Received: 14 March 2024 | Revised: 12 April 2024

DOI: 10.1002/bies.202400059

BioEssays

REVIEW ESSAY

Prospects & Overviews

Transposable elements as drivers of dedifferentiation: Connections between enhancers in embryonic stem cells, placenta, and cancer

Konsta Karttunen[1] | Divyesh Patel[1,2] | Biswajyoti Sahu[1,2,3,4]

[1]Applied Tumor Genomics Program, Research Programs Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland

[2]iCAN Digital Precision Cancer Medicine Flagship, University of Helsinki, Helsinki, Finland

[3]Centre for Molecular Medicine Norway, University of Oslo, Oslo, Norway

[4]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway

Correspondence
Biswajyoti Sahu, Centre for Molecular Medicine, University of Oslo, 0318 Oslo, Norway.
Email: biswajyoti.sahu@ncmm.uio.no

Funding information
Academy of Finland, Grant/Award Numbers: 317807, 320114, 346065; iCAN Digital Precision Cancer Medicine Flagship, Grant/Award Number: 320185; Cancer Foundation Finland; Sigrid Jusélius Foundation; Jane and Aatos Erkko Foundation; Centre for Molecular Medicine Norway, University of Oslo

Abstract

Transposable elements (TEs) have emerged as important factors in establishing the cell type-specific gene regulatory networks and evolutionary novelty of embryonic and placental development. Recently, studies on the role of TEs and their dysregulation in cancers have shed light on the transcriptional, transpositional, and regulatory activity of TEs, revealing that the activation of developmental transcriptional programs by TEs may have a role in the dedifferentiation of cancer cells to the progenitor-like cell states. This essay reviews the recent evidence of the *cis*-regulatory TEs (henceforth crTE) in normal development and malignancy as well as the key transcription factors and regulatory pathways that are implicated in both cell states, and presents existing gaps remaining to be studied, limitations of current technologies, and therapeutic possibilities.

KEYWORDS
cancer, dedifferentiation, development, epigenetics, genomics, immunotherapy, transposable elements

INTRODUCTION

Transposable elements (TEs), first discovered by McClintock,[1,2] are genomic elements capable of moving to new loci that are found in almost all eukaryotic genomes. About 50% of the human genome consists of TEs,[3,4] but out of these, only about 100 elements have retained their ability to move autonomously, that is, are transpositionally capable elements.[2,5] TEs are classified into two major classes, retrotransposons ("copy-paste") and DNA transposons ("cut-paste") by their mechanism of transposition, and further into four families known as long and short interspersed nuclear elements (LINEs and

Abbreviations: AML, acute myeloid leukemia; AR, androgen receptor; CRC, colorectal cancer; CRE, *cis*-regulatory element; DNMT, DNA methyltransferase; dsRNA, double-stranded RNA; EMT, epithelial-mesenchymal transition; ERV, endogenous retrovirus; ESC, embryonic stem cell; HDAC, histone deacetylase; HERV, human endogenous retrovirus; KRAB-ZNF, Krüppel-associated box zinc finger; LINE, long interspersed nuclear element; LTR, long terminal repeat; NSCLC, non-small cell lung cancer; piRNA, piwi-interacting RNA; PSC, pluripotent stem cell; SINE, short interspersed nuclear element; siRNA, silencing RNA; TAD, topologically associating domain; TE, transposable element; TF, transcription factor; TPRT, target-primed reverse transcription; TSC, trophoblast stem cells; UTR, untranslated region; VNTR, variable number tandem repeat.

SINEs), long terminal repeats or endogenous retroviruses (LTRs/ERVs), and DNA transposons by their replication and integration mechanisms and into further subfamilies by their phylogenetic relationships[6] (Figure 1A).

The potential role of TEs in gene regulation has been considered since their discovery.[7,8] The majority of TEs are fixed in the genome and incapable of transposition due to mutations or silencing mechanisms. Despite the tight control in normal cellular conditions, it has become evident that the *cis*-regulatory sequences in integrated TEs have been adapted into a myriad of different roles: TEs have evolved to possess multiple *cis*-regulatory roles in the human and other eukaryotic genomes, including promoters, enhancers, silencers, and boundary elements that demarcate topologically associated domains (comprehensively reviewed by, e.g., Sundaram and Wysocka,[9] the "gene-battery" model of TE adaptation to regulatory roles by Britten and Davidson[7] evaluated in light of modern evidence by Sundaram and Wang[10]).

Genome expansion is an important driver of the evolution of gene regulatory networks, providing new binding sites for transcription factors (TFs) in waves of expansions rather than by accumulation of point mutations.[11] TEs have had a key role in this process of regulatory evolution as evidenced by the majority of primate-specific and actively evolving *cis*-regulatory elements (CREs) being derived from TEs.[12–14] Around a quarter of the human *cis*-regulatory genome in normal tissues is comprised of TEs.[15] TEs have an especially outsized role in *cis*-regulation during embryonic and placental development: expansions of TEs have rewired regulatory networks for essential pluripotency factors[16] and placental *cis*-regulation,[17,18] including the direct exaptation of a retroviral envelope protein gene into syncytin, an essential factor in placental development.[19] Other examples of TF programs rewired by TEs include, for example, tumor suppressor p53[20] and innate immunity factor STAT1.[21]

Due to their mutagenicity and *cis*-regulatory potential, TE activity is tightly controlled in normal cellular conditions by multiple overlapping mechanisms. These include piwi-interacting RNAs (piRNAs), small interfering RNAs (siRNAs), Krüppel-associated box domain-containing zinc-finger proteins (KRAB-ZNFs) and p53 binding, DNA methylation, deposition of repressive histone modifications (Figure 1B), or ultimately domesticating the TE into functional, endogenous roles in the host genome (Figure 1C) (comprehensively reviewed by Almeida et al.[22]). TE silencing mechanisms are often multilayered: for example, a recent genome-wide CRISPR screen identified potential LINE1 regulators involved in diverse functional pathways such as chromatin remodeling, RNA splicing, N6-methylation of RNA, and RNA and protein degradation pathways.[23] However, the silencing is often disrupted in cancers, not only due to the loss of repressive chromatin modifications but also by transcript filtering mechanisms.[24,25] This can result in activation of TE transcription, transposition, and *cis*-regulatory activity that can have effects on cancer initiation and progression (Figure 2). Most notably, the causative role of transposition events in cancer initiation has been reported in colorectal cancer (CRC) with TE insertions resulting in disruptions of the *APC* gene.[26–28] On the other hand, activation of the *cis*-regulatory function through epigenetic de-regulation of cryptic promoter sites within TEs is known to drive the expression of oncogenes in multiple cancers.[29] However, the whole extent of crTE activation in driving cancer-specific gene regulatory programs, especially via the activation of cryptic enhancers, is still mostly unknown, despite the well-known dysregulation of epigenetic control mechanisms in cancer and the importance of TE regulation in development and other controlled cellular processes.

Transformation of cells into cancer by oncogene activation is highly dependent on cell state: progenitor-like cells are more susceptible to transformation compared to fully differentiated cells,[30,31] and metastatic progression is especially linked to activation of developmental (epi)genetic programs and enhancer reprogramming.[32,33] Unlocking phenotypic plasticity through dedifferentiation of cancer cells into progenitor-like states is considered one of the emerging hallmarks of cancer,[34] and recent evidence has shown remarkable similarities between cancers and pluripotent stem cells in TE activation and epigenetic dysregulation: Lynch-Sutherland et al.[35] have proposed the hypothesis that TE-driven regulation specifically reawakens the developmental transcriptional programs in cancers, and new advancements since then have provided intriguing new evidence on the roles of crTEs in cancers. This essay synthesizes the new evidence, identifying the key TEs and TFs that bind them, as well as regulatory pathways through which they may contribute to malignant states, providing some insights into what remains unknown and what needs to be researched further.

PLURIPOTENCY FACTORS COMMONLY BIND TEs IN EMBRYONIC STEM CELLS (ESCs)

There is a large body of evidence of TEs, especially LINE and LTR elements, as a major factor in gene regulatory network evolution of zygotic gene activation and post- and pre-implantation gene expression (comprehensively reviewed by Senft and Macfarlan[36]). The epigenetic landscape of embryonic development is especially permissive for TE activity, as embryonic stem cells undergo a complete reprogramming of TE methylation in the early stages of development.[37] In total, TEs contribute to about 20% of TF binding sites in both human and mouse.[38] TEs have also rewired new genes in embryonic development to the pluripotency-maintaining transcriptional programs, with only a small fraction of OCT4 and NANOG binding sites shared between human and mouse, whereas CTCF has a more similar binding profile between both.[16,39]

Multiple genome-wide studies have focused on the canonical pluripotency factors (OCT4, SOX2, and NANOG) and their occupancy at TEs. About 24% of OCT4 and SOX2 binding was found to occur at ERVK repeats in mice over an expected frequency of 9%,[40] and in humans roughly similar percentages of 21% and 15% were found for OCT4 and NANOG, respectively.[16] Analyses of large-scale genomics data have subsequently revealed that TEs contribute to a large fraction of open chromatin regions, with especially the LTR class overrepresented in ESCs,[13] and demethylated TEs in ESCs have been shown to associate with tissue-specific enhancers.[41]

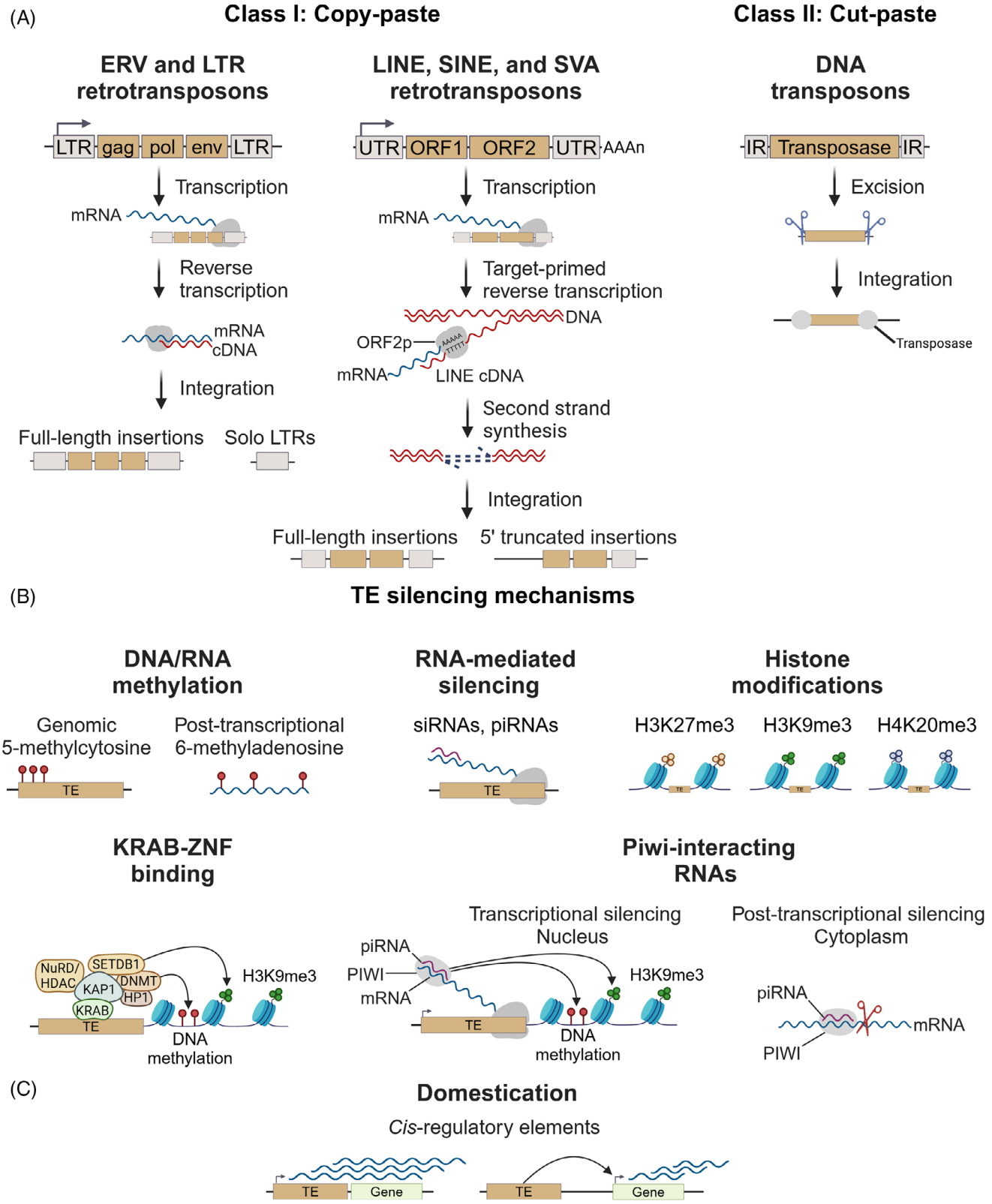

FIGURE 1 Mechanisms of TE transposition and silencing. (A) TEs are divided into two major classes based on their transposition mechanisms: Class I, that is, "copy-paste" and Class II, that is, "cut-paste." ERV/LTR elements transpose via an mRNA intermediate that encodes for gag, pol, and env proteins that are required for the transpositional machinery. The mRNA is reverse-transcribed to dsDNA in the cytoplasm and integrated back into the host genome. These insertions often appear as solo LTR insertions, where the internal sequence is lost due to recombination between 5′ and 3′ LTRs. LINE elements encode a bicistronic mRNA that contains reading frames for an RNA-binding ORF1p and ORF2p that has endonuclease and reverse transcriptase activity. LINEs transpose via a target-primed reverse transcription mechanism (TPRT), where the ORF2p endonuclease

The transient *cis*-regulatory and transcriptional activity of TEs during embryonic development is a carefully coordinated process with waves of transcriptional activation and deactivation of specific subfamilies of TEs. TEs are essential for hESC genome activation,[42,43] and LTR transcription is a hallmark of embryonic development that is directly regulated by the binding of pluripotency factors.[44,45] In human embryos, the earliest TE activation measured by RNA expression and chromatin accessibility seems to occur at the four-cell stage, with dynamic activation of different TE subfamilies depending on the stage of development.[46,47] In mouse embryos, LINE1 and MERVL are already transcribed at the two-cell stage and required for zygotic gene activation and progression from the two-cell stage.[23,48,49] Young LINE1s are also expressed in humans in the early embryo,[50] and the induction of transcription positively correlates with gene expression at the eight-cell stage: for example, two L1PA2 elements induced the expression of SIX2, a TF involved in early organ development.[23] Thus, the coordinated expression of specific TEs during development is not merely a side effect of the lax epigenetic control during reprogramming of DNA methylation but has concrete functional consequences for zygotic gene activation and the maintenance of pluripotency.[16,45,51,52] Collectively, the role of TEs in the developmental processes has been extensively characterized, and it is well established that specific TEs have crucial roles during development and in maintaining pluripotency.

PLACENTAL GENE EXPRESSION IS REGULATED BY TEs

During placentation, trophoblast stem cells (TSCs) invade the endometrium to form the placenta after embryonal implantation and connect the maternal blood circulation to the embryonic circulation. Placental development has necessitated remarkably similar cellular features that are also required for progression of cancer, such as tissue invasion via epithelial to mesenchymal transition (EMT), immunosuppression, and stimulation of angiogenesis.[34,53] Interestingly, comparative research has shown that organisms with invasive placentas may be more susceptible to cancers and metastases than metatherian non-placentals such as marsupials. However, other species-specific suppressive mechanisms could also explain the differences, since some placental organisms such as elephants and mole rats are also highly resistant to cancer.[54,55] Moreover, an invasive phenotype is also essential for other non-placental functions such as wound healing, so mechanisms such as placental immunomodulation can also contribute to the higher susceptibility in placental organisms.

The evolution of placental mammals has been largely facilitated by the rewiring of gene regulatory networks by TE expansions.[18] Multiple TE subfamilies have been reported to regulate placental gene expression in TSCs in both mice and in human, with the LTRs being highly overrepresented among the active elements.[17,56–58] One major player seems to be the elements of the MER50 subfamily, which were identified by both Yu et al.[58] and Sun et al.[57] as active TE-enhancers. Yu et al.[58] characterized the MER50 elements in more detail and found that they regulate genes essential for the formation of the syncytiotrophoblast, a continuous layer of epithelial cells that forms the interface between the fetal and maternal blood. About 20% of the TE enhancers that were identified by their H3K27ac profile in TSCs also had bivalent H3K9me3 methylation that was mostly lost during differentiation to syncytiotrophoblasts, indicating a poised enhancer state of these TEs that is activated during transition to another cell type.[58]

Other major regulators of placental gene expression are the MER41A and MER41B subfamilies identified in three independent studies as CREs.[56–58] MER41B seems to have various *cis*-regulatory roles: it has also been found to be adapted to innate immune responses,[21] and Frost et al.[56] reported it regulating genes that are essential for trophoblast development and corroborated the enrichment of STAT and SRF motifs as reported earlier.[21,57] Other active TEs identified in more than one of these studies are LTR10A, LTR8B, MER21A, MER39, and MER11D, all of which seem unique to TSCs as they show no activity in ESCs.[56] These TEs were found to bind many key TFs involved in stem cell state maintenance, such as ELF5, GATA3, TFAP2C, TP63, and TEAD4, speaking of the unique regulatory environment of the placenta molded by TEs. It is worth noting that the placenta is a transient organ that is discarded after parturition, and thus it is possible that there has been no selective pressure for controlling the deleterious effects of TE transposition, enabling widespread TE adaptation to *cis*-regulatory functions. Interestingly, in addition to serving as *cis*-regulatory elements, TEs contribute directly to placental gene expression: several genes that are functional in the placenta are directly adapted from TEs, such as the syncytin gene that is derived from a retroviral envelope protein gene.[19] Collectively, TEs have been instrumental for the development of the placenta, showing unique co-option from the syncytin protein to widespread *cis*-regulatory activity.

nicks DNA at a TTAAAA motif, freeing a 3′ OH group that primes the reverse transcription of the LINE RNA. The cDNA is integrated, and the second strand is subsequently synthesized. DNA elements encode for a transposase that can nick the DNA sequence and integrate it into a new locus in the host genome. The non-coding sequences in each TE type are marked in gray and the coding sequences are marked in brown. (B) TEs are silenced by different mechanisms. General mechanisms include DNA methylation, repressive histone modifications and RNA-mediated silencing. These silencing mechanisms are especially conferred by KRAB-ZNFs that recruit DNA methylases, histone deacetylases, and histone methylases at TEs, and piRNAs that are functional in the male germline that utilize similar mechanisms but can also post-transcriptionally nick TE-mRNA in the cytoplasm. (C) TEs can be domesticated or co-opted into new roles such as promoters or enhancers in the genome by accumulation of mutations, diminishing their deleterious effects and rewiring new gene regulatory networks.

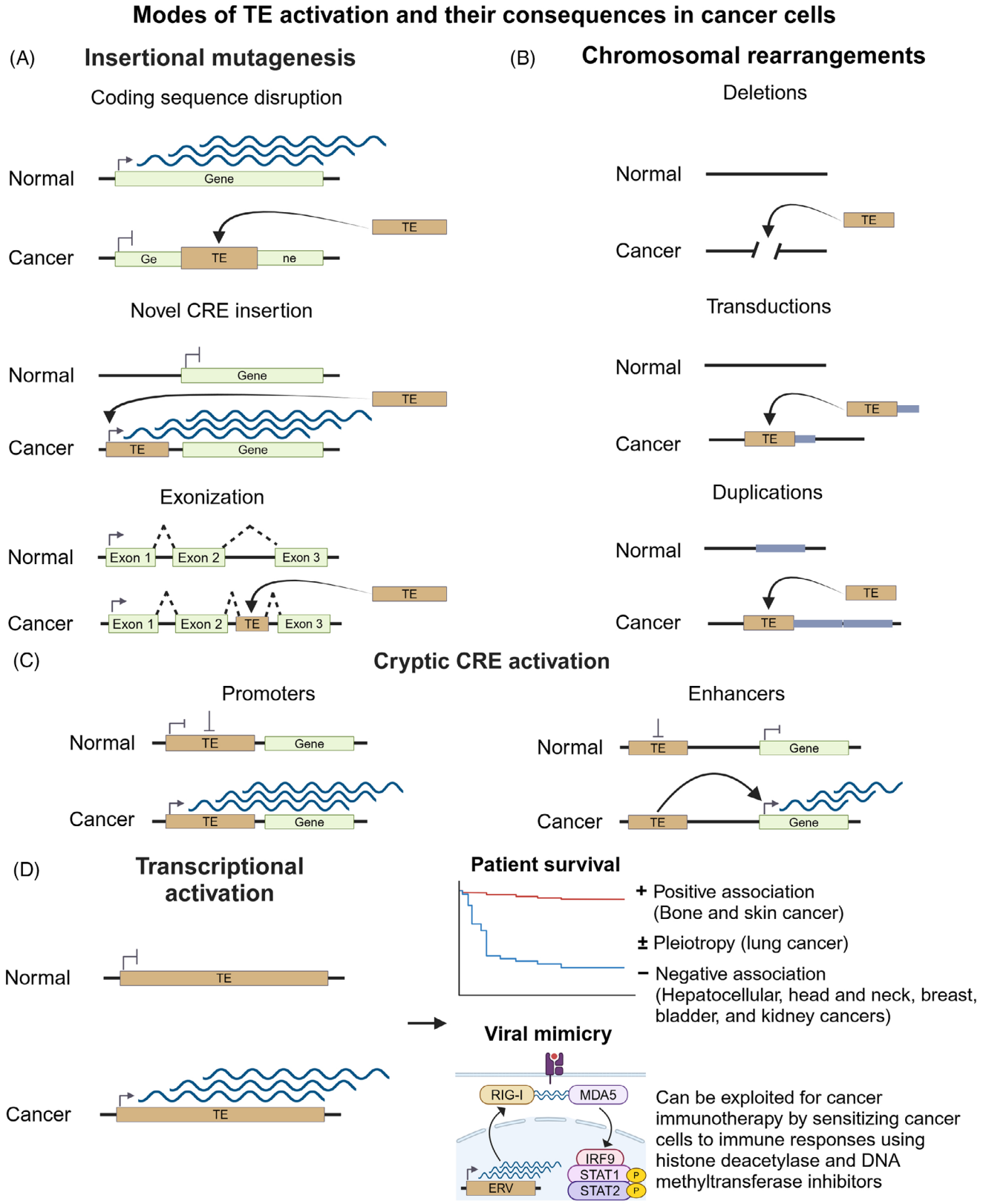

FIGURE 2 Modes of TE activation and their consequences in cancer cells. (A) De novo insertions resulting in from the ~100 active LINEs in cancer can cause insertional mutagenesis, where the coding sequences of tumor suppressors, such as *APC*, can be disrupted. Other mechanisms include novel *cis*-regulatory elements carried by insertions and exonization of insertions in gene sequences. (B) Insertions can cause megabase-scale chromosomal rearrangements, including deletions, transductions, duplications, and breakage-fusion-bridge cycles that can cause oncogene amplifications. (C) Extant germline TE insertions can be activated due to the dysregulation of epigenetic control mechanisms. These elements can contain *cis*-regulatory sequences, such as enhancers and cryptic promoters that can drive the expression of nearby oncogenes. (D)

MUTATIONAL AND *cis*-REGULATORY ROLES OF TEs IN CANCER

Activation of TE transcription and transposition has been extensively studied in pan-cancer analyses that have found large differences in TE activity and transposition rates between cancer types,[5] colorectal and esophageal carcinomas as well as head and lung small cell carcinomas showing the highest transposition activity. Transposition is known to cause insertional mutagenesis, genomic instability, and large-scale genomic rearrangements in cancer (Figure 2a,b).[5] The transpositional activity is driven by around 100 intact germline copies of the evolutionarily young LINE subfamilies (L1Hs, L1PA2) that still harbor full-length and transposition-capable sequences.[2,5,59] Whole genome sequencing studies can efficiently map TE transposition events and their mutagenic effects, but functional genomics and transcriptomics assays are needed for a comprehensive understanding of regulatory TE activity in the epigenomic level. TE transcription rates have also been reported to vary between cancer types, but largely the same repertoire of young LINEs are also transcriptionally overexpressed in the same cancer types and correlated with demethylation and a heightened interferon response.[60,61] However, TE transcription and *cis*-regulatory activity are not limited to full-length LINEs but can originate from all TE families upon their epigenetic derepression. Li et al.[23] also showed that L1 5′ UTR can transactivate nearby genes. The 5′ UTRs showed functional enhancer activity and were enriched for active enhancer marks such as H3K27ac and H3K4me1. Further systemic perturbation of L1 sequences through CRISPRa and CRISPRi influenced the expression of L1-regulated genes, and transcriptionally active L1 showed increased cohesin occupancy and physical contacts with their target genes.

CRC is one of the cancer types with the highest transposition rates. Interestingly, LINE transposition already starts during embryogenesis, also occurring in the normal colon epithelium, with CRC showing a threefold increase in transpositional activity compared to the normal colon.[62] Nam et al. reported that some of the LINE promoters remain demethylated after the global rearrangement of DNA methylation in the early embryo, resulting in increased LINE insertion burden with age. Thus, early stages of CRC already show transpositional activity, indicating that TE insertions very likely contribute to cancer driver mutations. For example, disruption of the *APC* gene has been shown to initiate CRC, and TE structural variants contribute to tumorigenesis.[5,26–28,63] The role of TE transposition events in the initiation of cancer also raises an interesting question whether this also applies to TEs with *cis*-regulatory activity—can they contribute to initiating cancer or does their activity only appear during later stages of cancer development.

Several epigenetic and filtering mechanisms contribute to TE silencing in somatic cells, and non-mutational epigenetic reprogramming is considered as one of the enabling characteristics in the development of cancers.[25,34] However, whether the mechanisms that cause transcriptional and transpositional TE activity in cancers also correlate with the activation of their *cis*-regulatory features is largely unknown. In recent years, evidence of TE co-option to multiple *cis*-regulatory roles in cancer, known as onco-exaptation, has been emerging (Figure 2C, left panel). Multiple studies have shown that DNMT and HDAC inhibition leads to transcriptional activation of TEs.[64] In particular, LTR12C elements become transcriptionally active upon inhibition of epigenetic enzymes in multiple cancers.[64–66] TEs have been shown to contribute to the development of cancer by driving the expression of oncogenes through chimeric TE-gene transcripts.[29] Chimeric TE-transcripts also encode immunogenic antigens that are presented on the surface of cancer cells,[67] presenting a promising opportunity to exploit these neoantigens and TE-produced double-stranded RNA (dsRNA) in cancer immunotherapies (reviewed recently by Reid Cahn et al.[68] and Liang et al.[69]). TEs are also known to harbor features of active gene regulatory sites, including open chromatin, enhancer-specific histone marks, and TF binding in multiple cancers, such as breast,[70] acute myeloid leukemia (AML),[71] prostate,[72] colorectal, and hepatocellular[73] cancers (Figure 2C, right panel). This evidence indicates that TEs can play active regulatory roles in different cancer types.

Onco-exaptation of TEs can have both pro- and anti-tumor effects depending on the cancer type. TE transcription in cancers and tumor-adjacent normal tissues has been mostly associated with worse survival in, for example, hepatocellular,[74] head and neck,[75] breast,[76] and kidney[77] cancers, but there is also some evidence of positive correlation on survival in osteosarcoma[78] and melanoma.[79] Attig et al.[80] showed recently that onco-exaptation can also have pleiotropic effects: in lung cancer, calbindin (*CALB1*) expression driven by HERVH alternative promoter prevents senescence in early stages of cancer but suppresses protumor inflammation in later stages, leading to a survival disadvantage (Figure 2D). These results indicate that the effects of TE dysregulation must be considered in the context of cancer type and disease stage and that further studies with a larger scale and a higher resolution are needed, for example, by utilizing single-cell sequencing technologies, to elucidate the role of TE activity in different cancers.

ESCs, placenta, and cancer cells have striking similarities in their epigenetic states that also reflect their transcriptional, transpositional, and crTE activity. During development, mammalian genomes undergo genome-wide hypomethylation that causes widespread transient TE activity, which is very similar to the global hypomethylation and increased TE activity observed in cancer. In both cases, the observed TE activity emerges predominantly from the LTR elements.[41,44,81–83] TEs adapted to *cis*-regulatory functions during development are often also active in cancers. For example, the chimeric HERVH-*CALB1*

Transcriptional activation of TEs has been extensively studied, and it has been found that transcription can have positive, pleiotropic, and negative effects on cancer survival depending on the cancer type. TE transcription can also mimic viral immune responses in cancer cells, also presenting TE-derived antigens on cell surfaces.

transcript expressed in lung cancer[80] is also expressed in ESCs and healthy neurons, indicating a possible connection between developmental processes and onco-exaptation.[45,84] Moreover, the syncytin gene that is essential for the placenta has also been adapted to induce cell fusions in breast and endometrial cancers.[85]

TE activity in ESCs raises an interesting question about the role of TE re-activation in cancer stem cells (CSC). In a recent report, TEs were shown to be essential for maintaining stemness in AML,[86] which is a key feature for cancer dedifferentiation.[87] CSCs are cell populations within cancers that have features that resemble embryonic stem cells, such as unlimited potential for proliferation via common pathways (e.g., *Wnt*) and telomere repair mechanisms. They are drivers of key cancer processes, such as tumor establishment, growth, and relapse via resistant post-treatment cell populations. CSCs also exhibit many of the same surface markers that are present in ESCs.[88,89] It is largely elusive which cells CSCs are derived from, but there is recent evidence that oncogenesis in prostate cancer arises from the co-option of pluripotent stem cell (PSC) crTEs that promote growth via androgen receptor (AR) binding.[72] However, more studies are needed to determine the connection between TEs and CSCs.

Recently, the role of crTEs in both ESCs and cancers has been comprehensively reviewed by, for example, Fueyo et al.,[90] Grundy et al.,[24] Hermant and Torres-Padilla,[91] and Sundaram and Wysocka.[9] Cancer as a disease of development and its connections to ESCs and placenta has also been reviewed by for example Costanzo et al.[54] and Stanger and Wahl.[92] This review focuses on specific LTR subfamilies (LTR7, LTR10, LTR12C, and MER11), as these elements have been shown to be especially rich in their *cis*-regulatory sequence content[9] and have the most definitive evidence of *cis*-regulatory activity in both development and cancer. In addition, these TEs have links to pathways that may have roles in both development and malignancy. However, this list is not exhaustive, and there are more potential TEs presented in the aforementioned reviews and Figure 3.

LTR7/B/Y

The five subfamilies, LTR7, LTR7A, LTR7B, LTR7C, and LTR7Y, can either be integrated as solo insertions in the genome or flank their proviral HERVH-int sequence. LTR7 elements have been extensively studied in ESCs, and they have been found to be essential for maintaining the pluripotency of ESCs and induced pluripotent stem cells by directly binding multiple pluripotency factors, such as TFCP2L1, NANOG, SOX2, OCT4 (POU5F1), and KLF4[16,42,45,52,93–95] and participating in TAD formation.[96] In a functional quantitative ChIP-STARR-seq assay, LTR7 was shown to have the highest enrichment in the most high-activity class of enhancers in hESCs.[97] Pluripotency factors NANOG and KLF4 were found to be essential for driving *cis*-regulation and expression from these elements. Specific LTR7 subfamilies seem to control different stages of development, with LTR7B enriched in the eight-cell and morula stage, LTR7Y in the blastocyst stage, and LTR7 in the epiblast, although all variants do not seem to play a functional role.[46,98] By using a combination of transcriptome and ChIP-seq data for TF binding and histone modifications, these studies found that LTR7-derived transcription and active histone modifications, such as H3K27ac, H3K4me1, and H3K4me3 strongly correlate with the stage of embryonic development. LTR7s also protect embryonic cells from deleterious LINE transposition.[50] There is evidence suggesting that LTR7 elements can have dual roles as both enhancers and promoters, with some studies showing that LTR7-HERVH is a marker of pluripotency (reviewed by Sexton et al.[99]).

In cancers, the role LTR7 elements is still largely elusive. LTR7Y was found to be significantly overexpressed in six cancer types—colorectal, stomach, bladder, head and neck, lung, and liver cancers—correlating with the overall high rates of TE transcription and transposition in these cancer types.[5,61] The proviral HERVH sequence co-expressed strongly with LTR7Y in these cancer types, suggesting that these sequences are co-regulated or that the full endogenous retroviral sequence is expressed. LTR7Y is associated with open chromatin in colorectal and lung cancers, and all three LTR7 subfamilies showed moderate activity in a functional STARR-seq enhancer assay.[73,100] Interestingly, a knockout of tumor suppressor ARID1A in CRC led to the increased transcription of LTR7, LTR7Y, and the proviral HERVH-int, also stimulating the activity of transcriptional regulator BRD4. The upregulation of LTR7-HERVH led to the increased formation of BRD4 nuclear foci, upregulating BRD4-mediated transcription.[101] BRD4 is also an important regulator of MYC, possibly indicating a connection between the LTR7 activation and the BRD4-MYC axis in driving cancer progression.[102] HERVH expression correlates with lymph node invasion and microsatellite instable subtype of CRC,[103] but as a whole, only preliminary evidence of the activity of LTR7s and the proviral HERVH in cancer exists. Due to their key roles in development and the connection to MYC regulation, further studies are important for elucidating the role and function of LTR7-HERVH in tumorigenic processes.

LTR10A/F

The provirus of LTR10 subfamilies is known as HERVIP10. Some members of these subfamilies, such as LTR10B1, LTR10B2, LTR10C, LTR10D, and LTR10E, are known to be highly enriched for p53 binding sites.[20] More recently, the LTR10A subfamily has emerged as an especially important *cis*-regulatory element in both ESCs and cancers. It has been shown to contribute to the expression of important placental genes,[56] and a recent report suggests that LTR10 elements function as tumor-specific enhancers in CRC that drive tumor progression: Both Ivancevic et al.[100] and Karttunen et al.[73] showed that LTR10 subfamilies are relatively widely enriched in accessible chromatin regions in multiple epithelial cancers, such as colorectal, stomach, prostate, and lung tumors. Moreover, both Frost et al.[56] and Ivancevic et al.[100] showed that the LTR10A elements are essential in the regulation of the AP-1 pathway, and the important role this pathway plays in both placenta and cancers was noted by both independent studies. The LTR10A and LTR10F elements were reported to contain an internal highly mutable variable number tandem repeat (VNTR) region consisting of

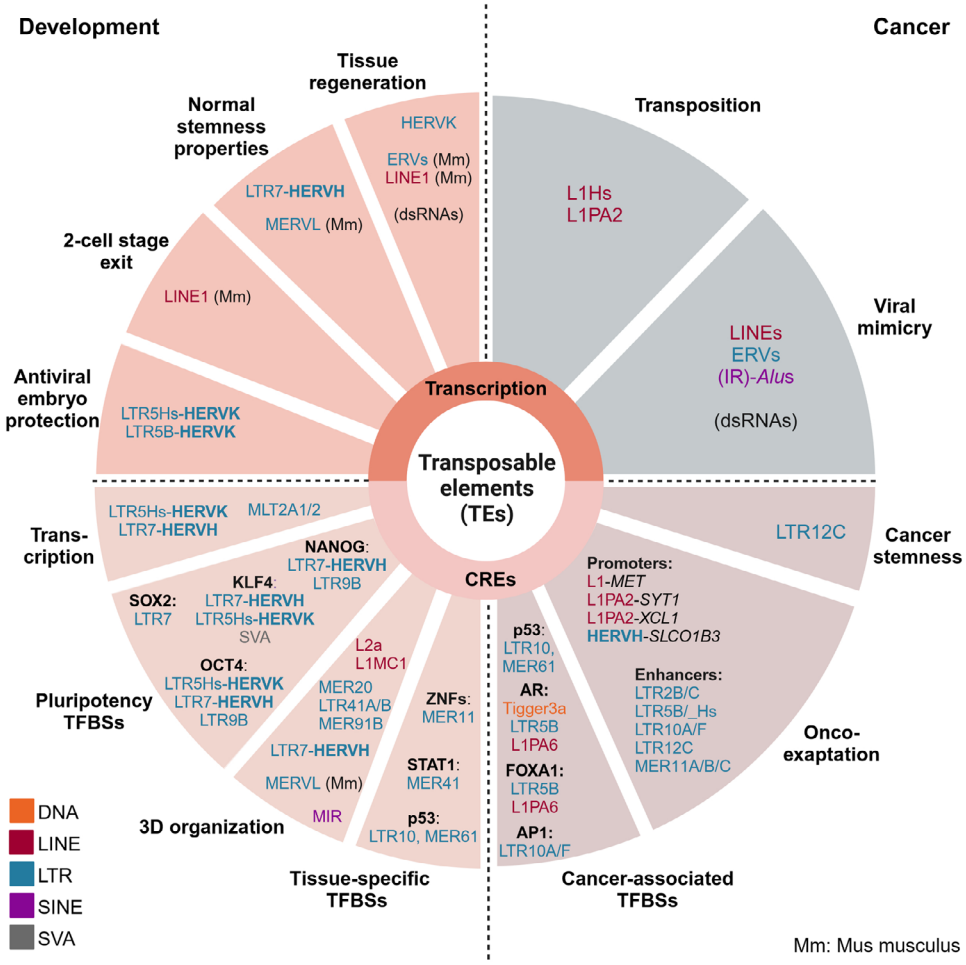

FIGURE 3 Summary of different functional properties of TEs in development and cancer.

28–30 bp repeats harboring AP-1 motif. The number of these repeats correlated with H3K27ac and FOSL1 binding and were found to have tumor-specific variations in CRC, possibly correlating with microsatellite instability-type tumors.[100] The AP-1 pathway contributes to cell motility and invasion,[104] suggesting that its activity can have large implications in both placental development and cancer. In conclusion, LTR10 is one of the most comprehensively studied crTE subfamilies in the context of cancer, and the AP-1 pathway has a close connection between the placenta and cancer via TE regulation. The VNTR region also suggests a novel mechanism of TE onco-exaptation via an expansion of TF motifs within their sequences, possibly due to microsatellite instability that occurs in some cancer types.

LTR12C

LTR12C is the long terminal repeat of the ERV9 family of endogenous retroviruses that was reported to have promoter activity since its discovery.[105] The LTR12C subfamily has been studied in multiple cellular contexts. Its transcription has been reported during embryonic development from two- to eight-cell stages and in the morula stage,[46,47] but its *cis*-regulatory activities and their significance in development are still mostly unclear. The LTR12-ERV9 has been exapted for *cis*-regulatory activities in normal cellular conditions, such as the promoter of a zinc finger gene *ZNF80*,[106] a male germ line-specific isoform of *TP63*,[107] and for long-range regulation of β-globin via a long non-coding RNA that stabilizes long-range LTR enhancer assemblies that regulate erythropoiesis.[108,109] The strong *cis*-regulatory functions of LTR12C elements are due to a strong TATA box and CCAAT boxes that can bind NFY-B, conferring BRD4-independent transcription and enhancer activity from the elements.[64,110,111] The LTR12C sequence also contains a highly variable tandem repeat region, which is possibly one of the mechanisms of sequence adaptation for the *cis*-regulatory activity of this subfamily.

LTR12C elements have been found to be strongly responsive to DNMT and HDAC inhibition that increases their expression and drives transcription of neopeptides that can cause an immunogenic response.[61,64–66,73] LTR12C has been reported to drive oncogene expression in multiple cancers,[29,112] to be essential for the maintenance of stemness,[86] and to have enhancer activity in a study focusing on AML.[71] LTR12C elements may function as oncogenic enhancers via NFYA activity since cancer type-specific binding of NFYA at LTR12C elements has been shown in liver cancer cells along with functional

enhancer activity and nascent RNA transcription (GRO-seq) signal.[73] However, LTR12C studies have largely focused on its promoter activity and its ability to drive transcription, and thus more data is needed to establish its distal enhancer activity.

MER11A/B/C/D

Another interesting TE group is the MER11, (medium reiteration frequency repetitive sequences) that are classified into MER11A, B, C, and D subfamilies. MER11 elements can be up to 1100 bp long, but their length varies depending on the number of 50 bp repeats.[113] As previously mentioned, expansions of VNTRs have amplified JUN/AP-1 motifs in LTR10A and LTR10F elements, suggesting that similar mechanism could have expanded TF binding sites also at MER11 elements, but this has not been confirmed in any studies so far to our knowledge. MER11 subfamilies, especially MER11A, B, and C, have been implicated as developmental enhancers with high activity during endodermal differentiation and in defining the segregation of inner cell mass and trophectoderm, and were found to contain binding sites for several pluripotency factors,[43,94,114] They also showed high activity in a functional enhancer assay in CRC cells, with the same enhancers also showing enhancer-like epigenetic features in developmental tissues.[73]

It has been suggested that MER11 elements may have a role in cancer via TFAP2A activity: the binding of TFAP2A at MER11 elements has been shown in colon cancer cells,[73] but whether it also binds to the same CREs during development has not been studied to our knowledge. However, TFAP2 family TFs have multiple roles in development: TFAP2C is a pioneer factor for pluripotency factor enhancers[115] and is considered a trophoblast marker.[116] TFAP2A and TFAP2C are essential in a core regulatory network for TGF-β-mediated EMT.[117] In CRC, TFAP2C is linked to stemness via upregulating stem cell factors, and its overexpression is correlated with chemotherapy resistance and poor prognosis.[118] However, both TFAP2A and TFAP2C have tumor suppressor and oncogene properties depending on the cancer type (reviewed in Kołat et al.[119]).

Interestingly, multiple copies of MER11D and LTR8B were found to be associated with a cluster of pregnancy-specific glycoprotein (*PSG*) genes that are expressed in the syncytiotrophoblast and are important markers of pregnancy.[56] TFAP2A can induce lung cancer EMT and metastasis by transactivating the *PSG9* gene that increases TGF-β signaling,[120] and TFAP2C was also shown to promote EMT in lung cancer in another study,[121] suggesting a potential connection between TE-regulated activity; however, neither of these reports studied the binding of TFAP2 to TEs nor whether the EMT promotion is mediated by TEs, so more studies are still needed.

OTHER TES AND POSSIBLE PATHWAYS

Grillo et al.[72] were the first to our knowledge to directly study the connections in TE activation between PSCs, normal prostate, and prostate cancer. Interestingly, they reported two subtypes of prostate cancer: the constant subtype had no enrichment of regulatory TEs, whereas in the reprogrammed subtype, 186 regulatory TEs were identified from two cohorts. Of the 186 identified TE subfamilies, 164 were found to be common with embryonic tissues but only 22 were common with the 97 subfamilies identified in benign prostate, highlighting the similarity of the developmental and malignant states. The most active subfamilies were Tigger3a, LTR5_Hs, and L1ME4b and these TEs were found to bind mostly lineage-specific TFs, such as the AR, which is an essential regulator of most prostate cancers. These results provide evidence of the connections between development and cancer and interesting leads for further studies. Other TE subfamilies with functional evidence of being active in different transcriptional and *cis*-regulatory roles in either developmental processes, cancer, or both, are summarized in Figure 3.

A summary of the specific TE subfamilies active in different transcriptional and *cis*-regulatory roles in development and cancer, including subfamilies that were not mentioned in the main text but were reported to have functional activity. All subfamilies are human-specific except for ones with only evidence in mice marked with Mm (Mus musculus). The major classes of TEs are marked in different colors, the proviral sequences of LTR elements are in bold, the gene names in chimeric TE-genes are in italics, and the TE-associated TFs in the relevant sections are also in bold. The figure is adapted from Grillo and Lupien.[122]

LIMITATIONS OF CURRENT APPROACHES

Studying TEs has been and still is hampered by their repetitive sequences that pose inherent difficulties for mapping them from short-read sequencing data. This is particularly challenging in the case of less divergent evolutionarily younger TE subfamilies, many of which seem to show high transpositional and *cis*-regulatory activity, for example, LTR5_Hs, L1Hs, and L1PA2.[123] Thus, the use of long-read sequencing methods for mapping transcription, methylation, and chromatin accessibility, such as NaNOMe-seq[31,124] and CELLO-seq[125], and improvements in mapping algorithms are important in uncovering the *cis*-regulatory activity emerging from TEs that are hard to map with short reads. However, short-read sequencing methods such as ChIP-seq, CUT&TAG, and ATAC-seq used to determine precise TF binding and chromatin accessibility are still essential and harder to replace with long-read sequencing.

Most of the studies on TE *cis*-regulatory activity thus far are based on biochemical measurements (e.g., RNA-seq, ChIP-seq, and ATAC-seq), which has been criticized as not being definitive evidence of functional *cis*-regulatory activity.[126,127] However, there is evidence that coordinated TE *cis*-regulation can be estimated solely from protein-coding gene expression and is consistently correlated with chromatin states.[128] In total, about 4% of the whole human genome is under evolutionary constraint,[129] which is a relatively reliable measure of biologically relevant functional activity. However, a comparative genomics approach may undervalue TEs that have recently expanded

and thus have no evident sequence constraint but have nevertheless been co-opted and are functionally active by using biochemical measures.[130]

Due to the uncertainties of determining *cis*-regulatory activity with biochemical measurements, large-scale functional genome screening methods are important for complementing them. Targeting TE sequences to directly assess their effect on gene expression is paramount, as enhancer redundancy is common, and some studies have also found little effects stemming from TEs.[131,132] The development of, for example, CRISPR-Cas9-based technologies for CRE deletion and silencing using CRISPR interference has been rapid, with most of the newer studies mentioned in this review utilizing them to support biochemical measurements, but higher throughput single-cell and locus-level methods, such as CROP-seq or Perturb-seq,[133,134] are also needed to achieve a more comprehensive view of TE activity, as individual copies of TEs may be active in different cellular contexts. Thus, we predict that in the coming years, we will see the understanding of TEs in cancer to take large leaps forward utilizing these new techniques and increasing amount of large-scale data that is available for public use.

CONCLUSIONS

Overall, this article provides evidence of the intriguing links between *cis*-regulatory activity of TEs in development and cancer. The evidence is still preliminary given the complexity and redundancy of the non-coding genome and the size and repetitiveness of the TE repertoire, covering almost half of the human genome. Thus, a better understanding of *cis*-regulatory activity of TEs in both development and cancer is warranted. The unsolved question is whether the *cis*-regulatory activity of TEs is a by-product of epigenetic dysregulation or whether TEs have an active role as cancer drivers stemming from their *cis*-regulatory activities. For active regulatory roles, TEs could be stochastically activated or co-opted after mutations or escape from epigenetic control due to natural aging and environmental factors. Another pertinent question is the heterogeneity in TE activity in different cancers, within cancer subtypes, in different tumor cell populations, and between patients. A better understanding of TE dysregulation in cancers could become useful in the development of epigenetic therapies targeting TEs to complement immunotherapies, as many ongoing clinical trials are already studying the efficacy of combined epigenetic- and immuno-therapy in different cancer types.[135] Thus, more research is needed to explain not only the normal physiological functions of TEs but also whether deviation from or exploitation of the normal TE-regulated developmental programs are consequential for the initiation, development, and progression of cancer.

AUTHOR CONTRIBUTIONS

The manuscript was conceptualized, written, and edited by Konsta Karttunen, Divyesh Patel, and Biswajyoti Sahu. Illustrations were created by KK with assistance from Biswajyoti Sahu and Divyesh Patel.

ACKNOWLEDGMENTS

This study was supported by grants to Biswajyoti Sahu from the Academy of Finland (317807, 320114, 346065), iCAN Digital Precision Cancer Medicine Flagship (320185), Cancer Foundation Finland, Sigrid Jusélius Foundation, Jane and Aatos Erkko Foundation, and Centre for Molecular Medicine Norway, University of Oslo. All figures were created in Biorender.com.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Konsta Karttunen https://orcid.org/0000-0003-3511-7156
Divyesh Patel https://orcid.org/0000-0003-4005-8429
Biswajyoti Sahu https://orcid.org/0000-0001-6576-5440

REFERENCES

1. McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America*, *36*(6), 344. https://doi.org/10.1073/pnas.36.6.344
2. Tubio, J. M. C., Li, Y., Ju, Y. S., Martincorena, I., Cooke, S. L., Tojo, M., Gundem, G., Pipinikas, C. P., Zamora, J., Raine, K., Menzies, A., Roman-Garcia, P., Fullam, A., Gerstung, M., Shlien, A., Tarpey, P. S., Papaemmanuil, E., Knappskog, S., Van Loo, P., … Campbell, P. J. (2014). Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science*, *345*(6196), 1251343. https://doi.org/10.1126/science.1251343
3. Hoyt, S. J., Storer, J. M., Hartley, G. A., Grady, P. G. S., Gershman, A., de Lima, L. G., Limouse, C., Halabian, R., Wojenski, L., Rodriguez, M., Altemose, N., Rhie, A., Core, L. J., Gerton, J. L., Makalowski, W., Olson, D., Rosen, J., Smit, A. F. A., Straight, A. F., … O'Neill, R. J. (2022). From telomere to telomere: The transcriptional and epigenetic state of human repeat elements. *Science*, *376*(6588), eabk3112. https://doi.org/10.1126/science.abk3112
4. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Patrinos, A., & Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. Article 6822. https://doi.org/10.1038/35057062
5. Rodriguez-Martin, B., Alvarez, E. G., Baez-Ortega, A., Zamora, J., Supek, F., Demeulemeester, J., Santamarina, M., Ju, Y. S., Temes, J., Garcia-Souto, D., Campbell, P. J., & Tubio, J. M. C. (2020). Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nature Genetics*, *52*(3), 306–319. Article 3. https://doi.org/10.1038/s41588-019-0562-0
6. Wells, J. N., & Feschotte, C. (2020). A field guide to eukaryotic transposable elements. *Annual Review of Genetics*, *54*, 539–561. https://doi.org/10.1146/annurev-genet-040620-022145
7. Britten, R. J., & Davidson, E. H. (1969). Gene regulation for higher cells: A theory. *Science*, *165*(3891), 349–357. https://doi.org/10.1126/science.165.3891.349
8. McClintock, B. (1956). Controlling elements and the gene. *Cold Spring Harbor Symposia on Quantitative Biology*, *21*, 197–216. https://doi.org/10.1101/sqb.1956.021.01.017
9. Sundaram, V., & Wysocka, J. (2020). Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. *Philosophical Transactions of the Royal Society B: Biological*

Sciences, *375*(1795), 20190347. https://doi.org/10.1098/rstb.2019.0347

10. Sundaram, V., & Wang, T. (2018). Transposable element mediated innovation in gene regulatory landscapes of cells: Re-visiting the "gene-battery" model. *BioEssays*, *40*(1), 1700155. https://doi.org/10.1002/bies.201700155
11. Marnetto, D., Mantica, F., Molineris, I., Grassi, E., Pesando, I., & Provero, P. (2018). Evolutionary rewiring of human regulatory networks by waves of genome expansion. *American Journal of Human Genetics*, *102*(2), 207–218. https://doi.org/10.1016/j.ajhg.2017.12.014
12. Andrews, G., Fan, K., Pratt, H. E., Phalke, N., ZOONOMIA, C., Karlsson, E. K., Lindblad-Toh, K., Gazal, S., Moore, J. E., & Weng, Z. (2023). Mammalian evolution of human cis-regulatory elements and transcription factor binding sites. *Science*, *380*(6643), eabn7930. https://doi.org/10.1126/science.abn7930
13. Jacques, P.-É., Jeyakani, J., & Bourque, G. (2013). The majority of primate-specific regulatory sequences are derived from transposable elements. *PLOS Genetics*, *9*(5), e1003504. https://doi.org/10.1371/journal.pgen.1003504
14. Trizzino, M., Park, Y., Holsbach-Beltrame, M., Aracena, K., Mika, K., Caliskan, M., Perry, G. H., Lynch, V. J., & Brown, C. D. (2017). Transposable elements are the primary source of novelty in primate gene regulation. *Genome Research*, *27*, 1623–1633. https://doi.org/10.1101/gr.218149.116
15. Pehrsson, E. C., Choudhary, M. N. K., Sundaram, V., & Wang, T. (2019). The epigenomic landscape of transposable elements across normal human development and anatomy. *Nature Communications*, *10*(1), 5640. https://doi.org/10.1038/s41467-019-13555-x
16. Kunarso, G., Chia, N.-Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.-S., Ng, H.-H., & Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature Genetics*, *42*(7), 631–634. Article 7. https://doi.org/10.1038/ng.600
17. Chuong, E. B., Rumi, M. A. K., Soares, M. J., & Baker, J. C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nature Genetics*, *45*(3), 325–329. https://doi.org/10.1038/ng.2553
18. Lynch, V. J., Leclerc, R. D., May, G., & Wagner, G. P. (2011). Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nature Genetics*, *43*(11), 1154–1159. Article 11. https://doi.org/10.1038/ng.917
19. Dupressoir, A., Lavialle, C., & Heidmann, T. (2012). From ancestral infectious retroviruses to bona fide cellular genes: Role of the captured syncytins in placentation. *Placenta*, *33*(9), 663–671. https://doi.org/10.1016/j.placenta.2012.05.005
20. Wang, T., Zeng, J., Lowe, C. B., Sellers, R. G., Salama, S. R., Yang, M., Burgess, S. M., Brachmann, R. K., & Haussler, D. (2007). Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proceedings of the National Academy of Sciences*, *104*(47), 18613–18618. Article 47. https://doi.org/10.1073/pnas.0703637104
21. Chuong, E. B., Elde, N. C., & Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*, *351*(6277), 1083–1087. https://doi.org/10.1126/science.aad5497
22. Almeida, M. V., Vernaz, G., Putman, A. L. K., & Miska, E. A. (2022). Taming transposable elements in vertebrates: From epigenetic silencing to domestication. *Trends in Genetics*, *38*(6), 529–553. https://doi.org/10.1016/j.tig.2022.02.009
23. Li, X., Bie, L., Wang, Y., Hong, Y., Zhou, Z., Fan, Y., Yan, X., Tao, Y., Huang, C., Zhang, Y., Sun, X., Li, J. X. H., Zhang, J., Chang, Z., Xi, Q., Meng, A., Shen, X., Xie, W., & Liu, N. (2024). LINE-1 transcription activates long-range gene expression. *Nature Genetics*, *56*, 1494–1502. https://doi.org/10.1038/s41588-024-01789-5
24. Grundy, E. E., Diab, N., & Chiappinelli, K. B. (2022). Transposable element regulation and expression in cancer. *The FEBS Journal*, *289*(5), 1160–1179. https://doi.org/10.1111/febs.15722
25. Radrizzani, S., Kudla, G., Izsvák, Z., & Hurst, L. D. (2024). Selection on synonymous sites: The unwanted transcript hypothesis. *Nature Reviews Genetics*, *25*, 431–448. https://doi.org/10.1038/s41576-023-00686-7
26. Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K. W., Vogelstein, B., & Nakamura, Y. (1992). Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Research*, *52*(3), 643–645.
27. Cajuso, T., Sulo, P., Tanskanen, T., Katainen, R., Taira, A., Hänninen, U. A., Kondelin, J., Forsström, L., Välimäki, N., Aavikko, M., Kaasinen, E., Ristimäki, A., Koskensalo, S., Lepistö, A., Renkonen-Sinisalo, L., Seppälä, T., Kuopio, T., Böhm, J., Mecklin, J.-P., … Aaltonen, L. A. (2019). Retrotransposon insertions can initiate colorectal cancer and are associated with poor survival. *Nature Communications*, *10*(1), 4022. https://doi.org/10.1038/s41467-019-11770-0
28. Scott, E. C., Gardner, E. J., Masood, A., Chuang, N. T., Vertino, P. M., & Devine, S. E. (2016). A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Research*, *26*(6), 745–755. Article 6. https://doi.org/10.1101/gr.201814.115
29. Jang, H. S., Shah, N. M., Du, A. Y., Dailey, Z. Z., Pehrsson, E. C., Godoy, P. M., Zhang, D., Li, D., Xing, X., Kim, S., O'Donnell, D., Gordon, J. I., & Wang, T. (2019). Transposable elements drive widespread expression of oncogenes in human cancers. *Nature Genetics*, *51*(4), 611–617. https://doi.org/10.1038/s41588-019-0373-3
30. Baggiolini, A., Callahan, S. J., Montal, E., Weiss, J. M., Trieu, T., Tagore, M. M., Tischfield, S. E., Walsh, R. M., Suresh, S., Fan, Y., Campbell, N. R., Perlee, S. C., Saurat, N., Hunter, M. V., Simon-Vermot, T., Huang, T.-H., Ma, Y., Hollmann, T., Tickoo, S. K., … White, R. M. (2021). Developmental chromatin programs determine oncogenic competence in melanoma. *Science*, *373*(6559), eabc1048. https://doi.org/10.1126/science.abc1048
31. Sahu, B., Pihlajamaa, P., Zhang, K., Palin, K., Ahonen, S., Cervera, A., Ristimäki, A., Aaltonen, L. A., Hautaniemi, S., & Taipale, J. (2021). Human cell transformation by combined lineage conversion and oncogene expression. *Oncogene*, *40*(36), 5533–5547. https://doi.org/10.1038/s41388-021-01940-0
32. Pomerantz, M. M., Qiu, X., Zhu, Y., Takeda, D. Y., Pan, W., Baca, S. C., Gusev, A., Korthauer, K. D., Severson, T. M., Ha, G., Viswanathan, S. R., Seo, J.-H., Nguyen, H. M., Zhang, B., Pasaniuc, B., Giambartolomei, C., Alaiwi, S. A., Bell, C. A., O'Connor, E. P., … Freedman, M. L. (2020). Prostate cancer reactivates developmental epigenomic programs during metastatic progression. *Nature Genetics*, *52*(8), 790–799. Article 8. https://doi.org/10.1038/s41588-020-0664-8
33. Roe, J.-S., Hwang, C.-I., Somerville, T. D. D., Milazzo, J. P., Lee, E. J., Silva, B. D., Maiorino, L., Tiriac, H., Young, C. M., Miyabayashi, K., Filippini, D., Creighton, B., Burkhart, R. A., Buscaglia, J. M., Kim, E. J., Grem, J. L., Lazenby, A. J., Grunkemeyer, J. A., Hollingsworth, M. A., … Vakoc, C. R. (2017). Enhancer reprogramming promotes pancreatic cancer metastasis. *Cell*, *170*(5), 875–888. e20.e20. https://doi.org/10.1016/j.cell.2017.07.007
34. Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discovery*, *12*(1), 31–46. https://doi.org/10.1158/2159-8290.CD-21-1059
35. Lynch-Sutherland, C. F., Chatterjee, A., Stockwell, P. A., Eccles, M. R., & Macaulay, E. C. (2020). Reawakening the developmental origins of cancer through transposable elements. *Frontiers in Oncology*, *10*, 468. https://www.frontiersin.org/articles/10.3389/fonc.2020.00468
36. Senft, A. D., & Macfarlan, T. S. (2021). Transposable elements shape the evolution of mammalian development. *Nature Reviews Genetics*, *22*(11), 691–711. Article 11. https://doi.org/10.1038/s41576-021-00385-1

37. Biechele, S., Lin, C.-J., Rinaudo, P. F., & Ramalho-Santos, M. (2015). Unwind and transcribe: Chromatin reprogramming in the early mammalian embryo. *Current Opinion in Genetics & Development*, *34*, 17–23. https://doi.org/10.1016/j.gde.2015.06.003
38. Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M. P., & Wang, T. (2014). Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Research*, *24*(12), 1963–1976. https://doi.org/10.1101/gr.168872.113
39. Xie, D., Chen, C.-C., Ptaszek, L. M., Xiao, S., Cao, X., Fang, F., Ng, H. H., Lewin, H. A., Cowan, C., & Zhong, S. (2010). Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome Research*, *20*(6), 804–815. https://doi.org/10.1101/gr.100594.109
40. Bourque, G., Leong, B., Vega, V. B., Chen, X., Lee, Y. L., Srinivasan, K. G., Chew, J.-L., Ruan, Y., Wei, C.-L., Ng, H. H., & Liu, E. T. (2008). Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Research*, *18*(11), 1752–1762. https://doi.org/10.1101/gr.080663.108
41. Xie, M., Hong, C., Zhang, B., Lowdon, R. F., Xing, X., Li, D., Zhou, X., Lee, H. J., Maire, C. L., Ligon, K. L., Gascard, P., Sigaroudinia, M., Tlsty, T. D., Kadlecek, T., Weiss, A., O'Geen, H., Farnham, P. J., Madden, P. A. F., Mungall, A. J., … Wang, T. (2013). DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nature Genetics*, *45*(7), 836–841. Article 7. https://doi.org/10.1038/ng.2649
42. Pontis, J., Planet, E., Offner, S., Turelli, P., Duc, J., Coudray, A., Theunissen, T. W., Jaenisch, R., & Trono, D. (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.e5.e5. https://doi.org/10.1016/j.stem.2019.03.012
43. Pontis, J., Pulver, C., Playfoot, C. J., Planet, E., Grun, D., Offner, S., Duc, J., Manfrin, A., Lutolf, M. P., & Trono, D. (2022). Primate-specific transposable elements shape transcriptional networks during human development. *Nature Communications*, *13*(1), 7178. https://doi.org/10.1038/s41467-022-34800-w
44. 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., Pera, R., & Wysocka, J. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*, *522*(7555), 221–225. Article 7555. https://doi.org/10.1038/nature14308
45. Wang, J., Xie, G., Singh, M., Ghanbarian, A. T., Raskó, T., Szvetnik, A., Cai, H., Besser, D., Prigione, A., Fuchs, N. V., Schumann, G. G., Chen, W., Lorincz, M. C., Ivics, Z., Hurst, L. D., & Izsvák, Z. (2014). Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature*, *516*(7531), 405–409. Article 7531. https://doi.org/10.1038/nature13804
46. Göke, J., Lu, X., Chan, Y.-S., Ng, H.-H., Ly, L.-H., Sachs, F., & Szczerbinska, I. (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. https://doi.org/10.1016/j.stem.2015.01.005
47. Liu, L., Leng, L., Liu, C., Lu, C., Yuan, Y., Wu, L., Gong, F., Zhang, S., Wei, X., Wang, M., Zhao, L., Hu, L., Wang, J., Yang, H., Zhu, S., Chen, F., Lu, G., Shang, Z., & Lin, G. (2019). An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nature Communications*, *10*(1), 364. https://doi.org/10.1038/s41467-018-08244-0
48. Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G. A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X., & Ramalho-Santos, M. (2018). A LINE1-nucleolin partnership regulates early development and ESC identity. *Cell*, *174*(2), 391–405.e19.e19. https://doi.org/10.1016/j.cell.2018.05.043
49. Sakashita, A., Kitano, T., Ishizu, H., Guo, Y., Masuda, H., Ariura, M., Murano, K., & Siomi, H. (2023). Transcription of MERVL retrotransposons is required for preimplantation embryo development. *Nature Genetics*, *55*(3), 484–495. https://doi.org/10.1038/s41588-023-01324-y
50. Singh, M., Kondrashkina, A. M., Widmann, T. J., Cortes, J. L., Bansal, V., Wang, J., Römer, C., Garcia-Canadas, M., Garcia-Perez, J. L., Hurst, L. D., & Izsvák, Z. (2023). A new human embryonic cell type associated with activity of young transposable elements allows definition of the inner cell mass. *PLoS Biology*, *21*(6), e3002162. https://doi.org/10.1371/journal.pbio.3002162
51. Lu, X., Sachs, F., Ramsay, L., Jacques, P.-É., Göke, J., Bourque, G., & Ng, H.-H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature Structural & Molecular Biology*, *21*(4), 423–425. https://doi.org/10.1038/nsmb.2799
52. Ohnuki, M., Tanabe, K., Sutou, K., Teramoto, I., Sawamura, Y., Narita, M., Nakamura, M., Tokunaga, Y., Nakamura, M., Watanabe, A., Yamanaka, S., & Takahashi, K. (2014). Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(34), 12426–12431. https://doi.org/10.1073/pnas.1413299111
53. Ferretti, C., Bruni, L., Dangles-Marie, V., Pecking, A. P., & Bellet, D. (2007). Molecular circuits shared by placental and cancer cells, and their implications in the proliferative, invasive and migratory capacities of trophoblasts. *Human Reproduction Update*, *13*(2), 121–141. https://doi.org/10.1093/humupd/dml048
54. Costanzo, V., Bardelli, A., Siena, S., & Abrignani, S. (2018). Exploring the links between cancer and placenta development. *Open Biology*, *8*(6), 180081. https://doi.org/10.1098/rsob.180081
55. Kshitiz Afzal, J., Maziarz, J. D., Hamidzadeh, A., Liang, C., Erkenbrack, E. M., Kim, H. N., Haeger, J.-D., Pfarrer, C., Hoang, T., Ott, T., Spencer, T., Pavličev, M., Antczak, D. F., Levchenko, A., & Wagner, G. P. (2019). Evolution of placental invasion and cancer metastasis are causally linked. *Nature Ecology & Evolution*, *3*(12), 1743–1753. https://doi.org/10.1038/s41559-019-1046-4
56. Frost, J. M., Amante, S. M., Okae, H., Jones, E. M., Ashley, B., Lewis, R. M., Cleal, J. K., Caley, M. P., Arima, T., Maffucci, T., & Branco, M. R. (2023). Regulation of human trophoblast gene expression by endogenous retroviruses. *Nature Structural & Molecular Biology*, *30*(4), 527–538. https://doi.org/10.1038/s41594-023-00960-6
57. Sun, M., Wolf, G., Wang, Y., Senft, A. D., Ralls, S., Jin, J., Dunn-Fletcher, C. E., Muglia, L. J., & Macfarlan, T. S. (2021). Endogenous retroviruses drive lineage-specific regulatory evolution across primate and rodent placentae. *Molecular Biology and Evolution*, *38*(11), 4992–5004. https://doi.org/10.1093/molbev/msab223
58. Yu, M., Hu, X., Pan, Z., Du, C., Jiang, J., Zheng, W., Cai, H., Wang, Y., Deng, W., Wang, H., Lu, J., Sun, M., & Cao, B. (2023). Endogenous retrovirus-derived enhancers confer the transcriptional regulation of human trophoblast syncytialization. *Nucleic Acids Research*, *51*(10), 4745–4759. https://doi.org/10.1093/nar/gkad109
59. Brouha, B., Schustak, J., Badge, R. M., Lutz-Prigge, S., Farley, A. H., Moran, J. V., & Kazazian, H. H. (2003). Hot L1s account for the bulk of retrotransposition in the human population. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5280–5285. https://doi.org/10.1073/pnas.0831042100
60. Chiappinelli, K. B., Strissel, P. L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N. S., Cope, L. M., Snyder, A., Makarov, V., Budhu, S., Slamon, D. J., Wolchok, J. D., Pardoll, D. M., Beckmann, M. W., Zahnow, C. A., Merghoub, T., Chan, T. A., … Strick, R. (2015). Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell*, *162*(5), 974–986. https://doi.org/10.1016/j.cell.2015.07.011
61. Kong, Y., Rose, C. M., Cass, A. A., Williams, A. G., Darwish, M., Lianoglou, S., Haverty, P. M., Tong, A.-J., Blanchette, C., Albert, M. L., Mellman, I., Bourgon, R., Greally, J., Jhunjhunwala, S., & Chen-Harris,

H. (2019). Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nature Communications*, *10*(1), 5228. https://doi.org/10.1038/s41467-019-13035-2

62. Nam, C. H., Youk, J., Kim, J. Y., Lim, J., Park, J. W., Oh, S. A., Lee, H. J., Park, J. W., Won, H., Lee, Y., Jeong, S.-Y., Lee, D.-S., Oh, J. W., Han, J., Lee, J., Kwon, H. W., Kim, M. J., & Ju, Y. S. (2023). Widespread somatic L1 retrotransposition in normal colorectal epithelium. *Nature*, *617*(7961), 540–547. https://doi.org/10.1038/s41586-023-06046-z
63. Ewing, A. D., Gacita, A., Wood, L. D., Ma, F., Xing, D., Kim, M.-S., Manda, S. S., Abril, G., Pereira, G., Makohon-Moore, A., Looijenga, L. H. J., Gillis, A. J. M., Hruban, R. H., Anders, R. A., Romans, K. E., Pandey, A., Iacobuzio-Donahue, C. A., Vogelstein, B., Kinzler, K. W., … Solyom, S. (2015). Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. *Genome Research*, *25*(10), 1536–1545. https://doi.org/10.1101/gr.196238.115
64. Brocks, D., Schmidt, C. R., Daskalakis, M., Jang, H. S., Shah, N. M., Li, D., Li, J., Zhang, B., Hou, Y., Laudato, S., Lipka, D. B., Schott, J., Bierhoff, H., Assenov, Y., Helf, M., Ressnerova, A., Islam, M. S., Lindroth, A. M., Haas, S., … Plass, C. (2017). DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nature Genetics*, *49*(7), 1052–1060. https://doi.org/10.1038/ng.3889
65. Goyal, A., Bauer, J., Hey, J., Papageorgiou, D. N., Stepanova, E., Daskalakis, M., Scheid, J., Dubbelaar, M., Klimovich, B., Schwarz, D., Märklin, M., Roerden, M., Lin, Y.-Y., Ma, T., Mücke, O., Rammensee, H.-G., Lübbert, M., Loayza-Puch, F., Krijgsveld, J., … Plass, C. (2023). DNMT and HDAC inhibition induces immunogenic neoantigens from human endogenous retroviral element-derived transcripts. *Nature Communications*, *14*(1), 6731. https://doi.org/10.1038/s41467-023-42417-w
66. Patel, D., Tiusanen, V., Pihlajamaa, P., & Sahu, B. (2024). *Cancer cell type-specific derepression of transposable elements by inhibition of chromatin modifier enzymes*. BioRxiv, 1–45, https://doi.org/10.1101/2024.01.15.575744
67. Shah, N. M., Jang, H. J., Liang, Y., Maeng, J. H., Tzeng, S.-C., Wu, A., Basri, N. L., Qu, X., Fan, C., Li, A., Katz, B., Li, D., Xing, X., Evans, B. S., & Wang, T. (2023). Pan-cancer analysis identifies tumor-specific antigens derived from transposable elements. *Nature Genetics*, *55*(4), 631–639. Article 4. https://doi.org/10.1038/s41588-023-01349-3
68. Reid Cahn, A., Bhardwaj, N., & Vabret, N. (2022). Dark genome, bright ideas: Recent approaches to harness transposable elements in immunotherapies. *Cancer Cell*, *40*(8), 792–797. https://doi.org/10.1016/j.ccell.2022.07.003
69. Liang, Y., Qu, X., Shah, N. M., & Wang, T. (2024). Towards targeting transposable elements for cancer therapy. *Nature Reviews Cancer*, *24*(2), 123–140. Article 2. https://doi.org/10.1038/s41568-023-00653-8
70. Jiang, J.-C., & Upton, K. R. (2019). Human transposons are an abundant supply of transcription factor binding sites and promoter activities in breast cancer cell lines. *Mobile DNA*, *10*, 16. https://doi.org/10.1186/s13100-019-0158-3
71. Deniz, Ö., Ahmed, M., Todd, C. D., Rio-Machin, A., Dawson, M. A., & Branco, M. R. (2020). Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nature Communications*, *11*(1), 3506. https://doi.org/10.1038/s41467-020-17206-4
72. Grillo, G., Keshavarzian, T., Linder, S., Arlidge, C., Mout, L., Nand, A., Teng, M., Qamra, A., Zhou, S., Kron, K. J., Murison, A., Hawley, J. R., Fraser, M., van der Kwast, T. H., Raj, G. V., He, H. H., Zwart, W., & Lupien, M. (2023). Transposable elements are co-opted as oncogenic regulatory elements by lineage-specific transcription factors in prostate cancer. *Cancer Discovery*, *13*(11), 2470–2487. https://doi.org/10.1158/2159-8290.Cd-23-0331
73. Karttunen, K., Patel, D., Xia, J., Fei, L., Palin, K., Aaltonen, L., & Sahu, B. (2023). Transposable elements as tissue-specific enhancers in cancers of endodermal lineage. *Nature Communications*, *14*(1), 5313. https://doi.org/10.1038/s41467-023-41081-4
74. Zhou, Y., Liu, L., Liu, Y., Zhou, P., Yan, Q., Yu, H., Chen, X., & Zhu, F. (2021). Implication of human endogenous retrovirus W family envelope in hepatocellular carcinoma promotes MEK/ERK-mediated metastatic invasiveness and doxorubicin resistance. *Cell Death Discovery*, *7*(1), 1–14. https://doi.org/10.1038/s41420-021-00562-5
75. Kolbe, A. R., Bendall, M. L., Pearson, A. T., Paul, D., Nixon, D. F., Pérez-Losada, M., & Crandall, K. A. (2020). Human endogenous retrovirus expression is associated with head and neck cancer and differential survival. *Viruses*, *12*(9), 956. Article 9. https://doi.org/10.3390/v12090956
76. Johanning, G. L., Malouf, G. G., Zheng, X., Esteva, F. J., Weinstein, J. N., Wang-Johanning, F., & Su, X. (2017). Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype. *Scientific Reports*, *7*(1), 41960. https://doi.org/10.1038/srep41960
77. Zapatka, M., Borozan, I., Brewer, D. S., Iskar, M., Grundhoff, A., Alawi, M., Desai, N., Sültmann, H., Moch, H., Cooper, C. S., Eils, R., Ferretti, V., & Lichter, P. (2020). The landscape of viral associations in human cancers. *Nature Genetics*, *52*(3), 320–330. https://doi.org/10.1038/s41588-019-0558-9
78. Wang, C., & Liang, C. (2022). The insertion and dysregulation of transposable elements in osteosarcoma and their association with patient event-free survival. *Scientific Reports*, *12*(1), 377. https://doi.org/10.1038/s41598-021-04208-5
79. Singh, M., Cai, H., Bunse, M., Feschotte, C., & Izsvák, 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), 1303. Article 11. https://doi.org/10.3390/v12111303
80. Attig, J., Pape, J., Doglio, L., Kazachenka, A., Ottina, E., Young, G. R., Enfield, K. S. S., Aramburu, I. V., Ng, K. W., Faulkner, N., Bolland, W., Papayannopoulos, V., Swanton, C., & Kassiotis, G. (2023). Human endogenous retrovirus onco-exaptation counters cancer cell senescence through calbindin. *The Journal of Clinical Investigation*, *133*(14), e164397. https://doi.org/10.1172/JCI164397
81. Smith, Z. D., Shi, J., Gu, H., Donaghey, J., Clement, K., Cacchiarelli, D., Gnirke, A., Michor, F., & Meissner, A. (2017). Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature*, *549*(7673), 543–547. https://doi.org/10.1038/nature23891
82. Kanholm, T., Rentia, U., Hadley, M., Karlow, J. A., Cox, O. L., Diab, N., Bendall, M. L., Dawson, T., McDonald, J. I., Xie, W., Crandall, K. A., Burns, K. H., Baylin, S. B., Easwaran, H., & Chiappinelli, K. B. (2023). Oncogenic transformation drives DNA methylation loss and transcriptional activation at transposable element loci. *Cancer Research*, *83*(15), 2584–2599. https://doi.org/10.1158/0008-5472.CAN-22-3485
83. Gomez, N. C., Hepperla, A. J., Dumitru, R., Simon, J. M., Fang, F., & Davis, I. J. (2016). Widespread chromatin accessibility at repetitive elements links stem cells with human cancer. *Cell Reports*, *17*(6), 1607–1620. https://doi.org/10.1016/j.celrep.2016.10.011
84. Singh, M., Kondraskhina, A. M., Hurst, L. D., & Izsvák, Z. (2023). Staring at the onco-exaptation: The two-faced medley of an ancient retrovirus. *HERVH. The Journal of Clinical Investigation*, *133*(14), 1–3. https://doi.org/10.1172/JCI172278
85. Bjerregaard, B., Holck, S., Christensen, I. J., & Larsson, L.-I. (2006). Syncytin is involved in breast cancer-endothelial cell fusions. *Cellular and Molecular Life Sciences CMLS*, *63*(16), 1906–1911. https://doi.org/10.1007/s00018-006-6201-9
86. Grillo, G., Nadorp, B., Qamra, A., Mitchell, A., Arlidge, C., Nand, A., Takayama, N., Murison, A., Tonekaboni, S. A. M., Arruda, A., Wang, J., Minden, M., Deniz, O., Boutzen, H., Dick, J., & Lupien, M. (2024). *Trans-*

posable elements shape stemness in normal and leukemic hematopoiesis. bioRxiv, 1–30. https://doi.org/10.1101/2021.02.16.431334

87. Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., Kamińska, B., Huelsken, J., Omberg, L., Gevaert, O., Colaprico, A., Czerwińska, P., Mazurek, S., Mishra, L., Heyn, H., Krasnitz, A., Godwin, A. K., Lazar, A. J., Caesar-Johnson, S. J., … Wiznerowicz, M. (2018). Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell*, *173*(2), 338–354.e15.e15. https://doi.org/10.1016/j.cell.2018.03.034
88. Kim, W.-T., & Ryu, C. J. (2017). Cancer stem cell surface markers on normal stem cells. *BMB Reports*, *50*(6), 285–298. https://doi.org/10.5483/BMBRep.2017.50.6.039
89. Bisht, S., Nigam, M., Kunjwal, S. S., Sergey, P., Mishra, A. P., & Sharifi-Rad, J. (2022). Cancer stem cells: From an insight into the basics to recent advances and therapeutic targeting. *Stem Cells International*, *2022*, 9653244. https://doi.org/10.1155/2022/9653244
90. Fueyo, R., Judd, J., Feschotte, C., & Wysocka, J. (2022). Roles of transposable elements in the regulation of mammalian transcription. *Nature Reviews Molecular Cell Biology*, *23*(7), 481–497. Article 7. https://doi.org/10.1038/s41580-022-00457-y
91. Hermant, C., & Torres-Padilla, M.-E. (2021). TFs for TEs: The transcription factor repertoire of mammalian transposable elements. *Genes & Development*, *35*(1–2), 22–39. https://doi.org/10.1101/gad.344473.120
92. Stanger, B. Z., & Wahl, G. M. (2024). Cancer as a disease of development gone awry. *Annual Review of Pathology: Mechanisms of Disease*, *19*(1), 397–421. https://doi.org/10.1146/annurev-pathmechdis-031621-025610
93. Fort, A., Hashimoto, K., Yamada, D., Salimullah, M., Keya, C. A., Saxena, A., Bonetti, A., Voineagu, I., Bertin, N., Kratz, A., Noro, Y., Wong, C.-H., de Hoon, M., Andersson, R., Sandelin, A., Suzuki, H., Wei, C.-L., Koseki, H., Hasegawa, Y., … Carninci, P. (2014). Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance. *Nature Genetics*, *46*(6), 558–566. Article 6. https://doi.org/10.1038/ng.2965
94. Ito, J., Sugimoto, R., Nakaoka, H., Yamada, S., Kimura, T., Hayano, T., & Inoue, I. (2017). Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses. *PLOS Genetics*, *13*(7), e1006883. https://doi.org/10.1371/journal.pgen.1006883
95. Ai, Z., Xiang, X., Xiang, Y., Szczerbinska, I., Qian, Y., Xu, X., Ma, C., Su, Y., Gao, B., Shen, H., Bin Ramli, M. N., Chen, D., Liu, Y., Hao, J.-J., Ng, H. H., Zhang, D., Chan, Y.-S., Liu, W., & Liang, H. (2022). Krüppel-like factor 5 rewires NANOG regulatory network to activate human naive pluripotency specific LTR7Ys and promote naive pluripotency. *Cell Reports*, *40*(8), 111240. https://doi.org/10.1016/j.celrep.2022.111240
96. Zhang, Y., Li, T., Preissl, S., Amaral, M. L., Grinstein, J. D., Farah, E. N., Destici, E., Qiu, Y., Hu, R., Lee, A. Y., Chee, S., Ma, K., Ye, Z., Zhu, Q., Huang, H., Fang, R., Yu, L., Izpisua Belmonte, J. C., Wu, J., … Ren, B. (2019). Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nature Genetics*, *51*(9), 1380–1388. Article 9. https://doi.org/10.1038/s41588-019-0479-7
97. Barakat, T. S., Halbritter, F., Zhang, M., Rendeiro, A. F., Perenthaler, E., Bock, C., & Chambers, I. (2018). Functional dissection of the enhancer repertoire in human embryonic stem cells. *Cell Stem Cell*, *23*(2), 276–288.e8.e8. https://doi.org/10.1016/j.stem.2018.06.014
98. Carter, T. A., Singh, M., Dumbović, G., Chobirko, J. D., Rinn, J. L., & Feschotte, C. (2022). Mosaic cis-regulatory evolution drives transcriptional partitioning of HERVH endogenous retrovirus in the human embryo. *eLife*, *11*, e76257. https://doi.org/10.7554/eLife.76257
99. Sexton, C. E., Tillett, R. L., & Han, M. V. (2022). The essential but enigmatic regulatory role of HERVH in pluripotency. *Trends in Genetics: TIG*, *38*(1), 12–21. https://doi.org/10.1016/j.tig.2021.07.007
100. Ivancevic, A., Simpson, D., Joyner, O., Stacey, B., Nguyen, L., Bitler, B., Pitts, T., & Chuong, E. (2024). Endogenous retroviruses mediate transcriptional rewiring in response to oncogenic signaling in colorectal cancer. *Science Advances*, *10*(29), eado1218 https://doi.org/10.1126/sciadv.ado1218
101. Yu, C., Lei, X., Chen, F., Mao, S., Lv, L., Liu, H., Hu, X., Wang, R., Shen, L., Zhang, N., Meng, Y., Shen, Y., Chen, J., Li, P., Huang, S., Lin, C., Zhang, Z., & Yuan, K. (2022). ARID1A loss derepresses a group of human endogenous retrovirus-H loci to modulate BRD4-dependent transcription. *Nature Communications*, *13*(1), 3501. https://doi.org/10.1038/s41467-022-31197-4
102. Kotekar, A., Singh, A. K., & Devaiah, B. N. (2023). BRD4 and MYC: Power couple in transcription and disease. *The FEBS Journal*, *290*(20), 4820–4842. https://doi.org/10.1111/febs.16580
103. Pérot, P., Mullins, C. S., Naville, M., Bressan, C., Hühns, M., Gock, M., Kühn, F., Volff, J.-N., Trillet-Lenoir, V., Linnebacher, M., & Mallet, F. (2015). Expression of young HERV-H loci in the course of colorectal carcinoma and correlation with molecular subtypes. *Oncotarget*, *6*(37), 40095–40111.
104. Song, D., Lian, Y., & Zhang, L. (2023). The potential of activator protein 1 (AP-1) in cancer targeted therapy. *Frontiers in Immunology*, *14*, 1224892. https://doi.org/10.3389/fimmu.2023.1224892
105. La Mantia, G., Maglione, D., Pengue, G., Di Cristofano, A., Simeone, A., Lanfrancone, L., & Lania, L. (1991). Identification and characterization of novel human endogenous retroviral sequences prefentially expressed in undifferentiated embryonal carcinoma cells. *Nucleic Acids Research*, *19*(7), 1513–1520.
106. Di Cristofano, A., Strazzullo, M., Longo, L., & La Mantia, G. (1995). Characterization and genomic mapping of the ZNF80 locus: Expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retrovrial family. *Nucleic Acids Research*, *23*(15), 2823–2830. https://doi.org/10.1093/nar/23.15.2823
107. Beyer, U., Moll-Rocek, J., Moll, U. M., & Dobbelstein, M. (2011). Endogenous retrovirus drives hitherto unknown proapoptotic p63 isoforms in the male germ line of humans and great apes. *Proceedings of the National Academy of Sciences*, *108*(9), 3624–3629. https://doi.org/10.1073/pnas.1016201108
108. Hu, T., Pi, W., Zhu, X., Yu, M., Ha, H., Shi, H., Choi, J.-H., & Tuan, D. (2017). Long non-coding RNAs transcribed by ERV-9 LTR retrotransposon act in cis to modulate long-range LTR enhancer function. *Nucleic Acids Research*, *45*(8), 4479–4492. https://doi.org/10.1093/nar/gkx055
109. Pi, W., Zhu, X., Wu, M., Wang, Y., Fulzele, S., Eroglu, A., Ling, J., & Tuan, D. (2010). Long-range function of an intergenic retrotransposon. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(29), 12992–12997. https://doi.org/10.1073/pnas.1004139107
110. Neumayr, C., Haberle, V., Serebreni, L., Karner, K., Hendy, O., Boija, A., Henninger, J. E., Li, C. H., Stejskal, K., Lin, G., Bergauer, K., Pagani, M., Rath, M., Mechtler, K., Arnold, C. D., & Stark, A. (2022). Differential cofactor dependencies define distinct types of human enhancers. *Nature*, *606*(7913), 406–413. https://doi.org/10.1038/s41586-022-04779-x
111. Yu, X., Zhu, X., Pi, W., Ling, J., Ko, L., Takeda, Y., & Tuan, D. (2005). The long terminal repeat (LTR) of ERV-9 human endogenous retrovirus binds to NF-Y in the assembly of an active LTR enhancer complex NF-Y/MZF1/GATA-2 *. *Journal of Biological Chemistry*, *280*(42), 35184–35194. https://doi.org/10.1074/jbc.M508138200
112. Hashimoto, K., Suzuki, A. M., Dos Santos, A., Desterke, C., Collino, A., Ghisletti, S., Braun, E., Bonetti, A., Fort, A., Qin, X.-Y., Radaelli, E., Kaczkowski, B., Forrest, A. R. R., Kojima, S., Samuel, D., Natoli, G.,

Buendia, M. A., Faivre, J., & Carnin, P. (2015). CAGE profiling of ncRNAs in hepatocellular carcinoma reveals widespread activation of retroviral LTR promoters in virus-induced tumors. *Genome Research*, *25*(12), 1812–1824. https://doi.org/10.1101/gr.191031.115

113. Kaplan, D. J., Jurka, J., Solus, J. F., & Duncan, C. H. (1991). Medium reiteration frequency repetitive sequences in the human genome. *Nucleic Acids Research*, *19*(17), 4731–4738.
114. Yu, H., Chen, M., Hu, Y., Ou, S., Yu, X., Liang, S., Li, N., Yang, M., Kong, X., Sun, C., Jia, S., Zhang, Q., Liu, L., Hurst, L. D., Li, R., Wang, W., & Wang, J. (2022). Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. *Cell Stem Cell*, *29*(7), 1031–1050.e12.e12. https://doi.org/10.1016/j.stem.2022.06.006
115. Pastor, W. A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T. J., Lukianchikov, A., Liu, X., Polo, J. M., Jacobsen, S. E., & Clark, A. T. (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. *Nature Cell Biology*, *20*(5), 553–564. Article 5. https://doi.org/10.1038/s41556-018-0089-0
116. Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., Kabayama, Y., Suyama, M., Sasaki, H., & Arima, T. (2018). Derivation of human trophoblast stem cells. *Cell Stem Cell*, *22*(1), 50–63.e6.e6. https://doi.org/10.1016/j.stem.2017.11.004
117. Dimitrova, Y., Gruber, A. J., Mittal, N., Ghosh, S., Dimitriades, B., Mathow, D., Grandy, W. A., Christofori, G., & Zavolan, M. (2017). TFAP2A is a component of the ZEB1/2 network that regulates TGFB1-induced epithelial to mesenchymal transition. *Biology Direct*, *12*(1), 8. https://doi.org/10.1186/s13062-017-0180-7
118. Wang, X., Sun, D., Tai, J., Chen, S., Yu, M., Ren, D., & Wang, L. (2018). TFAP2C promotes stemness and chemotherapeutic resistance in colorectal cancer via inactivating hippo signaling pathway. *Journal of Experimental & Clinical Cancer Research*, *37*(1), 27. https://doi.org/10.1186/s13046-018-0683-9
119. Kołat, D., Kałuzińska, Ż., Bednarek, A. K., & Płuciennik, E. (2019). The biological characteristics of transcription factors AP-2α and AP-2γ and their importance in various types of cancers. *Bioscience Reports*, *39*(3), BSR20181928. https://doi.org/10.1042/BSR20181928
120. Xiong, Y., Feng, Y., Zhao, J., Lei, J., Qiao, T., Zhou, Y., Lu, Q., Jiang, T., Jia, L., & Han, Y. (2021). TFAP2A potentiates lung adenocarcinoma metastasis by a novel miR-16 family/TFAP2A/PSG9/TGF-β signaling pathway. *Cell Death & Disease*, *12*(4), 352. https://doi.org/10.1038/s41419-021-03606-x
121. Kim, W., Kim, E., Lee, S., Kim, D., Chun, J., Park, K. H., Youn, H., & Youn, B. (2016). TFAP2C-mediated upregulation of TGFBR1 promotes lung tumorigenesis and epithelial-mesenchymal transition. *Experimental & Molecular Medicine*, *48*(11), e273. https://doi.org/10.1038/emm.2016.125
122. Grillo, G., & Lupien, M. (2022). Cancer-associated chromatin variants uncover the oncogenic role of transposable elements. *Current Opinion in Genetics & Development*, *74*, 101911. https://doi.org/10.1016/j.gde.2022.101911
123. Sexton, C. E., & Han, M. V. (2019). Paired-end mappability of transposable elements in the human genome. *Mobile DNA*, *10*(1), 29. https://doi.org/10.1186/s13100-019-0172-5
124. Lee, I., Razaghi, R., Gilpatrick, T., Molnar, M., Gershman, A., Sadowski, N., Sedlazeck, F. J., Hansen, K. D., Simpson, J. T., & Timp, W. (2020). Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing. *Nature Methods*, *17*(12), 1191–1199. https://doi.org/10.1038/s41592-020-01000-7
125. Berrens, R. V., Yang, A., Laumer, C. E., Lun, A. T. L., Bieberich, F., Law, C.-T., Lan, G., Imaz, M., Bowness, J. S., Brockdorff, N., Gaffney, D. J., & Marioni, J. C. (2022). Locus-specific expression of transposable elements in single cells with CELLO-seq. *Nature Biotechnology*, *40*(4), 546–554. https://doi.org/10.1038/s41587-021-01093-1
126. Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: From conflicts to benefits. *Nature Reviews Genetics*, *18*(2), 71–86. Article 2. https://doi.org/10.1038/nrg.2016.139
127. de Souza, F. S. J., Franchini, L. F., & Rubinstein, M. (2013). Exaptation of transposable elements into novel cis-regulatory elements: Is the evidence always strong? *Molecular Biology and Evolution*, *30*(6), 1239–1251. https://doi.org/10.1093/molbev/mst045
128. Pulver, C., Grun, D., Duc, J., Sheppard, S., Planet, E., Coudray, A., de Fondeville, R., Pontis, J., & Trono, D. (2023). Statistical learning quantifies transposable element-mediated cis-regulation. *Genome Biology*, *24*(1), 258. https://doi.org/10.1186/s13059-023-03085-7
129. Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M. F., Parker, B. J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., Ward, L. D., Lowe, C. B., Holloway, A. K., Clamp, M., Gnerre, S., Alföldi, J., Beal, K., Chang, J., Clawson, H., … Kellis, M. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature*, *478*(7370), 476–482. https://doi.org/10.1038/nature10530
130. Venuto, D., & Bourque, G. (2018). Identifying co-opted transposable elements using comparative epigenomics. *Development, Growth & Differentiation*, *60*(1), 53–62. https://doi.org/10.1111/dgd.12423
131. Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., Lee, E. A., Zhu, Y., Plajzer-Frick, I., Pickle, C. S., Kato, M., Garvin, T. H., Pham, Q. T., Harrington, A. N., Akiyama, J. A., Afzal, V., Lopez-Rios, J., Dickel, D. E., Visel, A., & Pennacchio, L. A. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*, *554*(7691), 239–243. Article 7691. https://doi.org/10.1038/nature25461
132. Todd, C. D., Deniz, Ö., Taylor, D., & Branco, M. R. (2019). Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. *eLife*, *8*, e44344. https://doi.org/10.7554/eLife.44344
133. Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L. C., Kuchler, A., Alpar, D., & Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. *Nature Methods*, *14*(3), 297–301. Article 3. https://doi.org/10.1038/nmeth.4177
134. Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R., Adamson, B., Norman, T. M., Lander, E. S., Weissman, J. S., Friedman, N., & Regev, A. (2016). Perturb-Seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell*, *167*(7), 1853–1866.e17.e17. https://doi.org/10.1016/j.cell.2016.11.038
135. Licht, J. D., & Bennett, R. L. (2021). Leveraging epigenetics to enhance the efficacy of immunotherapy. *Clinical Epigenetics*, *13*(1), 115. https://doi.org/10.1186/s13148-021-01100-x

How to cite this article: Karttunen, K., Patel, D., & Sahu, B. (2024). Transposable elements as drivers of dedifferentiation: Connections between enhancers in embryonic stem cells, placenta, and cancer. *BioEssays*, e2400059. https://doi.org/10.1002/bies.202400059