into oncogenes (Jang et al., 2019). This builds on the observations that pluripotent genes are oncogenic (Leis et al., 2012; Muller et al., 2016), and PSCs themselves have tumorigenic potential. However, the TE-gene fusions seen in

cancer were not observed in normal hPSCs (Zapatka et al., 2020; Babarinde et al., 2021). This observation suggests that some TEs are either normal or tolerated by the cell, whilst other TEs are specifically associated with disease.

TE transposition activity has been suggested to contribute to human genetic diseases, primarily through the transposition of LINE-1, SINE Alu, and SVA (SINE-VNTR-Alu) TEs (Cordaux and Batzer, 2009). Strikingly, nearly 100 retrotransposition events caused by polymorphic LINE-1 have been implicated in human

diseases (Hancks and Kazazian, 2012). Non-functional mutated TE sequences expressed inside RNAs have also been implicated in several conditions, including psychiatric disorders (Guffanti et al., 2014), neurofibromatosis (Payer and Burns, 2019), cancer (Chenais, 2013; Rodriguez-Martin et al., 2016) and aging (De Cecco et al., 2013b). However, until recently, there were surprisingly few concrete examples of the impact of TEs on human disease (Nakamura et al., 2015; Rodriguez-Martin et al., 2016), and the association with disease often remains correlative. Much of the challenge in studying TE-associated human diseases stems from the difficulty in accurately sequencing the locations of TEs inside an individual's genome, coupled with difficulties in establishing causation between novel insertions and disease. **Table 1** shows example TEs that have been implicated in selected diseases.

Transposable Elements in Cancer

LINE-1 elements are actively mobilized in cancer; however, untangling if this is a cause or consequence of tumorigenesis has been challenging. The first identified example of a LINE-1 disrupting a tumor suppressor gene was recognized in 1992 in a patient with colorectal cancer (Miki et al., 1992). It was found that a LINE-1 was inserted into the *APC* tumor suppressor gene. The insertion included the 3' part of a LINE-1 and around 180 base pairs of polyadenylated sequence (Miki et al., 1992). This was the first report of the disruption of a tumor suppressor gene caused by the somatic insertion of a mobile genetic element. Another independent study also found a novel somatic LINE-1 insertion in colorectal cancer which disrupted the *APC* gene (Scott et al., 2016). Moreover, using LINE-1-targeted sequencing in 16 colorectal tumors, Szilvia confirmed tumorigenesis-related genes were mutagenized by specific *de novo* LINE-1 insertions (Solyom et al., 2012). The study validated 69/107 tumor-specific insertions including 35 instances in which both 5' and 3' junctions were retrieved. This agrees with the study of cancer genomes from 244 patients, in which 53% of the patients had somatic retrotranspositions, including 24% 3' transductions (Tubio et al., 2014). By utilizing single-nucleotide resolution analysis, Lee et al. (2012) reported LINE-1 in genes that are commonly mutated in cancer. Identifying somatic TE insertions is experimentally challenging and a series of "Transposon-seq" methods have been developed that attempt to accurately place somatic TE insertions (Iskow et al., 2010). Employing a Transposon-seq method to interrogate 767 tumor samples with hybrid-capture exome data discovered 35 novel somatic transpositions, including one in the *PTEN* tumor suppressor gene (Helman et al., 2014). Also, LINE-1 ORF1p overexpression is observed in many human tumors (Burns, 2017). Indeed, LINE-1 expression has been positively identified as a biomarker in several cancers (**Table 1**).

Deregulation of TEs is a hallmark of many kinds of cancer (Jang et al., 2019; Zapatka et al., 2020). Genome-wide DNA hypomethylation is harmful to genomic stability and is often observed in cancer, particularly colorectal carcinogenesis (Jones and Gonzalvo, 1997; Lengauer et al., 1997; Breivik and Gaudernack, 1999; Ogino et al., 2008). In 2010, Wolff et al. (2010) for the first time demonstrated that the LINE-1 promoter

hypomethylation directly causes dysregulation of endogenous gene expression. They reported that hypomethylation of the LINE-1 promoter activated an alternate transcript of the *MET* oncogene in bladder tumors. Recent studies have implicated LINE-1 promoter hypomethylation in various cancers by enabling full-length LINE-1 mRNA translation from the LINE-1 mobilization machinery (Shukla et al., 2013; Burns, 2017; Zhang et al., 2020). Indeed, LINE-1 promoter hypomethylation, LINE-1 ORF1p protein expression, and somatic polymorphic LINE-1 retrotransposition have been linked to lung, colon, pancreatic, ovarian, and breast cancers (Lee et al., 2012; van Hoesel et al., 2012; Rodic et al., 2014; Ardeljan et al., 2017). Further, LINE-1 promoter hypomethylation leading to sense/antisense transcription is a marker for the progression of chronic myeloid leukemia (Roman-Gomez et al., 2005). The LINE-1 antisense promoter can also affect adjacent genomic states by generating chimeric RNAs which can interrupt transcription. Novel LINE-1 chimeric transcripts have been observed in breast cancer cell lines and colon cancer cells (Cruickshanks and Tufarelli, 2009). These studies demonstrate that hypomethylation of LINE-1s plays a role not only in human diseases but also in disease predisposition (Wolff et al., 2010). Consistently, LINE-1 hypomethylation caused by the inhibition of DNMTs in colon carcinoma cells or myeloid leukemia cells induces the expression of an irregular fusion transcript between an intronic LINE-1 element and the proto-oncogene *MET* (c-Met) (Weber et al., 2010). In addition, an analysis of 77 colorectal cancer patients showed that hypomethylation of LINE-1 led to the activation of proto-oncogenes in human colorectal cancer metastasis (Hur et al., 2014). Also, genome-wide DNA hypomethylation levels revealed by LINE-1 hypomethylation demonstrated that LINE-1 methylation levels can be used as a biomarker for identifying hepatocellular carcinoma patients who will experience poor clinical outcomes (Pattamadilok et al., 2008; Harada et al., 2015). In addition, LINE-1 hypomethylation has been reported to be associated with poor survival in more than 200 cases of gastric cancer and esophageal squamous cell carcinoma (Iwagami et al., 2013; Shigaki et al., 2013), suggesting its potential as a prognostic biomarker. Further, the study applying pyrosequencing in two independent cohorts of 643 colon cancer patients found that LINE-1 hypomethylation is associated with shorter survival (Ogino et al., 2008). Taken together, LINE-1 hypomethylation leading to aberrant transcription is associated with various cancer types.

In addition to LINE-1, the expression of ERVs has also been reported in cancer (Zapatka et al., 2020), including translated peptides and fragments of viral proteins. Transcripts derived from ERVs have been observed in many cancer types, including particularly ovarian cancer (Rycaj et al., 2015), melanoma (Schiavetti et al., 2002), pancreatic cancer (Li et al., 2017), breast cancer (Wang-Johanning et al., 2008), and prostate adenocarcinoma (Wang-Johanning et al., 2003). However, the mechanistic roles of the ERV-containing transcripts and peptides has not been well explored. HERV-K is one of the most well-studied subfamilies of ERVs. For instance, downregulation of HERV-K decreased cell proliferation and tumor growth in pancreatic cancer (Li et al.,

2017), suggesting a causal link between ERVs and tumor growth. Furthermore, the experimental depletion of HERV-K Rec in melanoma led to a lower level of melanocyte-inducing transcription factor (MITF), which may impact the transition from proliferative to invasive stages of melanoma (Singh et al., 2020). Targeting ERV sequences for downregulation has also been shown to have the potential in decreasing cancer proliferation in pancreatic cancers and melanoma (Li et al., 2017). These studies suggest that ERVs are not just markers for cancers, but they may also directly contribute to cancer progression, although the mechanisms behind these processes are not clear.

Transposable Elements in Inflammation and Neuroderegulation

TEs have also been shown to be a key player in immune regulation, such as cancer immune and autoimmune diseases (Ovejero et al., 2020). For example, LINE-1 can manipulate the immune system and immune microenvironment in many ways (Zhang et al., 2020). The analysis of the transcriptional activity of 1789 pathways in 112 TCGA cancer samples revealed that 49 of 176 immune pathways were significantly negatively correlated with LINE-1 expression (Zhang et al., 2020). Jung et al. (2018) reported that cancer immunity may contribute to genome stability by suppressing LINE-1 retrotransposition in gastrointestinal cancers. Furthermore, the LINE-1 methylation level was significantly associated with the peritumoral lymphocytic reaction in esophageal cancer (Zhang et al., 2020). Besides the involvement in immune alteration of tumors, LINE-1 also plays an important role in other inflammation-related diseases. For example, intragenic LINE-1s can also act as *cis*-regulatory elements to mediate the activation of the autoimmune gene expression in Fanconi anemia and Aicardi-Goutières syndrome (Wanichnopparat et al., 2013). Also, accumulation of LINE-1 in *TREX1*-deficient type I interferon apoptosis leads to autoimmune neuroinflammation disease (Thomas et al., 2017). In human neurodevelopmental diseases model, *MECP2* mutations can influence the frequency of LINE-1 retrotransposition (Muotri et al., 2010). Moreover, LINE-1 expression was significantly elevated in Autism (Shpyleva et al., 2018). Apart from LINE-1, ERVs have also been reported capable of initiating an immune response in disease models. One example is Syncytin-1, which is normally expressed in the placenta but has been reported to be upregulated in multiple sclerosis (MS) lesions (Antony et al., 2011). The abnormal expression was proposed to lead to proinflammatory cytokine release, oxidative damage, and eventually oligodendrocyte death. Another study has shown that HERVs of the H, K, and W subfamilies are overexpressed in the immune cells of fibromyalgia patients (Ovejero et al., 2020). Interestingly, patients infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), responsible for the COVID-19 pandemic, showed transcriptional upregulation of HERV-17 in T lymphocytes and leukocytes, which correlated with disease severity (Balestrieri et al., 2021; Garcia-Montojo & Nath, 2021).

Multiple studies have also implicated TEs in nervous system diseases such as neuropathy, psychiatric disorders, neurofibromatosis, and neurodegeneration. There is a particular interest in TE-derived peptides that induce an inflammatory

response. For HERVK/Ws both their sequences and protein products have been implicated in the development of amyotrophic lateral sclerosis (ALS) and other neurological diseases (Arru et al., 2018). In fact, the inhibition of HERV-K expression with anti-viral therapy decelerated ALS progression (Arru et al., 2018). Interestingly, one study identified a conotoxin-like protein that was produced from an ERVK sequence as a potential factor in ALS neuropathology (Curzio et al., 2020). The conotoxin-like protein could induce NF- κ B pro-inflammatory signaling, and its presence was linked with regions of damage in ALS samples (Curzio et al., 2020). Other studies have also documented various ways in which ERVs potentially contribute to multiple sclerosis (Perron et al., 1997; Antony et al., 2007). Moreover, the inhibition of LTR methylation may activate HERV-E transcription in systemic lupus erythematosus (Li et al., 2015; Wu et al., 2015), and elevated HERV-K expression in patients may contribute to neurodegeneration (Li et al., 2015). Indeed, the expression and DNA methylation patterns of ERV and LINE-1 are often disrupted in patients with autism, schizophrenia, and mood disorders (Misiak et al., 2019). These studies highlight various examples of the association between TEs and inflammation and neurodegeneration.

CONCLUSION

TEs are active in different cell types and developmental stages, they are hyperactive in hPSCs where they function in transcriptional regulation, transcript processing such as splicing, RNA stability, and translational processes. Intriguingly, TEs are broadly positive in early embryogenesis, contributing to gene regulation pathways, and acting as a substrate for evolutionary innovation. This is in contrast to the role of TEs in somatic tissues, which tend to be more negative, being associated with the development of human diseases. TEs and their derived peptides or sequence fragments have complex roles in the cell. However, many of these roles remain unclear. Considering the vast number of TEs in the human genome it has, and remains, challenging to study them. However, there is likely much critical information that remains to be discovered concerning both the advantageous and deleterious functions of TEs in both embryogenesis and somatic cells.

AUTHOR CONTRIBUTIONS

GM drafted the first version of the manuscript and revised the text. IAB and XZ revised the manuscript. APH revised the manuscript, approved the final text, and funded the work.

FUNDING

This work was supported by the National Natural Science Foundation of China (31970589), the Shenzhen Innovation Committee of Science and Technology (JCYJ20200109141018712 and ZDSYS20200811144002008 to the Shenzhen Key Laboratory of Gene Regulation and Systems Biology).

REFERENCES

- Agrawal, A., Eastman, Q. M., and Schatz, D. G. (1998). Transposition Mediated by RAG1 and RAG2 and its Implications for the Evolution of the Immune System. *Nature* 394(6695), 744–751. doi:10.1038/29457
- Antony, J. M., Deslauriers, A. M., Bhat, R. K., Ellestad, K. K., and Power, C. (2011). Human Endogenous Retroviruses and Multiple Sclerosis: Innocent Bystanders or Disease Determinants? *Biochimica Biophysica Acta (BBA) - Mol. Basis Dis.* 1812 (2), 162–176. doi:10.1016/j.bbadis.2010.07.016
- Antony, J. M., Zhu, Y., Izad, M., Warren, K. G., Vodjgani, M., and Mallet, F. (2007). Comparative Expression of Human Endogenous Retrovirus-W Genes in Multiple Sclerosis. *AIDS Res. Hum. Retroviruses* 23 (10), 1251–1256. doi:10.1089/aid.2006.0274
- Anwar, S. L., Wulaningsih, W., and Lehmann, U. (2017). Transposable Elements in Human Cancer: Causes and Consequences of Dereglulation. *Int. J. Mol. Sci.* 18 (5). doi:10.3390/ijms18050974
- Ardeljan, D., Taylor, M. S., Ting, D. T., and Burns, K. H. (2017). The Human Long Interspersed Element-1 Retrotransposon: An Emerging Biomarker of Neoplasia. *Clin. Chem.* 63 (4), 816–822. doi:10.1373/clinchem.2016.257444
- Arru, G., Mameli, G., Deiana, G. A., Rattu, A. L., Piredda, R., Sechi, E., et al. (2018). Humoral Immunity Response to Human Endogenous Retroviruses K/W Differentiates between Amyotrophic Lateral Sclerosis and Other Neurological Diseases. *Eur. J. Neurol.* 25 (8), 1076–e1084. doi:10.1111/ene.13648
- Ashapkin, V. V., Kutueva, L. I., and Vanyushin, B. F. (2019). Epigenetic Clock: Just a Convenient Marker or an Active Driver of Aging? *Adv. Exp. Med. Biol.* 1178, 175–206. doi:10.1007/978-3-030-25650-0_10
- Attig, J., Agostini, F., Gooding, C., Chakrabarti, A. M., Singh, A., Haberman, N., et al. (2018). Heteromeric RNP Assembly at LINEs Controls Lineage-specific RNA Processing. *Cell* 174 (5), 1067–1081. doi:10.1016/j.cell.2018.07.001
- Babarinde, I. A., Ma, G., Li, Y., Deng, B., Luo, Z., Liu, H., et al. (2021). Transposable Element Sequence Fragments Incorporated into Coding and Noncoding Transcripts Modulate the Transcriptome of Human Pluripotent Stem Cells. *Nucleic Acids Res.* 49 (16), 9132–9153. doi:10.1093/nar/gkab710
- Bakoulis, S., Krautz, R., Alcaraz, N., Salvatore, M., and Andersson, R. (2022). Endogenous Retroviruses Co-opted as Divergently Transcribed Regulatory Elements Shape the Regulatory Landscape of Embryonic Stem Cells. *Nucleic Acids Res.* 50 (4), 2111–2127. doi:10.1093/nar/gkac088
- Barakat, T. S., Halbritter, F., Zhang, M., Rendeiro, A. F., Perenthaler, E., Bock, C., et al. (2018). Functional Dissection of the Enhancer Repertoire in Human Embryonic Stem Cells. *Cell Stem Cell* 23 (2), 276–288 e278. doi:10.1016/j.stem.2018.06.014
- Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., et al. (2018). Ten Things You Should Know about Transposable Elements. *Genome Biol.* 19 (1), 199. doi:10.1186/s13059-018-1577-z
- Brevik, J., and Gaudernack, G. (1999). Genomic Instability, DNA Methylation, and Natural Selection in Colorectal Carcinogenesis. *Semin. Cancer Biol.* 9 (4), 245–254. doi:10.1006/scbi.1999.0123
- Brennecke, J., Malone, C. D., Aravin, A. A., Sachidanandam, R., Stark, A., and Hannon, G. J. (2008). An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. *Science* 322 (5906), 1387–1392. doi:10.1126/science.1165171
- Burns, K. H. (2017). Transposable Elements in Cancer. *Nat. Rev. Cancer* 17 (7), 415–424. doi:10.1038/nrc.2017.35
- Callinan, P. A., and Batzer, M. A. (2006). Retrotransposable Elements and Human Disease. *Genome Dyn.* 1, 104–115. doi:10.1159/000092503
- Carlevaro-Fita, J., Polidori, T., Das, M., Navarro, C., Zoller, T. I., and Johnson, R. (2019). Ancient Exapted Transposable Elements Promote Nuclear Enrichment of Human Long Noncoding RNAs. *Genome Res.* 29 (2), 208–222. doi:10.1101/gr.229922.117
- Castaneda, J., Genzor, P., and Bortvin, A. (2011). piRNAs, Transposon Silencing, and Germline Genome Integrity. *Mutat. Res.* 714 (1–2), 95–104. doi:10.1016/j.mrfmmm.2011.05.002
- Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., et al. (2014). Evolutionally Dynamic L1 Regulation in Embryonic Stem Cells. *Genes Dev.* 28 (13), 1397–1409. doi:10.1101/gad.241661.114
- Chenais, B. (2013). Transposable Elements and Human Cancer: a Causal Relationship? *Biochim. Biophys. Acta* 1835 (1), 28–35. doi:10.1016/j.bbcan.2012.09.001
- Chuong, E. B., Elde, N. C., and Feschotte, C. (2016). Regulatory Evolution of Innate Immunity through Co-option of Endogenous Retroviruses. *Science* 351 (6277), 1083–1087. doi:10.1126/science.1245497
- Cordaux, R., and Batzer, M. A. (2009). The Impact of Retrotransposons on Human Genome Evolution. *Nat. Rev. Genet.* 10 (10), 691–703. doi:10.1038/nrg2640
- Cruikshanks, H. A., and Tufarelli, C. (2009). Isolation of Cancer-specific Chimeric Transcripts Induced by Hypomethylation of the LINE-1 Antisense Promoter. *Genomics* 94 (6), 397–406. doi:10.1016/j.ygeno.2009.08.013
- Cruikshanks, H. A., Vafadar-Isfahani, N., Dunican, D. S., Lee, A., Sproul, D., Lund, J. N., et al. (2013). Expression of a Large LINE-1-Driven Antisense RNA Is Linked to Epigenetic Silencing of the Metastasis Suppressor Gene TFPI-2 in Cancer. *Nucleic Acids Res.* 41 (14), 6857–6869. doi:10.1093/nar/gkt438
- Curzio, D. D., Gurm, M., Turnbull, M., Nadeau, M. J., Meek, B., Rempel, J. D., et al. (2020). Pro-Inflammatory Signaling Upregulates a Neurotoxic Conotoxin-like Protein Encrypted within Human Endogenous Retrovirus-K. *Cells* 9 (7). doi:10.3390/cells9071584
- De Cecco, M., Criscione, S. W., Peckham, E. J., Hillenmeyer, S., Hamm, E. A., Manivannan, J., et al. (2013a). Genomes of Replicatively Senescent Cells Undergo Global Epigenetic Changes Leading to Gene Silencing and Activation of Transposable Elements. *Aging Cell* 12 (2), 247–256. doi:10.1111/ace.12047
- De Cecco, M., Criscione, S. W., Peterson, A. L., Neretti, N., Sedivy, J. M., and Kreiling, J. A. (2013b). Transposable Elements Become Active and Mobile in the Genomes of Aging Mammalian Somatic Tissues. *Aging (Albany NY)* 5 (12), 867–883. doi:10.18632/aging.100621
- De Cecco, M., Ito, T., Petrashen, A. P., Elias, A. E., Skvir, N. J., Criscione, S. W., et al. (2019). L1 Drives IFN in Senescent Cells and Promotes Age-Associated Inflammation. *Nature* 566 (7742), 73–78. doi:10.1038/s41586-018-0784-9
- Di Stefano, B., Luo, E.-C., Haggerty, C., Aigner, S., Charlton, J., Brumbaugh, J., et al. (2019). The RNA Helicase DDX6 Controls Cellular Plasticity by Modulating P-Body Homeostasis. *Cell Stem Cell* 25 (5), 622–638. doi:10.1016/j.stem.2019.08.018
- Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., et al. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Dev. Cell* 36 (6), 611–623. doi:10.1016/j.devcel.2016.02.024
- Faulkner, G. J., Kimura, Y., Daub, C. O., Wani, S., Plessy, C., Irvine, K. M., et al. (2009). The Regulated Retrotransposon Transcriptome of Mammalian Cells. *Nat. Genet.* 41 (5), 563–571. doi:10.1038/ng.368
- Finnegan, D. J. (1989). Eukaryotic Transposable Elements and Genome Evolution. *Trends Genet.* 5 (4), 103–107. doi:10.1016/0168-9525(89)90039-5
- Gerdes, P., Richardson, S. R., Mager, D. L., and Faulkner, G. J. (2016). Transposable Elements in the Mammalian Embryo: Pioneers Surviving through Stealth and Service. *Genome Biol.* 17 (1), 100. doi:10.1186/s13059-016-0965-5
- Gkoutoula, S., Zhang, K. X., Shafiq, T. A., Liao, W. W., Hargan-Calvopina, J., Chen, P. Y., et al. (2015). DNA Demethylation Dynamics in the Human Prenatal Germline. *Cell* 161 (6), 1425–1436. doi:10.1016/j.cell.2015.05.012
- Goke, J., Lu, X., Chan, Y. S., Ng, H. H., Ly, L. H., Sachs, F., et al. (2015). Dynamic Transcription of Distinct Classes of Endogenous Retroviral Elements Marks Specific Populations of Early Human Embryonic Cells. *Cell Stem Cell* 16 (2), 135–141. doi:10.1016/j.stem.2015.01.005
- Gong, C., and Maquat, L. E. (2011). lncRNAs Transactivate STAU1-Mediated mRNA Decay by Duplexing with 3' UTRs via Alu Elements. *Nature* 470 (7333), 284–288. doi:10.1038/nature09701
- Grandi, N., and Tramontano, E. (2018). HERV Envelope Proteins: Physiological Role and Pathogenic Potential in Cancer and Autoimmunity. *Front. Microbiol.* 9, 462. doi:10.3389/fmicb.2018.00462
- Gray, Y. H. (2000). It Takes Two Transposons to Tango: Transposable-Element-Mediated Chromosomal Rearrangements. *Trends Genet.* 16 (10), 461–468. doi:10.1016/s0168-9525(00)02104-1
- Grow, E. J., Flynn, R. A., Chavez, S. L., Bayless, N. L., Wossidlo, M., Wesche, D. J., et al. (2015). Intrinsic Retroviral Reactivation in Human Preimplantation Embryos and Pluripotent Cells. *Nature* 522 (7555), 221–225. doi:10.1038/nature14308

- Guffanti, G., Gaudi, S., Fallon, J. H., Sobell, J., Potkin, S. G., Pato, C., et al. (2014). Transposable Elements and Psychiatric Disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 165B (3), 201–216. doi:10.1002/ajmg.b.32225
- Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., et al. (2014). The DNA Methylation Landscape of Human Early Embryos. *Nature* 511 (7511), 606–610. doi:10.1038/nature13544
- Haig, D. (2016). Transposable Elements: Self-Seekers of the Germline, Team-Players of the Soma. *Bioessays* 38 (11), 1158–1166. doi:10.1002/bies.201600125
- Hancks, D. C., and Kazazian, H. H., Jr. (2012). Active Human Retrotransposons: Variation and Disease. *Curr. Opin. Genet. Dev.* 22 (3), 191–203. doi:10.1016/j.gde.2012.02.006
- Harada, K., Baba, Y., Ishimoto, T., Chikamoto, A., Kosumi, K., Hayashi, H., et al. (2015). LINE-1 Methylation Level and Patient Prognosis in a Database of 208 Hepatocellular Carcinomas. *Ann. Surg. Oncol.* 22 (4), 1280–1287. doi:10.1245/s10434-014-4134-3
- He, J., Babarinde, I. A., Sun, L., Xu, S., Chen, R., Shi, J., et al. (2021). Identifying Transposable Element Expression Dynamics and Heterogeneity during Development at the Single-Cell Level with a Processing Pipeline scTE. *Nat. Commun.* 12 (1), 1456. doi:10.1038/s41467-021-21808-x
- He, J., Fu, X., Zhang, M., He, F., Li, W., Abdul, M. M., et al. (2019). Transposable Elements Are Regulated by Context-specific Patterns of Chromatin Marks in Mouse Embryonic Stem Cells. *Nat. Commun.* 10 (1), 34. doi:10.1038/s41467-018-08006-y
- Hellen, E. H., and Brookfield, J. F. (2013). Transposable Element Invasions. *Mob. Genet. Elem.* 3 (1), e23920. doi:10.4161/mge.23920
- Helman, E., Lawrence, M. S., Stewart, C., Sougnez, C., Getz, G., and Meyerson, M. (2014). Somatic Retrotransposition in Human Cancer Revealed by Whole-Genome and Exome Sequencing. *Genome Res.* 24 (7), 1053–1063. doi:10.1101/gr.163659.113
- Heng, J. C., Orlov, Y. L., and Ng, H. H. (2010). Transcription Factors for the Modulation of Pluripotency and Reprogramming. *Cold Spring Harb. Symp. Quant. Biol.* 75, 237–244. doi:10.1101/sqb.2010.75.003
- Hormozdiari, F., Alkan, C., Ventura, M., Hajirasouliha, I., Malig, M., Hach, F., et al. (2011). Alu Repeat Discovery and Characterization within Human Genomes. *Genome Res.* 21 (6), 840–849. doi:10.1101/gr.115956.110
- Huang, X., Park, K. M., Gontarz, P., Zhang, B., Pan, J., McKenzie, Z., et al. (2021). OCT4 Cooperates with Distinct ATP-dependent Chromatin Remodelers in Naive and Primed Pluripotent States in Human. *Nat. Commun.* 12 (1), 5123. doi:10.1038/s41467-021-25107-3
- Hur, K., Cejas, P., Feliu, J., Moreno-Rubio, J., Burgos, E., Boland, C. R., et al. (2014). Hypomethylation of Long Interspersed Nuclear Element-1 (LINE-1) Leads to Activation of Proto-Oncogenes in Human Colorectal Cancer Metastasis. *Gut* 63 (4), 635–646. doi:10.1136/gutjnl-2012-304219
- Hurst, T. P., and Magiorkinis, G. (2017). Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (LTRs). *Viruses* 9 (6). doi:10.3390/v9060130
- Hutchins, A. P., and Pei, D. (2015). Transposable Elements at the Center of the Crossroads between Embryogenesis, Embryonic Stem Cells, Reprogramming, and Long Non-coding RNAs. *Sci. Bull. (Beijing)* 60 (20), 1722–1733. doi:10.1007/s11434-015-0905-x
- Iskow, R. C., McCabe, M. T., Mills, R. E., Torene, S., Pittard, W. S., Neuwald, A. F., et al. (2010). Natural Mutagenesis of Human Genomes by Endogenous Retrotransposons. *Cell* 141 (7), 1253–1261. doi:10.1016/j.cell.2010.05.020
- Iwagami, S., Baba, Y., Watanabe, M., Shigaki, H., Miyake, K., Ishimoto, T., et al. (2013). LINE-1 Hypomethylation Is Associated with a Poor Prognosis Among Patients with Curatively Resected Esophageal Squamous Cell Carcinoma. *Ann. Surg.* 257 (3), 449–455. doi:10.1097/SLA.0b013e31826d8602
- Izsvak, Z., Wang, J., Singh, M., Mager, D. L., and Hurst, L. D. (2016). Pluripotency and the Endogenous Retrovirus HERVH: Conflict or Serendipity? *Bioessays* 38 (1), 109–117. doi:10.1002/bies.201500096
- Jang, H. S., Shah, N. M., Du, A. Y., Dailey, Z. Z., Pehrsson, E. C., Godoy, P. M., et al. (2019). Transposable Elements Drive Widespread Expression of Oncogenes in Human Cancers. *Nat. Genet.* 51 (4), 611–617. doi:10.1038/s41588-019-0373-3
- Johnson, R., and Guigo, R. (2014). The RIDL Hypothesis: Transposable Elements as Functional Domains of Long Noncoding RNAs. *RNA* 20 (7), 959–976. doi:10.1261/rna.044560.114
- Jones, P. A., and Gonzalgo, M. L. (1997). Altered DNA Methylation and Genome Instability: a New Pathway to Cancer? *Proc. Natl. Acad. Sci. U. S. A.* 94 (6), 2103–2105. doi:10.1073/pnas.94.6.2103
- Jung, H., Choi, J. K., and Lee, E. A. (2018). Immune Signatures Correlate with L1 Retrotransposition in Gastrointestinal Cancers. *Genome Res.* 28 (8), 1136–1146. doi:10.1101/gr.231837.117
- Kami, D., Kitani, T., Nakamura, A., Wakui, N., Mizutani, R., Ohue, M., et al. (2018). The DEAD-Box RNA-Binding Protein DDX6 Regulates Parental RNA Decay for Cellular Reprogramming to Pluripotency. *PLoS One* 13 (10), e0203708. doi:10.1371/journal.pone.0203708
- Kapusta, A., Kronenberg, Z., Lynch, V. J., Zhuo, X., Ramsay, L., Bourque, G., et al. (2013). Transposable Elements Are Major Contributors to the Origin, Diversification, and Regulation of Vertebrate Long Noncoding RNAs. *PLoS Genet.* 9 (4), e1003470. doi:10.1371/journal.pgen.1003470
- Kelley, D. R., Hendrickson, D. G., Tenen, D., and Rinn, J. L. (2014). Transposable Elements Modulate Human RNA Abundance and Splicing via Specific RNA-Protein Interactions. *Genome Biol.* 15 (12), 537. doi:10.1186/s13059-014-0537-5
- Kelley, D., and Rinn, J. (2012). Transposable Elements Reveal a Stem Cell-specific Class of Long Noncoding RNAs. *Genome Biol.* 13 (11), R107. doi:10.1186/gb-2012-13-11-r107
- Kim, M. S., Lapkouski, M., Yang, W., and Gellert, M. (2015). Crystal Structure of the V(D)J Recombinase RAG1-RAG2. *Nature* 518 (7540), 507–511. doi:10.1038/nature14174
- Kunarsow, G., Chia, N.-Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.-S., et al. (2010). Transposable Elements Have Rewired the Core Regulatory Network of Human Embryonic Stem Cells. *Nat. Genet.* 42 (7), 631–634. doi:10.1038/ng.600
- Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., 3rd, et al. (2012). Landscape of Somatic Retrotransposition in Human Cancers. *Science* 337 (6097), 967–971. doi:10.1126/science.1222077
- Leis, O., Eguara, A., Lopez-Arribillaga, E., Alberdi, M. J., Hernandez-Garcia, S., Elorriaga, K., et al. (2012). Sox2 Expression in Breast Tumours and Activation in Breast Cancer Stem Cells. *Oncogene* 31 (11), 1354–1365. doi:10.1038/onc.2011.338
- Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1997). DNA Methylation and Genetic Instability in Colorectal Cancer Cells. *Proc. Natl. Acad. Sci. U. S. A.* 94 (6), 2545–2550. doi:10.1073/pnas.94.6.2545
- Li, M., Radvanyi, L., Yin, B., Rycak, K., Li, J., Chivukula, R., et al. (2017). Downregulation of Human Endogenous Retrovirus Type K (HERV-K) Viral Env RNA in Pancreatic Cancer Cells Decreases Cell Proliferation and Tumor Growth. *Clin. Cancer Res.* 23 (19), 5892–5911. doi:10.1158/1078-0432.CCR-17-0001
- Li, W., Lee, M. H., Henderson, L., Tyagi, R., Bachani, M., Steiner, J., et al. (2015). Human Endogenous Retrovirus-K Contributes to Motor Neuron Disease. *Sci. Transl. Med.* 7 (307), 307ra153. doi:10.1126/scitranslmed.aac8201
- Liu, C., Yang, Y., and Schatz, D. G. (2019). Structures of a RAG-like Transposase during Cut-And-Paste Transposition. *Nature* 575 (7783), 540–544. doi:10.1038/s41586-019-1753-7
- Lu, X., Sachs, F., Ramsay, L., Jacques, P. E., Goke, J., Bourque, G., et al. (2014). The Retrovirus HERVH Is a Long Noncoding RNA Required for Human Embryonic Stem Cell Identity. *Nat. Struct. Mol. Biol.* 21 (4), 423–425. doi:10.1038/nsmb.2799
- McClintock, B. (1950). The Origin and Behavior of Mutable Loci in Maize. *Proc. Natl. Acad. Sci. U. S. A.* 36 (6), 344–355. doi:10.1073/pnas.36.6.344
- Mi, S., Lee, X., Li, X., Veldman, G. M., Finnerty, H., Racie, L., et al. (2000). Syncytin Is a Captive Retroviral Envelope Protein Involved in Human Placental Morphogenesis. *Nature* 403 (6771), 785–789. doi:10.1038/35001608
- Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K. W., et al. (1992). Disruption of the APC Gene by a Retrotransposon Insertion of L1 Sequence in a Colon Cancer. *Cancer Res.* 52 (3), 643–645.
- Misiak, B., Ricceri, L., and Sasiadek, M. M. (2019). Transposable Elements and Their Epigenetic Regulation in Mental Disorders: Current Evidence in the Field. *Front. Genet.* 10, 580. doi:10.3389/fgene.2019.00580
- Muller, M., Hermann, P. C., Liebau, S., Weidgang, C., Seufferlein, T., Kleger, A., et al. (2016). The Role of Pluripotency Factors to Drive Stemness in Gastrointestinal Cancer. *Stem Cell Res.* 16 (2), 349–357. doi:10.1016/j.scr.2016.02.005

- Muotri, A. R., Marchetto, M. C., Coufal, N. G., Oefner, R., Yeo, G., Nakashima, K., et al. (2010). L1 Retrotransposition in Neurons Is Modulated by MeCP2. *Nature* 468 (7322), 443–446. doi:10.1038/nature09544
- Nakamura, Y., Murata, M., Takagi, Y., Kozuka, T., Nakata, Y., Hasebe, R., et al. (2015). SVA Retrotransposition in Exon 6 of the Coagulation Factor IX Gene Causing Severe Hemophilia B. *Int. J. Hematol.* 102 (1), 134–139. doi:10.1007/s12185-015-1765-5
- Ogino, S., Noshio, K., Kirkner, G. J., Kawasaki, T., Chan, A. T., Schernhammer, E. S., et al. (2008). A Cohort Study of Tumoral LINE-1 Hypomethylation and Prognosis in Colon Cancer. *J. Natl. Cancer Inst.* 100 (23), 1734–1738. doi:10.1093/jnci/djn359
- Ohno, S. (1972). So Much "junk" DNA in Our Genome. *Brookhaven Symp. Biol.* 23, 366–370.
- Ovejero, T., Sadones, O., Sanchez-Fito, T., Almenar-Perez, E., Espejo, J. A., Martin-Martinez, E., et al. (2020). Activation of Transposable Elements in Immune Cells of Fibromyalgia Patients. *Int. J. Mol. Sci.* 21 (4). doi:10.3390/ijms21041366
- Pattamadilok, J., Huapai, N., Rattanananyong, P., Vasurattana, A., Triratanachai, S., Tresukosol, D., et al. (2008). LINE-1 Hypomethylation Level as a Potential Prognostic Factor for Epithelial Ovarian Cancer. *Int. J. Gynecol. Cancer* 18 (4), 711–717. doi:10.1111/j.1525-1438.2007.01117.x
- Payer, L. M., and Burns, K. H. (2019). Transposable Elements in Human Genetic Disease. *Nat. Rev. Genet.* 20 (12), 760–772. doi:10.1038/s41576-019-0165-8
- Pehrsson, E. C., Choudhary, M. N. K., Sundaram, V., and Wang, T. (2019). The Epigenomic Landscape of Transposable Elements across Normal Human Development and Anatomy. *Nat. Commun.* 10 (1), 5640. doi:10.1038/s41467-019-13555-x
- Percharde, M., Lin, C. J., Yin, Y., Guan, J., Peixoto, G. A., Bulut-Karslioglu, A., et al. (2018). A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* 174 (2), 391–405. e319. doi:10.1016/j.cell.2018.05.043
- Perron, H., Garson, J. A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., et al. (1997). Molecular Identification of a Novel Retrovirus Repeatedly Isolated from Patients with Multiple Sclerosis. The Collaborative Research Group on Multiple Sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* 94 (14), 7583–7588. doi:10.1073/pnas.94.14.7583
- Perron, H., Suh, M., Lalande, B., Gratacap, B., Laurent, A., Stoebner, P., et al. (1993). Herpes Simplex Virus ICP0 and ICP4 Immediate Early Proteins Strongly Enhance Expression of a Retrovirus Harboured by a Leptomeningeal Cell Line from a Patient with Multiple Sclerosis. *J. Gen. Virol.* 74 (Pt 1), 65–72. doi:10.1099/0022-1317-74-1-65
- Pontis, J., Planet, E., Offner, S., Turelli, P., Duc, J., Coudray, A., et al. (2019). Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* 24 (5), 724–735. e725. doi:10.1016/j.stem.2019.03.012
- Rishishwar, L., Tellez Villa, C. E., and Jordan, I. K. (2015). Transposable Element Polymorphisms Recapitulate Human Evolution. *Mob. DNA* 6, 21. doi:10.1186/s13100-015-0052-6
- Rodic, N., Sharma, R., Sharma, R., Zampella, J., Dai, L., Taylor, M. S., et al. (2014). Long Interspersed Element-1 Protein Expression Is a Hallmark of Many Human Cancers. *Am. J. Pathol.* 184 (5), 1280–1286. doi:10.1016/j.ajpath.2014.01.007
- Rodriguez-Martin, C., Cidre, F., Fernandez-Teijeiro, A., Gomez-Mariano, G., de la Vega, L., Ramos, P., et al. (2016). Familial Retinoblastoma Due to Intronic LINE-1 Insertion Causes Aberrant and Noncanonical mRNA Splicing of the RB1 Gene. *J. Hum. Genet.* 61 (5), 463–466. doi:10.1038/jhg.2015.173
- Rodriguez-Terrones, D., and Torres-Padilla, M. E. (2018). Nimble and Ready to Mingle: Transposon Outbursts of Early Development. *Trends Genet.* 34 (10), 806–820. doi:10.1016/j.tig.2018.06.006
- Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Cervantes, F., Sanchez, J., Garate, L., et al. (2005). Promoter Hypomethylation of the LINE-1 Retrotransposable Elements Activates Sense/antisense Transcription and Marks the Progression of Chronic Myeloid Leukemia. *Oncogene* 24 (48), 7213–7223. doi:10.1038/sj.onc.1208866
- Rowe, H. M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T. S., Tarabay, Y., et al. (2013). TRIM28 Repression of Retrotransposon-Based Enhancers Is Necessary to Preserve Transcriptional Dynamics in Embryonic Stem Cells. *Genome Res.* 23 (3), 452–461. doi:10.1101/gr.147678.112
- Rycaj, K., Plummer, J. B., Yin, B., Li, M., Garza, J., Radvanyi, L., et al. (2015). Cytotoxicity of Human Endogenous Retrovirus K-specific T Cells toward Autologous Ovarian Cancer Cells. *Clin. Cancer Res.* 21 (2), 471–483. doi:10.1158/1078-0432.CCR-14-0388
- Saleh, A., Macia, A., and Muotri, A. R. (2019). Transposable Elements, Inflammation, and Neurological Disease. *Front. Neurol.* 10, 894. doi:10.3389/fneur.2019.00894
- Sanchez-Luque, F. J., Kempen, M. H. C., Gerdes, P., Vargas-Landin, D. B., Richardson, S. R., Troskie, R. L., et al. (2019). LINE-1 Evasion of Epigenetic Repression in Humans. *Mol. Cell* 75 (3), 590–604. e512. doi:10.1016/j.molcel.2019.05.024
- Schiavetti, F., Thonnard, J., Colau, D., Boon, T., and Coulie, P. G. (2002). A Human Endogenous Retroviral Sequence Encoding an Antigen Recognized on Melanoma by Cytolytic T Lymphocytes. *Cancer Res.* 62 (19), 5510–5516.
- Scott, E. C., Gardner, E. J., Masood, A., Chuang, N. T., Vertino, P. M., and Devine, S. E. (2016). A Hot L1 Retrotransposon Evades Somatic Repression and Initiates Human Colorectal Cancer. *Genome Res.* 26 (6), 745–755. doi:10.1101/gr.201814.115
- Shigaki, H., Baba, Y., Watanabe, M., Murata, A., Iwagami, S., Miyake, K., et al. (2013). LINE-1 Hypomethylation in Gastric Cancer, Detected by Bisulfite Pyrosequencing, Is Associated with Poor Prognosis. *Gastric Cancer* 16 (4), 480–487. doi:10.1007/s10120-012-0209-7
- Shpyleva, S., Melnyk, S., Pavliv, O., Pogribny, I., and Jill James, S. (2018). Overexpression of LINE-1 Retrotransposons in Autism Brain. *Mol. Neurobiol.* 55 (2), 1740–1749. doi:10.1007/s12035-017-0421-x
- Shukla, R., Upton, K. R., Munoz-Lopez, M., Gerhardt, D. J., Fisher, M. E., Nguyen, T., et al. (2013). Endogenous Retrotransposition Activates Oncogenic Pathways in Hepatocellular Carcinoma. *Cell* 153 (1), 101–111. doi:10.1016/j.cell.2013.02.032
- Sidorenko, L. V., Lee, T. F., Woosley, A., Moskal, W. A., Bevan, S. A., Merlo, P. A. O., et al. (2017). GC-rich Coding Sequences Reduce Transposon-like, Small RNA-Mediated Transgene Silencing. *Nat. Plants* 3 (11), 875–884. doi:10.1038/s41477-017-0040-6
- Singh, M., Cai, H., Bunse, M., Feschotte, C., and Izsvak, Z. (2020). Human Endogenous Retrovirus K Rec Forms a Regulatory Loop with MITF that Opposes the Progression of Melanoma to an Invasive Stage. *Viruses* 12 (11). doi:10.3390/v12111303
- Solyom, S., Ewing, A. D., Rahrmann, E. P., Doucet, T., Nelson, H. H., Burns, M. B., et al. (2012). Extensive Somatic L1 Retrotransposition in Colorectal Tumors. *Genome Res.* 22 (12), 2328–2338. doi:10.1101/gr.145235.112
- Spirito, G., Mangoni, D., Sanges, R., and Gustincich, S. (2019). Impact of Polymorphic Transposable Elements on Transcription in Lymphoblastoid Cell Lines from Public Data. *BMC Bioinforma.* 20 (Suppl. 9), 495. doi:10.1186/s12859-019-3113-x
- Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et al. (2015). An Integrated Map of Structural Variation in 2,504 Human Genomes. *Nature* 526 (7571), 75–81. doi:10.1038/nature15394
- Tang, Z., Steranka, J. P., Ma, S., Grivainis, M., Rodic, N., Huang, C. R., et al. (2017). Human Transposon Insertion Profiling: Analysis, Visualization and Identification of Somatic LINE-1 Insertions in Ovarian Cancer. *Proc. Natl. Acad. Sci. U. S. A.* 114 (5), E733–E740. doi:10.1073/pnas.1619797114
- Tangsuwanri, C., Saeli, T., Thongkorn, S., Chonchaiya, W., Suphapeetiporn, K., Mutirangura, A., et al. (2018). Investigation of Epigenetic Regulatory Networks Associated with Autism Spectrum Disorder (ASD) by Integrated Global LINE-1 Methylation and Gene Expression Profiling Analyses. *PLoS One* 13 (7), e0201071. doi:10.1371/journal.pone.0201071
- Thomas, C. A., Tejwani, L., Trujillo, C. A., Negraes, P. D., Herai, R. H., Mesci, P., et al. (2017). Modeling of TREX1-dependent Autoimmune Disease Using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. *Cell Stem Cell* 21 (3), 319–331. e318. doi:10.1016/j.stem.2017.07.009
- Thornburg, B. G., Gotea, V., and Makalowski, W. (2006). Transposable Elements as a Significant Source of Transcription Regulating Signals. *Gene* 365, 104–110. doi:10.1016/j.gene.2005.09.036
- Trizzino, M., Park, Y., Holsbach-Beltrame, M., Aracena, K., Mika, K., Caliskan, M., et al. (2017). Transposable Elements Are the Primary Source of Novelty in Primate Gene Regulation. *Genome Res.* 27 (10), 1623–1633. doi:10.1101/gr.218149.116
- Tubio, J. M. C., Li, Y., Ju, Y. S., Martincorena, I., Cooke, S. L., Tojo, M., et al. (2014). Mobile DNA in Cancer. Extensive Transduction of Nonrepetitive DNA

- Mediated by L1 Retrotransposition in Cancer Genomes. *Science* 345 (6196), 1251343. doi:10.1126/science.1251343
- Turelli, P., Castro-Diaz, N., Marzetta, F., Kapopoulou, A., Raclot, C., Duc, J., et al. (2014). Interplay of TRIM28 and DNA Methylation in Controlling Human Endogenous Retroelements. *Genome Res.* 24 (8), 1260–1270. doi:10.1101/gr.172833.114
- van Hoesel, A. Q., van de Velde, C. J., Kuppen, P. J., Liefers, G. J., Putter, H., Sato, Y., et al. (2012). Hypomethylation of LINE-1 in Primary Tumor Has Poor Prognosis in Young Breast Cancer Patients: a Retrospective Cohort Study. *Breast Cancer Res. Treat.* 134 (3), 1103–1114. doi:10.1007/s10549-012-2038-0
- Van Meter, M., Kashyap, M., Rezazadeh, S., Geneva, A. J., Morello, T. D., Seluanov, A., et al. (2014). SIRT6 Represses LINE1 Retrotransposons by Ribosylating KAP1 but This Repression Fails with Stress and Age. *Nat. Commun.* 5, 5011. doi:10.1038/ncomms6011
- Van Nostrand, E. L., Freese, P., Pratt, G. A., Wang, X., Wei, X., Xiao, R., et al. (2020). A Large-Scale Binding and Functional Map of Human RNA-Binding Proteins. *Nature* 583 (7818), 711–719. doi:10.1038/s41586-020-2077-3
- Wang, J., Huang, J., and Shi, G. (2020). Retrotransposons in Pluripotent Stem Cells. *Cell Regen.* 9 (1), 4. doi:10.1186/s13619-020-00046-4
- Wang, J., Xie, G., Singh, M., Ghanbarian, A. T., Rasko, T., Szvetnik, A., et al. (2014). Primate-specific Endogenous Retrovirus-Driven Transcription Defines Naive-like Stem Cells. *Nature* 516 (7531), 405–409. doi:10.1038/nature13804
- Wang, X., Liu, Z., Wang, P., Li, S., Zeng, J., Tu, X., et al. (2018). Syncytin-1, an Endogenous Retroviral Protein, Triggers the Activation of CRP via TLR3 Signal Cascade in Glial Cells. *Brain Behav. Immun.* 67, 324–334. doi:10.1016/j.bbi.2017.09.009
- Wang-Johanning, F., Frost, A. R., Jian, B., Azerou, R., Lu, D. W., Chen, D. T., et al. (2003). Detecting the Expression of Human Endogenous Retrovirus E Envelope Transcripts in Human Prostate Adenocarcinoma. *Cancer* 98 (1), 187–197. doi:10.1002/cncr.11451
- Wang-Johanning, F., Radvanyi, L., Rycaj, K., Plummer, J. B., Yan, P., Sastry, K. J., et al. (2008). Human Endogenous Retrovirus K Triggers an Antigen-specific Immune Response in Breast Cancer Patients. *Cancer Res.* 68 (14), 5869–5877. doi:10.1158/0008-5472.CAN-07-6838
- Wanichnopparat, W., Suwanwongse, K., Pin-On, P., Aporntewan, C., and Mutirangura, A. (2013). Genes Associated with the Cis-Regulatory Functions of Intrinsic LINE-1 Elements. *BMC Genomics* 14, 205. doi:10.1186/1471-2164-14-205
- Weber, B., Kimhi, S., Howard, G., Eden, A., and Lyko, F. (2010). Demethylation of a LINE-1 Antisense Promoter in the cMet Locus Impairs Met Signalling through Induction of Illegitimate Transcription. *Oncogene* 29 (43), 5775–5784. doi:10.1038/onc.2010.227
- Wolff, E. M., Byun, H. M., Han, H. F., Sharma, S., Nichols, P. W., Siegmund, K. D., et al. (2010). Hypomethylation of a LINE-1 Promoter Activates an Alternate Transcript of the MET Oncogene in Bladders with Cancer. *PLoS Genet.* 6 (4), e1000917. doi:10.1371/journal.pgen.1000917
- Wu, Z., Mei, X., Zhao, D., Sun, Y., Song, J., Pan, W., et al. (2015). DNA Methylation Modulates HERV-E Expression in CD4+ T Cells from Systemic Lupus Erythematosus Patients. *J. Dermatol. Sci.* 77 (2), 110–116. doi:10.1016/j.jdermsci.2014.12.004
- Yang, F., and Wang, P. J. (2016). Multiple LINEs of Retrotransposon Silencing Mechanisms in the Mammalian Germline. *Semin. Cell Dev. Biol.* 59, 118–125. doi:10.1016/j.semcdb.2016.03.001
- Yang, Z. N., Mueser, T. C., Kaufman, J., Stahl, S. J., Wingfield, P. T., and Hyde, C. C. (1999). The Crystal Structure of the SIV Gp41 Ectodomain at 1.47 Å Resolution. *J. Struct. Biol.* 126 (2), 131–144. doi:10.1006/j.sbi.1999.4116
- Zamudio, N., and Bourc'his, D. (2010). Transposable Elements in the Mammalian Germline: a Comfortable Niche or a Deadly Trap? *Hered. (Edinb)* 105 (1), 92–104. doi:10.1038/hdy.2010.53
- Zapatka, M., Borozan, I., Brewer, D. S., Iskar, M., Grundhoff, A., Alawi, M., et al. (2020). The Landscape of Viral Associations in Human Cancers. *Nat. Genet.* 52 (3), 320–330. doi:10.1038/s41588-019-0558-9
- Zhang, X., Zhang, R., and Yu, J. (2020). New Understanding of the Relevant Role of LINE-1 Retrotransposition in Human Disease and Immune Modulation. *Front. Cell Dev. Biol.* 8, 657. doi:10.3389/fcell.2020.00657
- Zhang, Y., Cheng, T. C., Huang, G., Lu, Q., Surleac, M. D., Mandell, J. D., et al. (2019). Transposon Molecular Domestication and the Evolution of the RAG Recombinase. *Nature* 569 (7754), 79–84. doi:10.1038/s41586-019-1093-7
- Zhou, X., Sam, T. W., Lee, A. Y., and Leung, D. (2021). Mouse Strain-specific Polymorphic Provirus Functions as Cis-Regulatory Element Leading to Epigenomic and Transcriptomic Variations. *Nat. Commun.* 12 (1), 6462. doi:10.1038/s41467-021-26630-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ma, Babarinde, Zhou and Hutchins. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.