The impact of transposable element activity on therapeutically relevant human stem cells

Gerald G. Schumann[1]* , Nina V. Fuchs[2], Pablo Tristán-Ramos[3,4], Attila Sebe[1], Zoltán Ivics[1] and Sara R. Heras[3,4]

Abstract

Human stem cells harbor significant potential for basic and clinical translational research as well as regenerative medicine. Currently ~ 3000 adult and ~ 30 pluripotent stem cell-based, interventional clinical trials are ongoing worldwide, and numbers are increasing continuously. Although stem cells are promising cell sources to treat a wide range of human diseases, there are also concerns regarding potential risks associated with their clinical use, including genomic instability and tumorigenesis concerns. Thus, a deeper understanding of the factors and molecular mechanisms contributing to stem cell genome stability are a prerequisite to harnessing their therapeutic potential for degenerative diseases. Chemical and physical factors are known to influence the stability of stem cell genomes, together with random mutations and Copy Number Variants (CNVs) that accumulated in cultured human stem cells. Here we review the activity of endogenous transposable elements (TEs) in human multipotent and pluripotent stem cells, and the consequences of their mobility for genomic integrity and host gene expression. We describe transcriptional and post-transcriptional mechanisms antagonizing the spread of TEs in the human genome, and highlight those that are more prevalent in multipotent and pluripotent stem cells. Notably, TEs do not only represent a source of mutations/CNVs in genomes, but are also often harnessed as tools to engineer the stem cell genome; thus, we also describe and discuss the most widely applied transposon-based tools and highlight the most relevant areas of their biomedical applications in stem cells. Taken together, this review will contribute to the assessment of the risk that endogenous TE activity and the application of genetically engineered TEs constitute for the biosafety of stem cells to be used for substitutive and regenerative cell therapies.

Keywords: Adult stem cells, Genomic destabilization, LINE-1, Methylation, Pluripotent stem cells, Restriction, Regenerative medicine, Transposable elements

Background

Origin of stem cell types

Regenerative medicine is a recent and emerging branch of medical science, addressing functional restoration of specific tissues and/or organs of patients suffering from severe injuries or chronic diseases in a condition where the organism's own regenerative responses do not suffice [1]. Stem cells are defined by their ability to regenerate multiple differentiated cell types, while retaining the capacity to self-replicate and self-renew. Those found in vivo have different origins and can be divided into three broad categories

accordingly: embryonic (ESCs), foetal (FSCs) and adult stem cells (ASCs). Among the latter are hematopoietic stem cells (HSCs) and mesenchymal stem cells (MSCs), which represent the most used stem cell types in current clinical trials.

ESCs are derived from the inner cell mass of human blastocysts, have the potential for self-renewal, and are considered pluripotent: they maintain the ability to differentiate into cells and tissues from the three main germ layers [2] and generate all the tissues found in an organism. Human embryonic stem cells (hESCs) emerged as a primary cell source for regenerative medicine in recent years, to repair tissue and organ anomalies that resulted from congenital defects, disease, age and environment-associated effects. Similar characteristics are associated with induced pluripotent stem cells

* Correspondence: Gerald.Schumann@pei.de
[1]Division of Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str.51-59, 63225 Langen, Germany
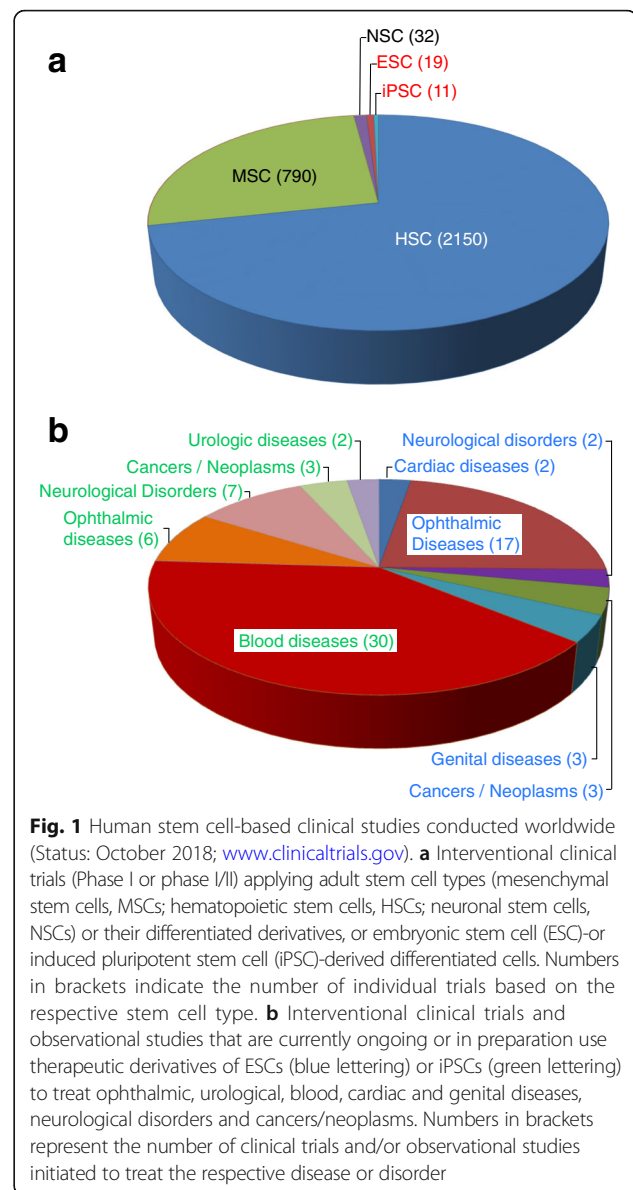Full list of author information is available at the end of the article

(iPSCs), although these are generated in vitro by ectopic expression of endogenous pluripotency factors which epigenetically transform terminally differentiated cells into ESC-like cells [3, 4].

FSCs, the second category of stem cells, are located in foetal tissues and embryonic annexes and are multipotent: they can differentiate into cell types from some, but not all of the three main germ layers [5]. FSCs have been subdivided into hematopoietic (blood, liver, bone marrow), mesenchymal (blood, liver, bone marrow, lung, kidney and pancreas), endothelial (bone marrow, placenta), epithelial (liver, pancreas) and neuronal stem cells (neurons, glia and oligodendrocytes [6].

The third category, termed ASCs or progenitor cells, comprises multipotent tissue-resident stem cells found in fully developed tissues. They reside in niches that create a special microenvironment for their replication and self-renewal. Domiciled in most tissues of the human body, discrete populations of ASCs generate cells to replace those that are lost through normal repair, disease, or injury. ASCs are found throughout the lifetime of the organism and were identified in tissues such as the umbilical cord, placenta, bone marrow, muscle, brain, fat tissue, skin, gut, etc. ASCs are thought to act to repair and regenerate tissues in which they reside, helping to maintain tissue homeostasis.

Clinical application of adult and pluripotent stem cells

Adult stem cells such as HSCs, MSCs, and neural stem cells (NSCs) are the most frequently used cell types in interventional clinical trials (www.clinicaltrials.gov) (Fig. 1a). HSCs are essential for the generation and homeostasis of the blood system, and give rise to all blood cell types, including lymphocytes, erythrocytes, monocytes, granulocytes, and platelets [7]. Indeed, HSC transplantation is the accepted therapy of choice for a variety of malignant and non-malignant blood-related diseases in children and adults. Initially developed as rescue therapy for a patient with cancer after high doses of chemotherapy and radiation, as well as the correction of severe deficiencies in the hematopoietic system, it has evolved into an adoptive immune therapy for malignancies and autoimmune disorders [8]. Human MSCs are multipotent stem cells with the capacity to differentiate into the mesodermal lineage such as osteocytes, adipocytes and chondrocytes, as well as some ectodermal (neurocytes) and endodermal lineages (hepatocytes) [9]. MSCs have been isolated from various tissues including bone marrow, adipose tissue, amniotic fluid, endometrium, dental tissues, umbilical cord and Wharton's jelly [9]. MSCs have immunomodulatory features and secrete cytokines and immune receptors which regulate the microenvironment in the host tissue [10]. Multilineage potential, immunomodulation and secretion of

Fig. 1 Human stem cell-based clinical studies conducted worldwide (Status: October 2018; www.clinicaltrials.gov). **a** Interventional clinical trials (Phase I or phase I/II) applying adult stem cell types (mesenchymal stem cells, MSCs; hematopoietic stem cells, HSCs; neuronal stem cells, NSCs) or their differentiated derivatives, or embryonic stem cell (ESC)-or induced pluripotent stem cell (iPSC)-derived differentiated cells. Numbers in brackets indicate the number of individual trials based on the respective stem cell type. **b** Interventional clinical trials and observational studies that are currently ongoing or in preparation use therapeutic derivatives of ESCs (blue lettering) or iPSCs (green lettering) to treat ophthalmic, urological, blood, cardiac and genital diseases, neurological disorders and cancers/neoplasms. Numbers in brackets represent the number of clinical trials and/or observational studies initiated to treat the respective disease or disorder

anti-inflammatory molecules makes MSCs an effective tool in the treatment of chronic diseases. Consequently, there are several clinical trials harnessing their innate characteristics or other in vitro observations in a series of diseases (for a review see [11]). NSCs are a group of ectodermal progenitor cells which can differentiate into neural subtypes, such as neurons, astrocytes or oligodendrocytes. Neural stem cell derivatives are used in a number of clinical trial applications including treatments of ALS and Parkinson's Disease [12–14] and currently ongoing interventional clinical trials (Fig. 1a).

In contrast to MSCs and other types of tissue-specific stem cells, pluripotent stem cells (PSCs) are derived either from human pre-implantation embryos, giving rise to hESCs [2, 15], or from somatic cells that are reprogrammed

Schumann *et al. Mobile DNA*   (2019) 10:9

Page 3 of 23

to a primitive pluripotent state (hiPSCs). PSCs are immortal and highly expandable in culture in vitro, and can be differentiated to almost any cell type of the body. Their potential for regenerative medicine is therefore unique and extraordinary. Indeed, cellular products derived from hESCs are now in clinical trials for cardiac and ophthalmic diseases and neurological disorders, with some other applications registered for clinical trial approval (Fig. 1b) [12–14]. Initially, hiPSCs have been used in one experimental procedure in an autologous approach on an individual in Japan with macular degeneration [16, 17]. In March 2017, the first study was initiated involving 5 AMD (Age-related macular degeneration) patients who received retina cells derived from banked hiPSCs in an allogeneic approach [18]. To date, 11 interventional clinical trials and 25 observational studies are based on the application of iPSCs (Fig. 1). However, and despite these trials in the frontier of knowledge, relatively little is known about undesired long-term effects of such approaches.
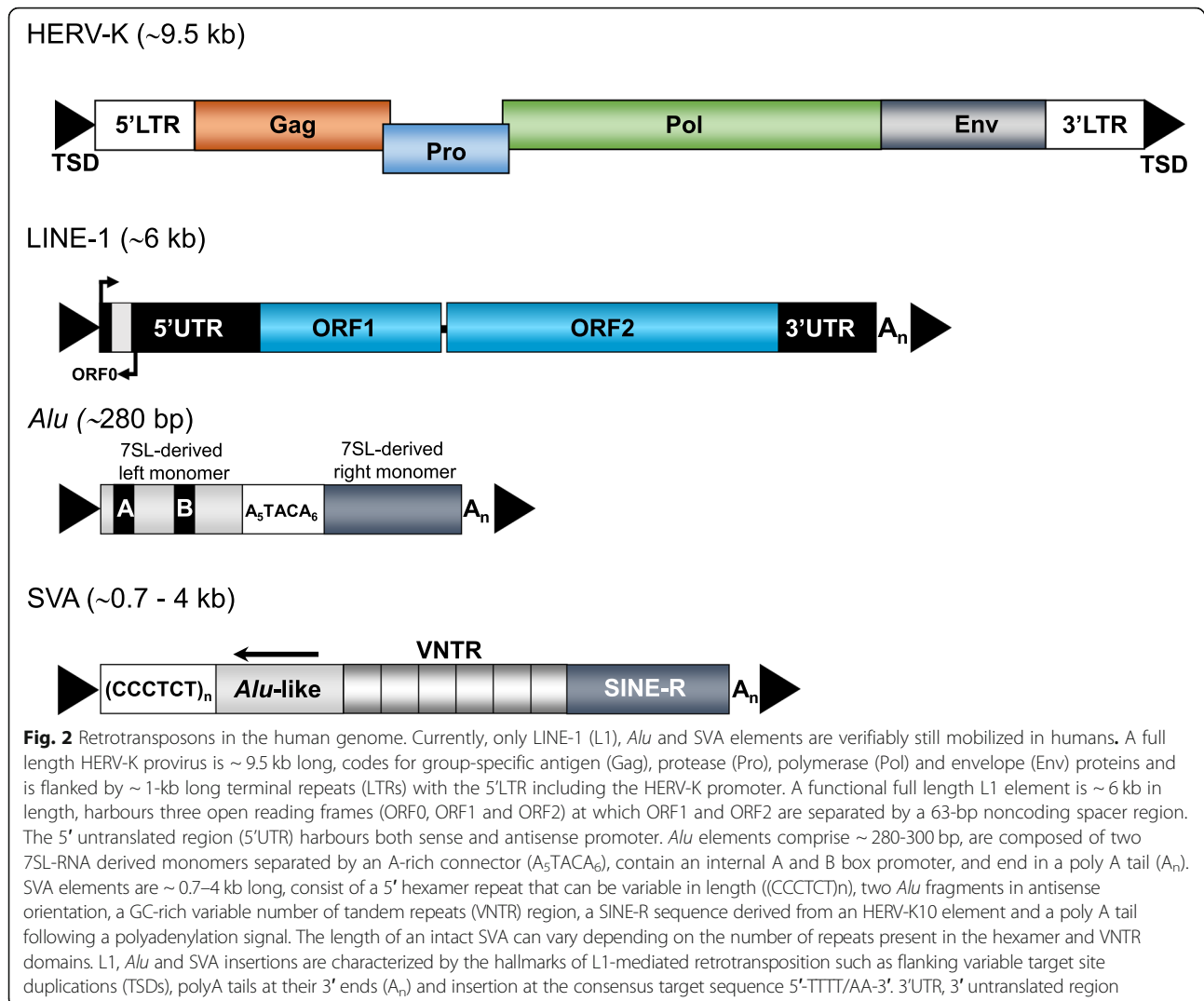
The issue of genomic integrity

The promise for human disease treatment using differentiated cells derived from multipotent ASCs and pluripotent stem cells, such as hESCs and hiPSCs, also carries the threat of genomic instability of the cells to be administered. Firstly, cultivation of multipotent and pluripotent stem cells exposes the cells to selection pressures that often result in the acquisition and manifestation of genomic alterations, varying in size from point mutations, through copy number changes in small genomic elements (e.g. amplification of repetitive sequences and retroelement mobility), to large chromosomal aberrations, trisomies and monosomies [19–21]. Previous reviews reported several factors that contribute to differences in genomic and epigenomic stabilities of stem cells, including derivation source (embryonic vs. somatic cells), derivation methods (direct isolation vs. reprogramming), and culture conditions [22]. Much attention has been drawn in recent years to the genomic aberrations acquired by hESCs and hiPSCs, ranging from point mutations to whole-chromosome trisomies [23–30]. Similarly, human ASCs that are expanded in culture were also shown to be prone to acquire chromosomal aberrations [24]. Secondly, the treatment of many human diseases often involve genetic manipulation of stem cells prior to transplantation, which may further jeopardize their genomic stability. Overall, genomic aberrations can affect identity, differentiation capability and tumorigenicity of stem cells, and should thus be routinely evaluated for their proper use in basic research and in clinical trials. In the promising era of stem cell research and therapy, ensuring genomic stability of stem cells and their derivatives remains one of the highest priorities prior to clinical translation.

In this review, we focus on one specific source of genomic instability in human therapeutically relevant stem cells that has been mostly ignored by the stem cell community to date, namely the activity of endogenous non-Long Terminal Repeat (non-LTR)-retrotransposons, and the consequences for genomic integrity and host gene expression. Non-LTR retrotransposons constitute our center of attention because in contrast to most TEs in our genome, a small fraction of this group of TEs is currently active and mobilized in the human population [31, 32]. We provide an overview of the impact of endogenous TEs in pluripotent and adult stem cells, discuss new roles of TEs in regulating pluripotency, and describe host defense systems counteracting TE activity in stem cells. Furthermore, we address the application of DNA-transposons to genetically engineer human stem cells for medical applications. In order to fulfill the standards for safe clinical applications and evaluate the risk for biosafety inflicted on therapeutically relevant cells by TE activity, we propose that the extent of the activity of potentially mutagenic TEs in pluripotent stem cells needs to be elucidated, and perhaps used as an additional quality control check point in the future.

Non–LTR retrotransposons are currently active in the human genome
Retroelement families

More than 45% of the human genome is made of TE-derived sequences, but only a specific subset of these TEs that belongs to the group of retrotransposons is currently propagating in the human genome. Retrotransposons spread by a 'copy-and-paste' mechanism that involves reverse transcription and insertion of an intermediate RNA at a new site in the genome. They are classified in two different groups: LTR-retrotransposons, which include human endogenous retroviruses (HERVs) and are flanked by long terminal repeats harbouring transcriptional regulatory domains, and non-LTR retrotransposons that lack LTRs (Fig. 2). Among the 31 human endogenous retrovirus subfamilies covering 8% of the human genome, replication-competent HERVs have not been described to date, although their existence cannot be ruled out [33, 34]. The only TEs that are currently verifiably mobilized in the human genome are members of the group of non-LTR retrotransposons (Fig. 2). They comprise two main types of elements: i) autonomous Long interspersed class 1 elements (LINE-1s or L1s), and ii) non-autonomous short interspersed elements (SINEs) such as *Alu* and SVA (SINE-VNTR [variable number of tandem repeats] -*Alu* elements) elements. In humans, > 500,000 L1 copies occupy ~ 17% of the haploid genome [35]. A functional, full-length L1 is 6 kb in length and initiates transcription using an internal sense-promoter located within its 5′

Fig. 2 Retrotransposons in the human genome. Currently, only LINE-1 (L1), *Alu* and SVA elements are verifiably still mobilized in humans. A full length HERV-K provirus is ~ 9.5 kb long, codes for group-specific antigen (Gag), protease (Pro), polymerase (Pol) and envelope (Env) proteins and is flanked by ~ 1-kb long terminal repeats (LTRs) with the 5'LTR including the HERV-K promoter. A functional full length L1 element is ~ 6 kb in length, harbours three open reading frames (ORF0, ORF1 and ORF2) at which ORF1 and ORF2 are separated by a 63-bp noncoding spacer region. The 5' untranslated region (5'UTR) harbours both sense and antisense promoter. *Alu* elements comprise ~ 280-300 bp, are composed of two 7SL-RNA derived monomers separated by an A-rich connector (A_5TACA_6), contain an internal A and B box promoter, and end in a poly A tail (A_n). SVA elements are ~ 0.7–4 kb long, consist of a 5' hexamer repeat that can be variable in length ((CCCTCT)n), two *Alu* fragments in antisense orientation, a GC-rich variable number of tandem repeats (VNTR) region, a SINE-R sequence derived from an HERV-K10 element and a poly A tail following a polyadenylation signal. The length of an intact SVA can vary depending on the number of repeats present in the hexamer and VNTR domains. L1, *Alu* and SVA insertions are characterized by the hallmarks of L1-mediated retrotransposition such as flanking variable target site duplications (TSDs), polyA tails at their 3' ends (A_n) and insertion at the consensus target sequence 5'-TTTT/AA-3'. 3'UTR, 3' untranslated region

untranslated region (UTR) [36–38] (Fig. 2). In functional L1s the sense promoter drives expression of two proteins, named ORF1p and ORF2p, that catalyze L1 retrotransposition in *cis* [39]. ORF1p is a 40 kDa RNA binding protein with nucleic acid chaperone activity, while ORF2p is a 150 kDa protein that contains reverse transcriptase (RT) and endonuclease activities (reviewed in [32]). Both proteins are strictly required for retrotransposition [40] by a process termed Target-Primed Reverse Transcription (TPRT; [41, 42]). A conserved antisense-promoter in the L1 5'UTR [36, 37] drives expression of the *trans*-acting polypeptide ORF0, that can stimulate L1 mobilization [43]. Although the RT activity of human L1 was shown to be highly processive [44], most de novo L1 insertions are rendered immobile by 5′ truncation. These truncations together with internal mutations accumulated over time, have resulted in only 80–100 retrotransposition-competent or functional L1s per individual human genome [45–47]. Of these, fewer

than 10 mobilize efficiently when tested in vitro: they are termed 'hot' L1s and responsible for the bulk of retrotransposition in the human population [45, 47–50]. The vast majority of hot L1s belong to the youngest subfamily, termed L1Hs (for Homo sapiens-specific, also known as L1-Ta, for transcribed-active) [45, 47, 51]. L1 subfamilies can be distinguished by nucleotide substitutions that are shared by all members of a specific subfamily. Five subfamilies are thought to have amplified in hominoid primates within the past 25 myrs, named L1PA1 to L1PA5. Functional L1 s are also responsible for the mobilization of *Alu* and SVA elements which lack protein-coding capacity [52–55]. Indeed, L1 activity *in trans* can act at low frequency on any cellular polyadenylated mRNA, resulting in the generation of processed pseudogenes [56, 57]. Members of the primate-specific *Alu* family are approximately 300-bp long, have a dimeric structure, and are transcribed by RNA polymerase III (reviewed in [58, 59]) (Fig. 2). The

family of *Alu* elements is, with > 1 million copies, the most abundant TE in our genome [35], and represents the currently most active group of retrotransposons in humans with ~ one thousand mobilization-competent copies per genome. The SVAs are hominid-specific composite non-coding retroelements, with an average length of ~ 2 kb, and represent the youngest family of active human TEs (Fig. 2). There are roughly 2700 SVA copies per human genome, most of which are full-length and about 20–50 of which may be active [60–67].

In sum, all above listed mobilization-competent non-LTR retrotransposons may collectively account for thousands of somatic insertions [68]. It was calculated that up to 5% of newborns harbour a new retrotransposition event, and > 120 known human disease-causing insertions of L1 s, *Alu*s and SVAs have been identified to date [69–73]. L1-mediated insertions have sporadically resulted in a wide number of human genetic disorders, including hemophilia A (L1), cystic fibrosis (*Alu*) [74] and X-linked dystonia-parkinsonism (SVA) to cite a few reviewed in [71]. Altered expression of TEs and animation of genomic L1 sequences also appear to be hallmarks of cancer, and can be responsible for driving mutations in tumorigenesis [75, 76]. Indeed, the random mobilization of L1 in our genome implies that virtually any disorder can be generated by the mutagenic activity of non-LTR retrotransposons, and L1-mediated retrotransposition events occur approximately once in every 250 pathogenic human mutations [77].

TE insertions can affect host gene expression by various mechanisms and impact stem cell development

When expressed, TEs can affect developmental processes either by their encoded gene products, which could influence phenotype and/or function of the host cell, or through their genomic de novo insertion, that could result in genetic changes in the host genome. L1-mediated retrotransposition events can be associated with the induction of local genomic instability, including DNA deletions at the insertion site, 3′- or 5′-transductions, and non-allelic homologous recombination (reviewed in [32, 78]). Additionally, new retrotransposon insertions into genes can act as insertional mutagens, alter gene expression and interfere with host gene function [79].

Although fewer than 1 of 10,000 TEs is still capable of mobilization [71], a far greater proportion can influence gene expression in *cis* and in *trans* (reviewed in [80]). This is because TEs, such as ERVs, L1 and SVA elements, contain transcription factor binding sites that promote transcription by RNA polymerase II or, in the case of *Alu* elements, by RNA polymerase III. Therefore, the promoter(s) of resident and new TE(s) that integrated into or close to host genes, can alter gene expression, and even lead to the appearance of novel transcripts that encompass part of the coding region.

For instance, because transcripts of several host genes in iPSCs originate from a nearby L1 antisense promoter, the most abundant transcript isoforms expressed contain L1 ORF0 sequences whereas for other genes, these isoforms contribute only little to the bulk transcripts. Interestingly, ORF0 transcript levels were dramatically downregulated in the source fibroblasts relative to their derived iPSCs [43]. De novo genic insertions can also induce premature termination of transcription by introducing TE-derived polyadenylation signals [81], and sporadically result in the exonization of TEs. Exonized TEs can expand the mammalian transcriptome and proteome, but can also be used to fine-tune gene regulation [82]. Indeed, nucleotide sequences of several classes of TEs have been found inserted within mammalian RNAs [82–84] raising the possibility that they affect gene regulation and function by providing or interfering with regulatory elements in those RNAs. While the decrease in host gene RNA expression due to inadequate transcript elongation after L1 insertion in an intron of a transcribed gene was originally attributed to the A/T richness of the L1 sequence [85], it was suggested recently that transcriptional attenuation of host gene expression could be a consequence of epigenetic silencing mediated by the human silencing hub (HUSH) complex [86]. TE insertions can also introduce TE-derived splice acceptor or donor sites that alter splicing, and thereby generate non-functional or nonsense transcripts [87] or can be incorporated into mRNAs and thus introduce frameshifts or premature termination codons. It has also been shown recently that spliced L1 mRNAs are also used as templates generating spliced integrated retrotransposed elements (SpIREs) which lack part of the 5′UTR and, consequently, lack some of the regulatory sequences [88]. L1-encoded proteins can sporadically generate processed pseudogenes [39, 56], and thereby affect the stem cell transcriptome and proteome. Finally, ongoing L1 retrotransposition itself can copy regulatory sequences such as promoter and enhancer regions and/or protein coding regions and introduce them into new sites in the host genome by a mechanism termed "exon shuffling" [89]. Exon shuffling occurs when an L1 insert resides within a gene, L1 transcription bypasses the inherent weak L1 polyadenylation signal, and transcribes downstream exons generating a chimeric RNA, which could then retrotranspose to a new genomic location. It is estimated that ~ 15–20% of the L1Ta subfamily members in the human genome contain 3′-transduced sequences. Such L1-mediated 3′ transductions are responsible for as much as 1% of the human genome [90, 91], and this process has probably increased the repertoire of the human proteome. In summary, there are multiple manners in which TE insertions can impact the functioning and regulation of the human genome.

Activity of non-LTR retrotransposons in human pluripotent stem cells

Primary milieu for the evolutionary struggle between TEs and the host genome are pluripotent cells of the early mammalian embryo and germ cells

To ensure their evolutionary success, TEs need to mobilize in germ cells and/ or during early development, before germline partitioning, to guarantee transmission of new retrotransposition events to the next generation. A genome-wide epigenetic switch in the early mammalian embryo, characterized by global DNA demethylation, is necessary to activate the programme of embryonic development. This 'epigenomic reset' is thought to provide a window of opportunity for retrotransposons to mobilize and create heritable insertions [92], as L1 expression is also controlled by DNA methylation [93]. Consistently, previous studies have shown that i) L1 mRNAs are present in human immature diploid oocytes from in vitro fertilization donors [94], ii) both L1-encoded proteins, ORF1p and ORF2p, are expressed in prespermatogonia of fetal testes and in germ cells of adult testes [95], and iii) hESCs overexpress a combination of potentially functional full-length L1 and core *Alu* elements as well as non-functional L1 and *Alu* RNAs (Fig. 3) [79, 96–100]. The observed L1 expression levels in these pluripotent

stem cells inversely correlate with both their low global DNA methylation levels, and their levels of L1 promoter methylation [97, 101, 102]. It has been reported that only a specific subset of L1 retrotransposons, likely under the combined influence of activators or repressors and local chromatin constraints, is active at any given period of early embryogenic development type [67, 96, 103, 104]. While expression of members of the youngest L1 lineages, L1Hs and L1PA2, is highest in hESCs, expression of a subset of L1 lineages, primarily L1PA3, L1PA4, L1PA5 and L1PA6, which were predicted to have entered the ancestral genome between 26.8 and 7.6 million years ago, is reduced [98]. Thus, these data demonstrated that RNAs from active and "hot" L1 s are expressed in hESCs and hiPSCs. Using an engineered L1 retrotransposition reporter assay in cultured cells [40], it was demonstrated that the cellular environment of human and murine ESCs support L1 retrotransposition (Fig. 3), and that L1 de novo insertions in cultivated hESCs can occur into genes and lead to small deletions of genomic DNA at the target site [97, 105, 106]. In agreement with these data demonstrating that host factors essential for retrotransposition of functional L1 mRNAs are present in hESCs, it was verified that endogenous *Alu*Ya5 elements are *trans*-mobilized in cultured hiPSCs [79].
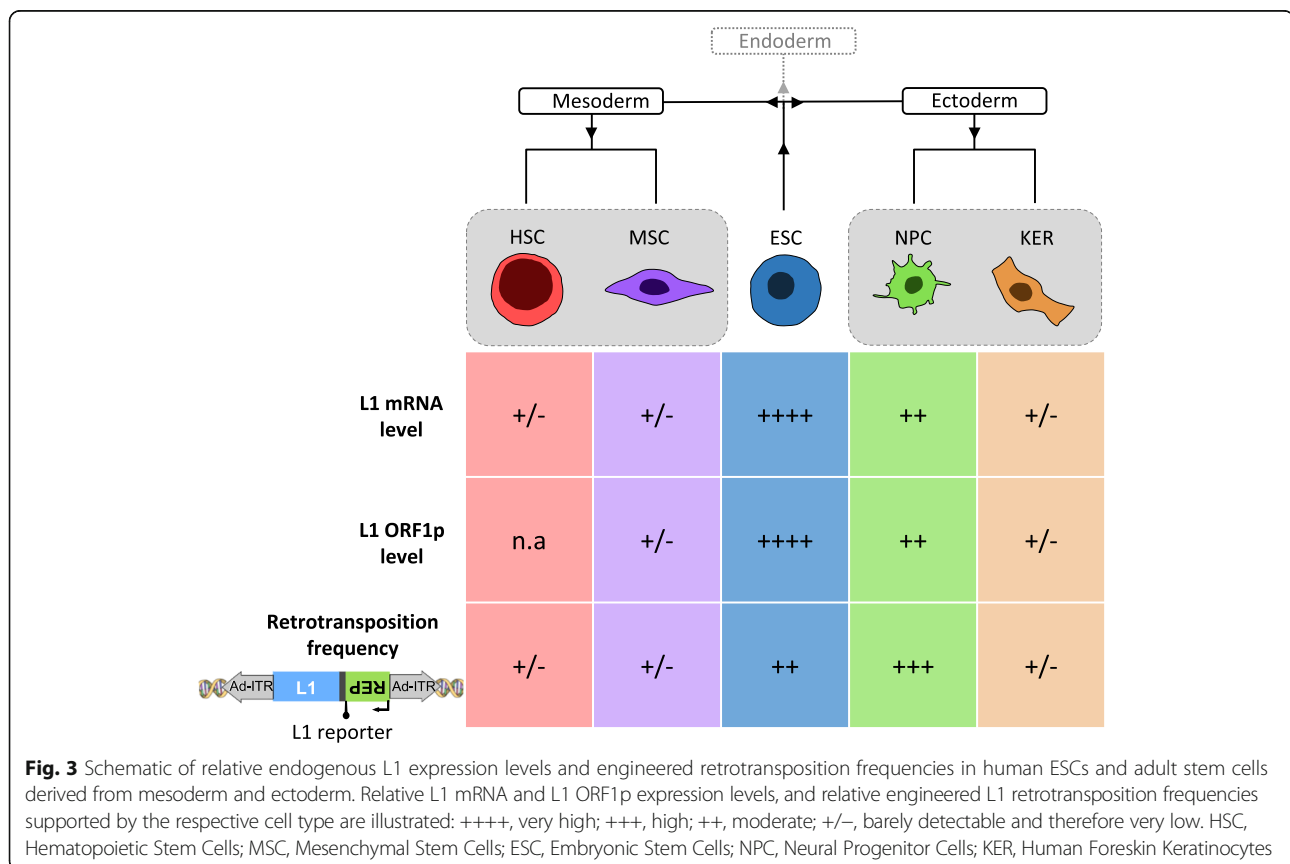

Fig. 3 Schematic of relative endogenous L1 expression levels and engineered retrotransposition frequencies in human ESCs and adult stem cells derived from mesoderm and ectoderm. Relative L1 mRNA and L1 ORF1p expression levels, and relative engineered L1 retrotransposition frequencies supported by the respective cell type are illustrated: ++++, very high; +++, high; ++, moderate; +/–, barely detectable and therefore very low. HSC, Hematopoietic Stem Cells; MSC, Mesenchymal Stem Cells; ESC, Embryonic Stem Cells; NPC, Neural Progenitor Cells; KER, Human Foreskin Keratinocytes

Somatic and heritable L1-mediated retrotransposition in pluripotent stem cells during early development

Studies of transgenic mouse models established to investigate engineered L1 retrotransposition in vivo, and, more recently, the sequencing of mouse pedigrees, confirmed that the majority of heritable de novo retrotransposition events in mice occur in preimplantation embryos, and indicate that embryonic cells are likely to be the natural habitat for L1-mediated retrotransposition [107–113]). Notably, such L1 mobilization in pluripotent cells of pre-implantation embryos could result in the new insertion being a mosaic within somatic and germline cells in adult mice. Although it was reported that L1 transcript levels peak after fertilization at the 2-cell-stage and decrease from the 2- to the 8-cell stages in mice [114], it remains unclear at which time point after fertilization L1-retrotransposition is initiated. Endogenous L1 elements were shown to be active in mouse early primordial germ cells, before the PIWI/piRNA retrotransposon defense pathway is activated in male embryonic gonads [113, 115].

By contrast, we currently know relatively little about TE expression and timing of retrotransposition in humans. An example that unambiguously proves that endogenous L1 mobilization can occur early in human embryogenesis was reported by van den Hurk and colleagues: they characterized a mutagenic L1 insertion resulting in choroideremia, a rare recessive X-linked eye disorder, in an affected male patient whose mother was a somatic and germline mosaic for the pathogenic L1 insertion [116]. Due to a 3′ transduction carried by the mutagenic L1 insertion, the authors were able to identify the specific donor L1 element and verify that it was retrotransposition-competent [116, 117]. Taken together, data strongly suggest that, in humans, new heritable L1 and *Alu* insertions can likely accumulate during early embryogenesis, and to some extent in germ cells [97, 107, 110]. Retrotransposition events mediated by L1 were estimated to occur, at a minimum, in 1 of 20 meioses for *Alu*, 1 of 20 to 200 meioses for L1, and 1 of 900 meioses for SVA [118]. Clearly, future studies should determine the timing of L1 retrotransposition during embryogenesis and a rate of retrotransposition in modern humans.

Very recently, studies in mice provided evidence that the observed transcriptional L1 activation may have a role in mouse ESCs and pre-implantation embryos. Activation of L1 transcription after fertilization regulates global chromatin accessibility at the beginning of development, and therefore could be important for normal embryonic development [119]. Also, the L1 RNA seems to act as a nuclear scaffold to repress the transcriptional program specific to the 2-cell embryo and, consequently, the L1 transcript could be required to exit from the 2-cell stage [120]. These data suggest that L1-encoded RNAs and proteins have additional roles in embryonic development. These findings seem paradoxical, as despite the mutagenic potential of L1 during embryogenesis, they suggest that L1 activity could be required for proper embryonic development and even to maintain ESC identity.

TE activity as parameter to define the naïve pluripotent state in hESCs

Mouse ESCs and iPSCs were proposed to represent a naïve state of pluripotency corresponding to the inner cell mass, whereas hESCs/iPSCs correspond to a more advanced, or 'primed', state of pluripotency found in the postimplantation epiblast [121]. Indeed, it has been challenging to define the naïve state of pluripotency in hESCs, particularly in view of the expanding number of protocols for deriving putative naïve human cells [122, 123]. Putative naïve human cultured cells were examined according to their resemblance to human pluripotent cells in vivo. Three comparative parameters for evaluating the naïve state of human pluripotent stem cells were established: (1) the expression profile of TEs based on single-cell RNA-seq data [124], (2) the DNA methylation landscape of human preimplantation development [125], and (3) the chromosome X inactivation status of female human ESCs [126, 127]. Comprehensive examination of the "transcriptome" was found to offer a more sensitive measure of the correspondence between pluripotent stem cells and the early human embryo than gene expression profiling [128]. Constituting a unique model system to dissect mechanisms of early human development, naïve hESCs provide an excellent cellular model to interrogate the function of transposable elements that become activated during early human development [128].

Cultured human pluripotent stem cells support expression and mobilization of endogenous TEs

Human iPSCs are currently successfully applied for disease modeling, the study of cell development and function, and for in vitro screening of drug candidates on healthy and diseased cells; thus, cellular products derived from hiPSCs hold substantial promise for substitutive and regenerative, autologous and allogenic cell therapies. Unlike hESCs, hiPSCs are a potential source of autologous cells compatible with the immune system of transplant recipients [129]. hiPSCs also circumvent ethical issues associated with the use of human embryos to derivate hESCs [129]. However, genetic and epigenetic aberrations that occur during reprogramming and in vitro expansion [25–28, 79, 130] may hinder the use of hiPSCs in regenerative medicine because they could affect their identity and differentiation capability, and could elevate the risk of tumorigenesis upon implantation [131]. Thus, identifying the full spectrum of aberrant mutational processes accumulated in hiPSCs, and

their functional consequences, is of paramount significance. A direct consequence of reprogramming, due to genomewide epigenomic changes [130], is the activation of full-length L1 mRNA expression; indeed, L1-ORF1p and L1 RNA transcript levels are upregulated by up to four orders of magnitude compared to their respective parental cells [79, 100, 132, 133]. To investigate if hiPSCs provide the environment for L1 retrotransposition, Wissing and colleagues exploited an in vitro L1 retrotransposition assay, and detected similar levels of L1-engineered retrotransposition as those previously reported in hESCs [97, 100]. These data demonstrated that the cellular milieu of both hESCs and hiPSCs contains the host factors required for L1 retrotransposition. In order to clarify if the transcriptional activation of endogenous L1 elements in hiPSCs leads to L1-mediated mobilization, targeted sequencing was performed on cultured hESC and hiPSC lines, followed by PCR validation of candidate de novo L1, *Alu* and SVA insertions in multiple laboratories [79]. Notably, the sequencing of a small number of lines revealed ongoing L1, *Alu* and SVA retrotransposition in hiPSCs; Similarly, one *Alu* de novo insertion was found to have occurred during the cultivation of the hESC line H9, which was the only hESC line analyzed in this study [79]. Thus, these studies confirmed that hESCs and hiPSCs are a natural habitat for L1-mediated retrotransposition, consistent with heritable retrotransposition during early human embryogenesis. However, the currently available datasets on L1 mobilization in hiPSCs vs hESCs are not sufficient to allow conclusions on the frequency of L1-mediated retrotransposition events in individual hESCs or hiPSCs. Remarkably, an earlier whole-genome sequencing (WGS)-based study in mouse ESCs failed to detect endogenous L1 retrotransposition [134]. Lack of retrotransposition in mESCs when compared to hESCs/hiPSCs might reflect distinctly different properties of human primed [79] and mouse naïve [134] pluripotency states; or the limitations of the method employed for detecting de novo retrotransposition events, namely genome-wide DNA sequencing with [79] or without [134] a previous retrotransposon enrichment step. Overall, these data might suggest that the genome of mESCs would be more stable than hESCs/hiPSCs, although several studies have presented compelling evidence demonstrating ongoing (and heritable) L1 retrotransposition during early mouse embryogenesis [107, 113]. Thus, it is also possible that cultured mESCs fail to recapitulate this aspect of mouse early embryonic biology, and additional studies are required to solve this contradiction.

Occurrence and structures of de novo retrotransposition events in cultured hESCs and hiPSCs

An intriguing aspect of L1 retrotransposition studies in hESCs/hiPSCs is that most validated insertions are

full-length. In hiPSCs, 57–66% of all validated endogenous and engineered L1 de novo retrotransposition events were full-length [79, 100]. Another report found seven potential de novo L1 insertions in two hiPSC lines using targeted L1 sequencing, but could not PCR-validate or fully characterize the genomic integration sites of these events [133]. The success in the identification of L1 mobilization events depends on multiple factors such as heterogeneity of the investigated cell population, and methodological factors such as stem cell culture conditions, population bottlenecks in cultured cells, high-throughput sequencing method used, bioinformatic parameters, and how candidate L1 insertions are validated, which can affect the results decisively. For example, a recent study applied WGS to nine hiPSC lines [135] and did not identify any de novo retrotransposon insertions, and far fewer mutations overall when compared to earlier studies [25, 26]. While an early report in hESCs validated a small number of engineered but 5′-truncated L1 insertions [97], additional insertion characterization in hESCs has recently revealed that > 70% of them are full-length (Cano and Garcia-Perez, personal communication). The finding that 66–75% of engineered L1 retrotranposition events in hESCs and hiPSCs are full-length is in stark contrast to the situation in cancer cell lines, where only ~ 5% of engineered L1 de novo insertions represent full-length elements [136]. While this discrepancy warrants additional research, it is tempting to speculate that due to L1 overexpression, host factors involved in 5′ truncation might be titrated/sequestered in hESCs/hiPSCs allowing a higher rate of full-length insertions in these cells. Notably, other features of de novo L1 insertions events do not differ significantly between cancer cells and hESCs/hiPSCs; For example, ~ 50% of all validated L1 retrotransposition events in both cell types occurred into introns of host genes [79, 100, 136]. Similar to transformed cells, intronic L1 insertions into hiPSCs can interfere with the transcription of the affected host gene [79].

Epigenomic reprogramming causes a constant derepression of endogenous TEs

Despite the unknown aspects of L1 biology in ESCs, available data indicate that reprogramming elicits a dynamic but consistent derepression of endogenous L1 s, and that they experience a relaxation of host genome control in pluripotent stem cells obtained from embryonic tissues [111, 128, 132, 137, 138]. A recent study examining transcriptional control of endogenous retroelements during the reprogramming of human hematopoietic stem cells from cord blood and primary hepatocytes into hiPSCs and subsequent cultivation uncovered a marked transcriptional upregulation of active SVA_A and L1Hs

retrotransposons, but also of currently inactive HERV-K, HERV-H and HERV-W elements [132]. A sharp increase of L1 and SVA transcription was observed several days after transduction with reprogramming vectors, and both RNAs remained highly expressed upon culturing hiPSCs [132]. While there was only little heterogeneity in L1 expression levels between tested hESC lines, considerable heterogeneity between hiPSC clones was observed [132]. Comparison of expression levels of individual endogenous retroelements between hiPSC clones derived from a single donor and issued from the same reprogramming experiment uncovered striking differences, notably for HERV-H, HERV-K and L1Hs. Thus, these data suggest that the genome wide epigenomic reprogramming might not occur to the same extent in all reprogrammed cells, and that hESCs might represent a more realistic model to study human early embryogenesis processes. It was hypothesized that transcriptionally activated functional L1Hs elements, whose antisense promoters can affect the expression of neighbouring genes [139], have a similar effect during or after reprogramming of hiPSCs [132]. The interference of *cis-* and *trans*-acting transcriptional regulatory elements of preexisting or de novo TE insertions with the transcription of TE-close genes may result in phenotypic anomalies difficult to detect through conventional assays, such as blockade of differentiation to particular lineages, predisposition to oncogenic changes, aberrant release of bioactive molecules, altered immunogenicity or ectopic activation of disease-related genes in iPSCs or their progeny [67, 132]. In hiPSCs, unsteady production of transcripts from TE-integrants normally silenced in ESCs and secondary activation of neighbouring genes was noted [67]. This phenomenon could contribute to the inefficiency of reprogramming by stochastically activated genes that affect the path to pluripotency [140]. Transcriptional perturbation of TE-close genes may also plague iPSCs and their progeny with phenotypic anomalies [141]. These findings argue for an in-depth survey of the genomic, transcriptional, and epigenetic state of the repetitive genome of iPSC clones, that are to be used for basic research or clinical applications. In these studies, bona fide established hESC lines could offer an excellent reference to detect harmful changes in hiPSCs that might limit their applications. Indeed, and while the use of hESCs on regenerative medicine is clearly limited, hESCs have become an excellent benchmark to establish clinically safer human pluripotent embryonic stem cell-like cells.

Activity of endogenous non-LTR retrotransposons in adult stem cells

Breaking the dogma – TE activity in adult stem cells of the brain

TEs are the prototype of "selfish DNA", whose only purpose is to generate more copies of themselves. Thus, it was assumed for a long time that most TE activity in mammals might occur only in cellular niches that transmit genetic information to the next generation, as this would ensure the evolutionary success of TEs. Although pioneer work by Barbara McClintock discovered TEs in somatic tissues of corn [142], the lack of TE activity in the mammalian soma has been a long-standing dogma. However, this "dogma" was recently challenged with the description of TE activity in mammalian healthy somatic cells, specifically in adult stem cells of the brain. The first evidence for ongoing mobility of L1 s in adult stem cells was obtained from neural precursor cells (NPCs) derived from rat hippocampus neural stem cells [112]. It was demonstrated that rodent NPCs can support elevated levels of retrotransposition using engineered human L1 reporter elements. By applying a transgenic mouse model it was demonstrated that L1 s generate genomic somatic mosaicism during brain development and in adult mice. Consistently, human NPCs isolated from fetal brain or derived from hESCs were shown to express moderate levels of L1 RNA and L1-ORF1p (Fig. 3), and to significantly support elevated levels of engineered L1 retrotransposition in vitro [143]. These and previous studies in mice (i.e., [112]) represented a major breakthrough in the TE field, because it was demonstrated for the first time that the endogenous L1 copy number is increased in several regions of the human brain (including hippocampus) when compared to other human somatic tissues [143, 144]. More recently, the use of an adenoviral/L1 hybrid vector [145] has allowed demonstrating L1 retrotransposition in non-dividing mature neuronal cells, differentiated from hESCs [99] (Fig. 3). While L1 s are usually repressed in somatic tissues, the mechanism responsible for L1 activation in neuronal progenitor cells has been studied (for review see [146, 147]). While the L1 promoter whose activity is known to be regulated by CpG methylation [93, 148], and is hypermethylated in mature neurons, it is less methylated than in fibroblasts [143] and these data suggest that additional factors beside DNA methylation might allow L1 to be expressed in neurons. The most accepted current model for the regulation of L1 expression and retrotransposition in neurons and NPCs involves the activity of Sox2 and WNT3A in combination with DNA methylation. Briefly, Sox2, a negative regulator of neuronal differentiation, is suggested to interact with the L1 promoter and repress L1 expression in rodent and human neural stem cells [112, 143]. Methyl-CpG-binding protein 2 (MeCP2) has been demonstrated to associate with the L1 promoter in a methylation-dependent manner, and to repress L1 expression in neural stem cells [149, 150]. Upon differentiation, Sox2 is downregulated, the promoter of L1 is mildly demethylated and thereby decreases binding of MeCP2 and increases L1 expression. At the same time, activation of WNT3A could

stimulate L1 expression through the canonical Wnt pathway [151].

Somatic L1 retrotransposition events accumulate during various developmental steps

Above studies mostly used engineered L1 reporter constructs, although more recently sequencing-based approaches have been used to demonstrate ongoing endogenous L1 retrotransposition in the human brain, not only in NPCs but also in neurons [144, 152–154]. Several high-throughput sequencing methods for the detection of endogenous retrotransposition events have been recently developed and applied to brain-derived samples (reviewed in [155]). Retrotransposon capture sequencing (RC-seq, [156]), and L1-seq [72] are two of the methods used to demonstrate L1 retrotransposition in neuronal cells isolated from cerebral cortex, caudate nucleus and hippocampus. Although these studies have reported significant variability in the inferred rate of L1 mobilization in the brain [144, 152–154, 157], the main conclusion from these studies is that somatic L1 retrotransposition events are accumulated during a variety of neural development stages, including early progenitor cells and mature neuronal cells. In summary, and while current estimates of L1 retrotransposition rates in neurons range from 0.04–0.6 to 13.7 L1 insertions per cell (see [111, 146, 157]), it is clear now that our brain is composed of a mosaic of genomes.

L1 retrotransposition occurs at very low level in the majority of tested adult stem cell types

Is L1 active in other somatic tissues besides the brain, or can all adult somatic stem cell types of the human body support L1 retrotransposition? In order to address this question, Macia and colleagues compared L1 expression and engineered L1 retrotransposition among various human cell types using isogenic cells [99]. To this end, the differentiation capability of hESCs was exploited to produce relatively pure populations of adult stem cells, which were then used to study L1 expression and retrotransposition. Notably, this approach allowed comparing different cell types that contain the same genetic background. Thus, L1 expression and retrotransposition was analyzed in keratinocytes (KERs, considered multipotent embryonic progenitor cells generating epidermal barrier, hair and nails), NPCs, fully differentiated neurons, MSCs and HSCs (Fig. 3). This allowed comparing NPCs and neurons to another ectoderm-derived cell type (KERs) and to multipotent stem cell types derived from mesoderm (MSCs and HSCs). Remarkably, data revealed that all cell types tested express endogenous L1-mRNAs and encoded L1-ORF1p, although there were significant differences in their expression levels (Fig. 3): L1-RNAs and L1-ORF1p were expressed at similar low levels in all cell types but NPCs,

which expressed moderate levels of gene products as revealed by RT-qPCR, immunoblot analysis and confocal microscopy [99]. Consistently, the absence of detectable full-length L1 transcripts from human bone marrow-derived MSCs was demonstrated [158]. Analysis of the DNA-methylation status of L1 promoters in the various differentiated cell types tested uncovered an inverse correlation with L1 expression levels [99], which is consistent with the hypothesis that DNA methylation is a major mechanism to inhibit L1 retrotransposition in somatic cells [159].

To investigate whether KERs, HFFs, neurons and multipotent NPCs, MSCs and HSCs support L1 retrotransposition, Macia and colleagues took advantage of a hybrid adenovirus-based L1 vector that was specifically designed to conduct experiments in non-dividing cells [160]. Consistent with L1 expression data, L1 retrotransposition in all cell types tested, was extremely low, except for NPCs (Fig. 3). However, this study confirmed elevated retrotransposition levels in NPCs, and demonstrated that non-dividing mature neurons support L1 retrotransposition using engineered L1 reporter elements in vitro. Thus, this data suggests that L1 activity is not a generic property of all somatic stem cell types of the human body. Intriguingly, a recent report using transgenic mice carrying a human engineered L1 reporter element suggests that ionizing radiation does increase L1 expression in mouse HSCs, and that these cells might express higher levels of L1 RNAs than human HSCs [161]. Low levels of L1 retrotransposition observed in non-irradiated HSCs were also increased upon irradiation [161]. However, it could not be ruled out that some or most of the L1 retrotransposition events had accumulated already during embryonic development [161].

The above studies suggest that, among the cell types analyzed, L1 retrotransposition in adult multipotent stem cells occurs at a very low level, with the exception of NPCs (Fig. 3). These data also indicate that, even at low levels, L1 activity in multipotent stem cells would result in new L1 insertions carried by the derived specific tissues. Consistently, several putative somatic L1 insertions have been identified in healthy tissues, including liver [162], stomach [163] and esophagus [164]. Although in this report, one de novo insertion has been validated in tissues of colon [163] and esophagus [164], respectively, additional research is required to investigate to what extent endogenous L1 mobilization occurs in healthy adult tissues or at any earlier developmental stage. Considering that (just as neuronal cells) the gastrointestinal tract is derived from endoderm, extending L1 mobilization analyses to adult stem cells derived from this germ layer may clarify if the capability to support L1 retrotransposition is restricted to NPCs or also applies to endodermal stem cell populations. It was shown recently that heritable L1 insertions occur not only in pluripotent embryonic cells, but also in early

Schumann *et al. Mobile DNA*   (2019) 10:9

Page 11 of 23

primordial germ cells (PGCs), at least in mice [113]. In humans, PGCs represent the primary undifferentiated stem cell type that differentiates towards gametes (spermatozoa and oocytes). Due to the therapeutic potential of PGCs for the treatment of infertility [165], the impact of L1 activity in those cells should be given due regard. In summary, among the adult stem cell types tested to date, somatic L1 retrotransposition is restricted to NPCs and PGCs, and does not seem to be a generic property of multipotent adult stem cells.

Regulation of TE activity in pluripotent stem cells

The human host is in a constant battle with TEs such as L1 to prevent their amplification and overreaching activity, because functional, active L1 elements carry substantial mutagenic potential. Somatic L1-mediated retrotransposition in somatic cells, such as stem cells, has also been linked to disease, especially in the context of cancer [75, 76, 162, 166]. Mobility restriction of these elements is therefore critical to maintain genome stability, especially during germline establishment and early embryogenesis, as new TE insertions in these cells can potentially be transmitted to the next generation [167].

During the first few days of embryogenesis, most TEs are targeted by silencing mechanisms and kept under control during early embryogenesis by epigenetic silencing and post-transcriptional regulation of L1 mRNAs [99, 168, 169]. Epigenetic silencing of TE expression is mediated by 1) Krüppel-associated box domain-containing zinc finger proteins (KRAB-ZFPs), mediators of heterochromatic formation [98] 2) cytosine methylation, a process governed by the action of DNA methyltransferase-3 like protein in germ cells [93], and 3) ten-eleven translocation (TET) family proteins [170]. TEs are recognized by sequence-specific proteins- or RNA-based repressors, with subsequent recruitment of heterochromatin-inducing complexes [171]. Histone methylation, histone deacetylation, and DNA methylation control TE expression and repress their *cis*-acting transcriptional components, which would otherwise activate neighboring genes via promoter or enhancer effects [172–178]. Histone repressive marks play a major role in this process, and trimethylation of histone 3 at lysine 9 (H3K9me3) is a modification found in a wide variety of TEs, including ERVs, LINEs, SINEs, and SVAs in hESCs [98, 179–184]. Excellent reviews have been published recently on how the host controls the activity of active TEs [67, 110, 155]. In the following section, we will address those host-encoded mechanisms that were reported to interfere with the retrotransposition cycle specifically in pluripotent stem cells.

Restriction of TE transcription

KRAB-ZFPs are the largest family of transcriptional regulators in higher vertebrates [185] and play a major role in the early embryonic control of vertebrate TE expression, including ERV, LINE, SINE and SVA elements [98, 178, 184]. Harbouring an N-terminal KRAB domain and a C-terminal array of DNA-binding zinc fingers, they mediate silencing by recruiting the cofactor KRAB-associated protein 1 (KAP1, also known as TRIM28), which acts as a docking protein for proteins with heterochromatin-forming activities, including the H3K9 methyltransferase SETDB1 [178, 186]. The KRAB/KAP1 system represses transcription of endogenous retroelements primarily via histone deacetylation, H3K9 trimethylation, and HP1 recruitment, with DNA methylation occurring only secondarily to ensure the permanence of the silencing process [187, 188].

Evolutionary old L1 subfamilies that are between ~ 7 and 25 million years of age are silenced in hESCs [98, 185, 189] through specific KRAB-ZFPs such as ZNF93/ KAP1, and there is evidence suggesting that KRAB-ZFPs also bind the younger and retrotransposition-competent L1Hs elements [185]. Beside L1s, in hESCs two-thirds of endogenous retroelements bound by KAP1 are SVAs, suggesting KAP1 is a major repressor of SVA activity. In summary, it is now well established that KRAB-ZFP/ KAP1 complexes control TE activity in hESCs, including retrotransposition-competent elements harbouring the potential to affect genomic integrity, and help maintaining transcriptional homeostasis and normal differentiation of ESCs.

DNA-Methylation at CpG dinucleotides plays a pivotal role in TE silencing in both somatic and germ cells, and has been suggested to have evolved primarily to protect the host against TEs. [93, 190, 191]. In mammals, CpG methylation is established by DNA methyltransferases DNMT3a, 3b, 3L and 3C and maintained by DNA methyltransferase DNMT1 and cofactors. DNMTs are required to reestablish TE DNA methylation after the two waves of reprogramming undergone by germ cells and early embryos [192–195]. DNA-methylation is known to control L1 expression, and the L1 promoter has a canonical CpG island whose methylation is inversely correlated with L1 expression [93, 148, 149, 196].

As early embryonic models, hESCs/hiPSCs show significant hypomethylation of L1 promoters, overexpress full-length L1 mRNA, L1 ORF1p and L1-RNPs [79, 96–101, 132, 133, 197, 198]. However, DNA methylation is also involved in the long-term transcriptional control of currently inactive L1 s, and in hESCs, L1PA4 and L1PA5 exhibit higher methylation levels than the currently active L1Hs elements [98]. Investigation of the correlation between CpG methylation and TE expression in naïve and primed human ESCs uncovered that TEs in naïve cells were hypomethylated relative to primed cells [128]. Despite the overall low methylation in naïve hESCs, endogenous SVAs, which represent the most overexpressed

TE family in these cells, tended to be more hypomethylated compared to non-overexpressed copies in naïve cells [128]. Thus, DNA methylation is key to control L1 expression and mobilization, and is the major TE controlling mechanism during gastrulation [125].

TET proteins: While the primary mechanism underlying DNA demethylation during the early embryonic period is replication-coupled dilution of 5mC [199, 200], an active mechanism dependent on TET enzymes has recently been uncovered to erase DNA methylation at specific loci in this period [170, 201–203]. TET enzymes catalyze the oxidation of 5-methylcytosine to 5-hydroxymethylcytosine (5hmC), and further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which can be replaced with unmodified cytosine by base excision repair [204, 205]. Some of the TET proteins are highly expressed in ESCs and blastocysts [206], and it was reported that TET binding and demethylation at particular TE classes, such as LTR retrotransposons, acts in concert with pluripotency factors Nanog, Oct4 and Sox2 to maintain expression of a subset of genes in ESCs [170]. Depletion of TET1 and TET2 in mESCs has been shown to cause loss of 5hmC in the 5′ region of L1 [207]. In hESCs, TET proteins were shown to preferentially bind to evolutionary young, functional L1 elements, and participate in their active demethylation, but do not interact with older, inactive subfamilies. Although TETs drive L1 demethylation, L1s can be kept repressed through the TET-dependent recruitment of the transcriptional repressor SIN3A. The SIN3A co-repressive complex binds functional L1s in mouse and human ESCs ensuring their repression in a TET-dependent manner [92, 170]. Thus, and instead of being only positive regulators of L1 expression, TET enzymes may have a dual role in TE regulation by also recruiting SIN3A to demethylated L1 elements. A recent report provided evidence that the methyl-CpG binding domain as well as the adjacent non-sequence specific DNA binding domain of MeCP2 mediate repression of TET1-induced L1 mobilization [208]. Also, the KZFP/KAP1 complex was recently reported to maintain heterochromatin and DNA methylation at TEs in naïve murine ESCs partly by protecting these loci from TET-mediated demethylation [209].

Small RNAs classes include PIWI-interacting RNAs (piRNAs), endogenously produced small interfering RNAs (siRNAs) or micro RNAs (miRNAs), and the importance of small RNA-based repression in the control of L1 expression in human pluripotent stem cells has been reported and reviewed [114, 210, 211]. Double-stranded RNAs (dsRNAs) with sequence specificity to portions of the transcriptome that originate from TEs with bidirectional promoters, such as L1s, was suggested to serve as substrate for the production of siRNAs [212]. Moreover, dsRNA produced from the 5'UTR of L1 may trigger the interferon-dependent restriction factor ribonuclease L in some cancer cells, leading to L1 mRNA degradation [213].

piRNAs are a complex class of small non-coding RNAs that comprise mostly 24–32 nucleotides, specifically interact with the PIWI protein subfamily of the ARGONAUTE family [214], and act to repress mobile genetic elements in the germline of *Drosophila* and mammals [169]. piRNAs are generated by transcription of long TE clusters, resulting in the accumulation of short mature piRNAs in the cytoplasm by the ping-pong mechanism [215]. piRNAs then act as guides to destroy complementary TE transcripts by endonucleolytic cleavage. PIWI-mediated control is indeed triggered by the recognition of L1-proximal sequences by a complex encompassing a member of the PIWI subclade of Argonaute proteins and L1-derived piRNAs, which leads to L1 transcriptional inhibition via DNA methylation [216–218]. The piRNAs–PIWI system and DNA methyltransferases, acting downstream of piRNA action, play a crucial role in the early embryonic control of the youngest and mobile L1 lineages in human pluripotent cells [98, 186, 211, 219].

A role for miRNAs and the miRNA biogenesis machinery in controlling human non-LTR retrotransposons has been suggested by demonstrating that a complex termed Microprocessor, which comprises the RNase III type enzyme Drosha and its partner DGCR8, catalyzes the nuclear step of microRNA biogenesis [220, 221] and binds L1, *Alu* and SVA-derived small RNAs in human cells [210]. The results suggest that Microprocessor recognizes and processes structural regions within retrotransposition-competent L1 and *Alu* elements leading to a decrease of functional L1 and *Alu* RNAs which could result in a reduction of L1 and *Alu* retrotransposition frequencies [169, 210]. miR-128, a member of the class of miRNAs encompassing 20- to 24-nt-long noncoding RNAs that inhibit translational initiation and stimulate decay of mRNA targets [222, 223], was recently reported to decrease both amounts of full-length L1 RNA and engineered L1 retrotransposition frequencies in a cell culture based assay in human tumor cells and induced pluripotent stem cells [198].

Post-transcriptional control of TEs
APOBEC3 proteins The human APOBEC3 (Apolipoprotein B mRNA Editing Enzyme Catalytic Polypeptide 3, A3) protein family of cytidine deaminases comprises seven members (APOBEC3A [A3A], APOBEC3B [A3B], APOBEC3C [A3C], APOBEC3DE [A3DE], APOBEC3F [A3F], APOBEC3G [A3G], APOBEC3H [A3H]) [224–227], that inhibit the human non-LTR retrotransposons L1 and *Alu* with varying degrees of efficiency [reviewed in [228–230]. While A3B, A3C, A3DE, A3F and A3G were shown to be expressed in hESCs, A3A is absent from hESCs [105], but transiently overexpressed by up

to 10-fold during reprogramming [132]. A3B is a nuclear protein that is also expressed in hiPSCs [105, 231, 232] and was demonstrated to restrict L1 retrotransposition in both hESCs [105] and hiPSCs in a deaminase-independent manner [211]. While it is likely that A3A inhibits L1 retrotransposition by deaminating transiently exposed L1 DNA [233], both A3C and A3DE were reported to restrict L1 retrotransposition through editing-independent mechanisms by interacting with ORF1p, thereby interfering with L1 RT activity [234, 235]. Additional mechanistic studies are required, because contradictory results were found among cell lines and studies.

SAMHD1, a dGTP-activated deoxynucleoside triphosphate triphosphohydrolase that inhibits L1 retrotransposition in actively dividing cells by reducing the L1 ORF2p level [236], was demonstrated to be transiently induced during the initial 10 days of reprogramming [132], and could therefore play a role in the restriction of L1 mobilization during this period. In addition to diminishing L1 reverse transcription by reducing L1 ORF2p expression [236], SAMHD1 was recently reported to stimulate the formation of stress granules and thus enhance the sequestration of L1 RNP complexes in these granules [237].

TEX19.1 is expressed in germ cells, pluripotent cells and the placenta of mammals [238] and was recently shown to regulate L1-ORF1p levels and mobilization of engineered L1 elements in pluripotent mouse embryonic stem cells [106]. Mouse TEX19.1, and its human ortholog TEX19, physically interact with L1-ORF1p, and can regulate L1-ORF1p abundance at postranslational level through stimulating its polyubiquitylation and proteasome-dependent degradation [106].

For the sake of completeness, we also mention the following host-encoded factors which also restrict L1 retrotransposition mostly by post-transcriptional mechanisms, but have not been reported to restrict L1 mobilization in stem cells so far: (i) MOV10, an RNA helicase that mediates access of the RNA-induced silencing complex to messenger RNAs [239, 240] (ii) ZAP, an antiviral member of the poly (ADP-ribose) polymerase (PARP) protein family [241, 242] (iii) TREX-1, a three-prime DNA exonuclease [243, 244].

DNA transposons as tools for genetic engineering of stem cells

An important point of intersection between stem cells and transposable elements is genetic engineering for the purposes of deciphering disease mechanisms and for gene- and cell-based therapeutics. DNA (Class 2) transposons (Fig. 4a) became major tools for genome manipulations in a wide range of cell types, including stem cells (reviewed by [245–247]. The mobility of DNA transposons invariably depends on two functional components: a protein component called the transposase and a DNA component called inverted terminal repeats (ITRs) supporting binding of and cleavage by the transposase (Fig. 4a). Due to their attributes including
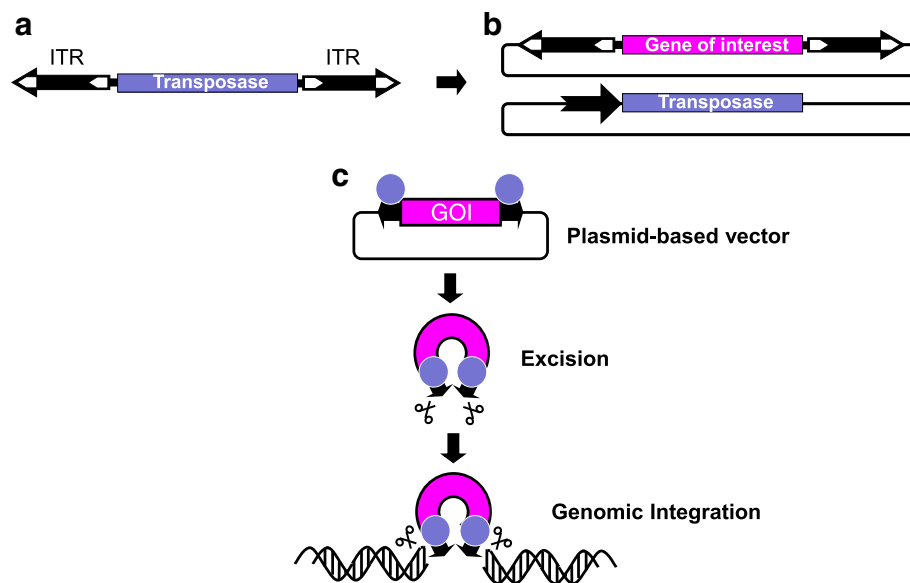

Fig. 4 Use of class II transposons as gene vectors. **a** Autonomous DNA transposons consist of a transposase-coding gene that is flanked by inverted terminal repeats (ITR; black arrows flanked by white arrows). **b** Bi-component transposon vector system for delivering plasmid-encoded transgenes. One component consists of a plasmid containing a gene of interest (GOI) flanked by transposon ITRs. The second component is a transposase expression plasmid. Black arrow, promoter driving expression of transposase gene. **c** The transposon carrying a GOI is excised from the donor plasmid and integrated at a chromosomal site by the transposase

Schumann *et al. Mobile DNA*   (2019) 10:9

Page 14 of 23

efficient gene transfer in human cell types, ease of experimental manipulation, and detailed understanding of the transposition process, the *Sleeping Beauty*, *piggyBac* and *Tol2* transposons have become the transposons of choice for stem cell engineering. What is common to all gene transfer applications by any transposon is the setup of a conditional, two-component gene delivery system, in which a gene of interest flanked by the transposon ITRs is mobilized out of standard plasmid vectors by the transposase that is conditionally and transiently expressed in the relevant cells (Fig. 4b and c). Stem cells are ideal targets for gene therapy applications. It is of great hope to consider stem cells to achieve tissue repair or to restore and replenish cells in the background of a genetic disease. Following a transposon-based genetic manipulation step to introduce a gene of interest, such as a therapeutic gene rendering stable phenotypic correction, stem cells can be expanded in vitro and then subjected to differentiation into particular cell lineages according to the specific therapeutic need. There is now widespread evidence for robust transposon-mediated gene transfer in several, clinically relevant stem cell types, such as ESCs, iPSCs, CD34+ hematopoietic stem cells (HSC), MSCs, or myoblasts.

Transgene expression in hESCs by means of the *Sleeping Beauty* transposon system

A proof of concept study established efficient and stable transgene expression in hESCs using the *Sleeping Beauty* transposon system. In this study, transposons, carrying different transgenes coding for genetic markers, were effectively delivered to undifferentiated hESCs. The source of transposase was either DNA or mRNA. Molecular analysis indicated that 98% of stable gene transfer resulted from transposition events. These genetically engineered hESCs were then differentiated into teratomas in vivo and mature hematopoietic cells in vitro and the progeny maintained stable transgene expression [248]. Similarly successful gene transfer in hESCs was obtained with the *piggyBac* system as well, where fluorescent reporters were introduced into ESCs, and these remained functional following in vitro differentiaton. The *piggyBac* transposon system also allowed for seamless restoration of the insertion site following transposon removal by a second round of transient transposase expression [249]. The *Sleeping Beauty* transposon was also used to introduce a CAG promoter-driven GFP expression cassette to hESCs. During differentiation experiments the CAG promoter yielded outstanding GFP expression in cardiomyocytes allowing for specific labeling of cardiomyocytes in a spontaneously differentiated, mixed cell population [250]. The same transposon was also used to introduce a genetically encoded Ca^{2+} ions indicator (GCaMP2) into hESCs allowing for real-time Ca imaging without the need for dyes. Cardiomyocytes differentiated

from these ESCs were characterized by their spontaneous contractions and Ca^{2+} signal oscillations, presenting a powerful tool for pharmacological screening assays [251].

Multiple applications of transposon systems in the hiPSC field

During the early developments in the iPSC field, a significant step was achieved when non-viral methods were used to achieve reprogramming. Both *piggyBac* [252] and *Sleeping Beauty* [253, 254] were successfully used to generate hiPSCs, and transposons remained relevant gene transfer tools in the context of hiPSCs generated from patient-derived cells with a disease-causing genetic background. In this scenario transposons can be effectively used to correct a disease phenotype by introducing functional genes into hiPSCs followed by differentiating the relevant cell types that can be used for therapeutic purposes. In a preclinical mouse model of Duchenne muscular dystrophy, hiPSCs were corrected with the *Sleeping Beauty* transposon system carrying the micro-utrophin gene. Cells were then differentiated into skeletal muscle progenitors. Following transplantation of these progenitors into dystrophic mice, engrafted muscles displayed large numbers of micro-utrophin-positive myofibers, with biochemically restored dystrophin-glycoprotein complex and improved contractile strength [255]. Transposon-based gene delivery is not only useful for correcting a genetic disease, but can also be applied to coax differentiation into a particular cell lineage. For example, drug-inducible expression of MYOD1 from a *piggyBac* transposon vector in hiPSCs for at least 5 days was demonstrated to lead to highly efficient differentiation into myocytes [256]. In a similar experiment, *Sleeping Beauty*-mediated overexpression of PAX3 in iPSCs induced differentiation into MYOD positive myogenic progenitors and multinucleated myofibers [257]. Transposons can also be used in combination with designer nucleases in iPSCs to correct gene defects. For example, the efficiency of endonuclease-based gene targeting can be enhanced by using the *piggyBac* transposon as an efficient, transient drug selection tool due to the possibility of seamless removal of the drug marker enabled by the re-transfection of the transposase [258]. In ß-thalassemia patient-derived iPSCs, *piggyBac* transposon-mediated puroΔtk-based drug selection was used in combination with CRISPR/Cas9 to achieve correction of the mutation in the Hemoglobin Beta Chain (HBB) gene [259]. A similar strategy was used in combination with CRISPR/Cas9 in Huntington disease iPSCs to correct mutations in the Huntigtin (HTT) gene. The corrected cells were then successfully differentiated into excitable, synaptically active forebrain neurons [260]. A different combination of CRISPR/Cas9 and *piggyBac* transposon demonstrated the versatility of combinations with different genome editing tools in iPSCs. The Cas9 gene driven by an inducible

promoter was delivered by the transposon, and genomic modification was achieved following sgRNA delivery. Following a transient transposase expression the inducible Cas9 cassette was removed yielding a genome-edited iPSC clone with seamless transgene removal [261].

Efficient gene transfer into mobilized HSCs and human CD34+ cord blood cells

A rapidly progressing area for transposon-mediated gene therapy applications is represented by HSCs. Transposon-based tools hold great promise for rendering efficient gene correction without the potential risks inherent to viral delivery methods, which resulted in severe adverse reactions in clinical trials in the past [262, 263]. HSCs have the potential for self-renewal and maintenance of the ability to differentiate into hematopoietic lineages, and are thus ideal targets for gene therapy applications in hematologic diseases. HSCs can be efficiently modified by the *Sleeping Beauty* and *piggyBac* transposon systems [264]. Especially, developments of *Sleeping Beauty* transposons were taking shape in recent years enabled by the highly efficient SB100X hyperactive transposase in CD34+ HSCs [265]. The *Sleeping Beauty* transposon system not only supports the efficient gene transfer into mobilized CD34+ HSCs, but also into human cord blood CD34+ cells as shown in a model of sickle cell disease [266]. Further developments of the transposon-mediated delivery were also undertaken in recent years to target CD34+ HSCs. For example, an interesting and promising optimization step was undertaken when the transposon was delivered to the cells as a minicircle in combination with transposase supplied as mRNA. This approach led to improved cell survival and reduced cytotoxicity in the HSCs providing also several biosafety advantages over conventional delivery methods [267]. The combination of the *Sleeping Beauty* transposon system and adenoviral vectors in a hybrid vector system was also proposed as a promising method to achieve in vivo gene delivery. This approach holds broader clinical application potential for gene therapy as it may circumvent the need for myeloablation and transplantation. In this study HSCs were mobilized into peripheral blood in a transgenic, humanized mouse model, and were targeted using a hybrid adenovirus/transposon vector system injected intravenously in vivo, resulting in functional HSCs homing back to the bone marrow stably expressing the transgene [268].

MSCs are also in the focus of regenerative medicine, however, there are no breakthrough applications yet, and the field is still hampered by many controversies. Nevertheless, it has been evidenced that both *Sleeping Beauty* and *piggyBac* transposon-based gene transfer is applicable in MSCs as well. The genetically engineered MSCs are still characterized by the ability to undergo osteogenic, myogenic, and adipogenic differentiation after modifications with the transposons [257, 269]. *piggyBac*-mediated gene transfer of IFNγ into adipose-derived MSCs was used in a mouse model of melanoma to show that the IFNγ-expressing MSCs engrafted into tumor stroma, inhibited tumor growth and angiogenesis, and prolonged the survival of mice [270].

Myoblasts are self-renewing adult muscle progenitor cells, which differentiate into skeletal muscle cells and could potentially be harnessed for cell therapy of muscle disorders. Both *Sleeping Beauty*- and *piggyBac*-mediated gene transfer methods can be applied to efficiently modify myoblasts and were shown to be useful tools in delivering therapeutic genes into myoblasts. Proper dysferlin expression as well as highly efficient engraftment of the engineered myoblasts was evidenced in the skeletal muscle of dysferlin-deficient mice [271], whereas transposon vectors encoding microdystrophins and delivered into myoblasts have been shown to yield proper expression levels in differentiated multinucleated myotubes [272].

Insertional mutagenesis and genetic labeling studies

DNA transposons have also been successfully used in stem cells to identify different mechanisms of stemness and differentiation in insertional mutagenesis and genetic labeling studies. For such applications, the transposon vectors are equipped with mutagenic and reporting features (e. g., gene traps) that allow conditional expression of a marker upon integration into a gene (reviewed in [246]. For example, a modified *Sleeping Beauty* transposon was generated to randomly trap genes in the neural stem cell genome and modify their expression or tag them with fluorescent markers and selectable genes, allowing recombinant cells to be isolated and expanded clonally. This approach may facilitate the identification of novel determinants of stem cell biology and neural cell fate specification in NSCs [273]. A modified *Tol2* transposon was also designed to allow for conditional disruption of a broad spectrum of genes. The system was used in mouse ESCs, and relied on differentially tagged *Tol2* transposons to discern individual integrations within a single cell [274]. A *piggyBac* transposon-based insertional mutagenesis method has been developed to efficiently generate genome-wide mutant libraries in mouse haploid ESCs [275]. In hESCs a *piggyBac* insertional mutagenesis screen was used to identify the role of nuclear RhoA during stem cell differentiation [276]. The *Sleeping Beauty* transposon has been also used to explore the clonal dynamics of native haematopoiesis in vivo allowing for a specific fate tracking approach based on in situ labelling of HSCs. DsRed-positive HSCs harboring distinct insertion sites resulted from inducible transposition events fully reconstituted myeloid and lymphoid lineages evidencing successful tagging. These experiments established that

Schumann *et al. Mobile DNA*   (2019) 10:9

Page 16 of 23

steady-state blood production is maintained by the successive recruitment of thousands of clones, each with a minimal contribution to mature progeny. These results demonstrated that large numbers of long-lived progenitors are the main drivers of steady-state haematopoiesis during most of adulthood [277].

The risk of genotoxicity caused by DNA transposon activity
As discussed above in the context of endogenous TEs, one of the most important risk factors also associated with the use of transposon-based gene transfer tools in stem cells is genotoxicity. In the context of a transposon vector system, at least two fundamental properties can contribute to genotoxicity: i) interaction of the transposase with endogenous human DNA sequences or human proteins with the transposon vector sequences and ii) the genome-wide insertion profile of the vector. With respect to "off-target" cleavage of the transposase, the use of the *Sleeping Beauty* and *Tol2* systems appears to be safe in human cells. Both of these transposons originate from fish genomes, and the mammalian lineage does not contain transposons sufficiently similar to allow cleavage by these transposases. Vice versa, human cells do not express proteins that could re-mobilize a genomically integrated *Sleeping Beauty* or *Tol2* vector. In contrast, the human PGBD5 transposase-derived protein was reported to mobilize *piggyBac* transposon vectors in human cells [278], thereby presenting potential implications for human applications [279].

Characterization of the target site selection properties of different vector systems is highly useful for ranking the different vector types and designs according to their genotoxic potential [280]. The *Sleeping Beauty* transposon displays the least deviation from random with respect to genome-wide distribution: no apparent bias was seen for either heterochromatin marks or euchromatin marks and only a weak correlation with transcriptional status of targeted genes was detected [281]. This is in marked contrast to target site distributions of several other transposons including *Tol2* [264, 282], and *piggyBac* [264, 283, 284] that favor integration into genes and near chromatin marks characteristic of active transcription units (e.g., H3K27 acetylation and H3K4 monomethylation). The *piggyBac* transposon, in particular, has been shown to favor open chromatin, expressed genes and TSSs (±5 kb) associated with DNaseI hypersensitive sites, H3K4Me3 marks and Pol II-bound regions in mouse and human cells [283–288]. These observations collectively suggest that *Sleeping Beauty* might be the safest currently available transposon for therapeutic gene delivery in clinical trials.

Conclusions
Stem cell therapies have been expected to bring substantial benefit to patients suffering a wide range of diseases and injuries. It was expected that the benefits of bone marrow transplants for patients needing reconstruction of their hematopoietic and immune system would apply to stem cell transplants of other cell types, and optimism has been high for the utilization of embryonic and induced pluripotent stem cells for a variety of applications. However, before these promising stem cells or their differentiated derivatives are administered to patients, genomic integrity of these cells has to be ensured to guarantee that these cells remain therapeutically functional and are not tumorigenic. In recent years, a remarkable amount of data accumulated showing that the activity of endogenous TEs can be one source of genomic destabilization in stem cells, and constitutes a risk for the biosafety of stem cell-based therapies. By giving an overview of the potentially mutagenic activity of TEs in human multipotent and pluripotent stem cells, the consequences of their activity for the genomic integrity and host gene expression, we provide arguments for a thorough characterization of TE activity and its consequences in the individual stem cell lines before their therapeutic utilization in patients in order to ensure biosafety of these stem cells cells and/or their applied derivatives.

Availability of data and materials
The dataset used to generate Fig. 1 is available in the repository of the US Food and Drug Administration (FDA), www.clinicaltrials.gov.

Authors' contributions
SRH, AS, ZI and GGS wrote the manuscript. PTR and NVHF generated the figures. All authors read and approved the submitted manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Division of Medical Biotechnology, Paul-Ehrlich-Institut, Paul-Ehrlich-Str.51-59, 63225 Langen, Germany. [2]Host-Pathogen Interactions, Paul-Ehrlich-Institut, Paul-Ehrlich-Str. 51-59, 63225 Langen, Germany. [3]GENYO. Centre for Genomics and Oncological Research, Pfizer/University of Granada/ Andalusian Regional Government, PTS Granada-Avenida de la Ilustración, 114, 18016 Granada, Spain. [4]Department of Biochemistry and Molecular

Biology II, Faculty of Pharmacy, University of Granada, Campus Universitario de Cartuja, 18071 Granada, Spain.

References

1. Mason C, Dunnill P. A brief definition of regenerative medicine. Regen Med. 2008;3:1–5. https://doi.org/10.2217/17460751.3.1.1.
2. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts // Embryonic Stem Cell Lines Derived from Human Blastocysts. Science. 1998; 282:1145–7. https://doi.org/10.1126/science.282.5391.1145.
3. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126: 663–76. https://doi.org/10.1016/j.cell.2006.07.024.
4. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318:1917–20. https://doi.org/10.1126/science.1151526.
5. Lodi D, Iannitti T, Palmieri B. Stem cells in clinical practice: Applications and warnings. J Exp Clin Cancer Res. 2011;30:9. https://doi.org/10.1186/1756-9966-30-9.
6. O'Donoghue K, Fisk NM. Fetal stem cells. Best Pract Res Clin Obstet Gynaecol. 2004;18:853–75. https://doi.org/10.1016/j.bpobgyn.2004.06.010.
7. Avots A, Harder F, Schmittwolf C, Petrovic S, Muller AM. Plasticity of hematopoietic stem cells and cellular memory. Immunol Rev. 2002;187:9–21. https://doi.org/10.1034/j.1600-065X.2002.18702.x.
8. Barriga F, Ramírez P, Wietstruck A, Rojas N. Hematopoietic stem cell transplantation: Clinical use and perspectives. Biol Res. 2012;45:307–16. https://doi.org/10.4067/S0716-97602012000300012.
9. Ullah I, Subbarao RB, Rho GJ. Human mesenchymal stem cells - current trends and future prospective. Biosci Rep. 2015. https://doi.org/10.1042/BSR20150025.
10. Chaudhary D, Trivedi RN, Kathuria A, Goswami TK, Khandia R, Munjal A. In vitro And In vivo Immunomodulating Properties of Mesenchymal Stem Cells. Recent Patents Inflamm Allergy Drug Discov. 2018;12:59–68. https://doi.org/10.2174/1872213X12666180227105924.
11. Squillaro T, Peluso G, Galderisi U. Clinical Trials With Mesenchymal Stem Cells: An Update. Cell Transplant. 2016;25:829–48. https://doi.org/10.3727/096368915X689622.
12. Trounson A, McDonald C. Stem Cell Therapies in Clinical Trials: Progress and Challenges. Cell Stem Cell. 2015;17:11–22. https://doi.org/10.1016/j.stem.2015.06.007.
13. Ratcliffe E, Glen KE, Naing MW, Williams DJ. Current status and perspectives on stem cell-based therapies undergoing clinical trials for regenerative medicine: Case studies. Br Med Bull. 2013;108:73–94. https://doi.org/10.1093/bmb/ldt034.
14. Kimbrel EA, Lanza R. Current status of pluripotent stem cells: Moving the first therapies to the clinic. Nat Rev Drug Discov. 2015;14:681–92. https://doi.org/10.1038/nrd4738.
15. Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: Somatic differentiation in vitro. Nat Biotechnol. 2000;18:399–404. https://doi.org/10.1038/74447.
16. Trounson A, DeWitt ND. Pluripotent stem cells progressing to the clinic. Nat Rev Mol Cell Biol. 2016;17:194–200. https://doi.org/10.1038/nrm.2016.10.
17. Mandai M, Kurimoto Y, Takahashi M. Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. N Engl J Med. 2017;377:792–3. https://doi.org/10.1056/NEJMc1706274.
18. Cyranoski D. Japanese man is first to receive 'reprogrammed' stem cells from another person. Nature. 2017. https://doi.org/10.1038/nature.2017.21730.
19. Ben-David U, Benvenisty N. Analyzing the genomic integrity of stem cells. Cambridge: StemBook; 2008. https://doi.org/10.3824/stembook.1.150.1.
20. Lefort N, Perrier AL, Laâbi Y, Varela C, Peschanski M. Human embryonic stem cells and genomic instability. Regen Med. 2009;4:899–909. https://doi.org/10.2217/rme.09.63.
21. Maitra A, Arking DE, Shivapurkar N, Ikeda M, Stastny V, Kassauei K, et al. Genomic alterations in cultured human embryonic stem cells. Nat Genet. 2005;37:1099–103. https://doi.org/10.1038/ng1631.
22. Martins-Taylor K, Xu R-H. Concise review: Genomic stability of human induced pluripotent stem cells. Stem Cells. 2012;30:22–7. https://doi.org/10.1002/stem.705.
23. Baker DEC, Harrison NJ, Maltby E, Smith K, Moore HD, Shaw PJ, et al. Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. Nat Biotechnol. 2007;25:207–15. https://doi.org/10.1038/nbt1285.
24. Ben-David U, Mayshar Y, Benvenisty N. Large-scale analysis reveals acquisition of lineage-specific chromosomal aberrations in human adult stem cells. Cell Stem Cell. 2011;9:97–102. https://doi.org/10.1016/j.stem.2011.06.013.
25. Gore A, Li Z, Fung H-L, Young JE, Agarwal S, Antosiewicz-Bourget J, et al. Somatic coding mutations in human induced pluripotent stem cells. Nature. 2011;471:63–7. https://doi.org/10.1038/nature09805.
26. Hussein SM, Batada NN, Vuoristo S, Ching RW, Autio R, Närvä E, et al. Copy number variation and selection during reprogramming to pluripotency. Nature. 2011;471:58–62. https://doi.org/10.1038/nature09871.
27. Laurent LC, Ulitsky I, Slavin I, Tran H, Schork A, Morey R, et al. Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. Cell Stem Cell. 2011;8:106–18. https://doi.org/10.1016/j.stem.2010.12.003.
28. Mayshar Y, Ben-David U, Lavon N, Biancotti J-C, Yakir B, Clark AT, et al. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell. 2010;7:521–31. https://doi.org/10.1016/j.stem.2010.07.017.
29. Liang G, Zhang Y. Genetic and epigenetic variations in iPSCs: Potential causes and implications for application. Cell Stem Cell. 2013;13:149–59. https://doi.org/10.1016/j.stem.2013.07.001.
30. Tapia N, Schöler HR. Molecular Obstacles to Clinical Translation of iPSCs. Cell Stem Cell. 2016;19:298–309. https://doi.org/10.1016/j.stem.2016.06.017.
31. Kazazian HH, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature. 1988;332:164–6. https://doi.org/10.1038/332164a0.
32. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 elements in structural variation and disease. Annu Rev Genomics Hum Genet. 2011;12:187–215. https://doi.org/10.1146/annurev-genom-082509-141802.
33. Marchi E, Kanapin A, Magiorkinis G, Belshaw R. Unfixed endogenous retroviral insertions in the human population. J Virol. 2014;88:9529–37. https://doi.org/10.1128/JVI.00919-14.
34. Naveira H, Bello X, Abal-Fabeiro JL, Maside X. Evidence for the persistence of an active endogenous retrovirus (ERVE) in humans. Genetica. 2014;142: 451–60. https://doi.org/10.1007/s10709-014-9789-y.
35. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921. https://doi.org/10.1038/35057062.
36. Speek M. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol. 2001;21:1973–85. https://doi.org/10.1128/MCB.21.6.1973-1985.2001.
37. Swergold GD. Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol. 1990;10:6718–29. https://doi.org/10.1128/MCB.10.12.6718.
38. Athanikar JN, Badge RM, Moran JV. A YY1-binding site is required for accurate human LINE-1 transcription initiation. Nucleic Acids Res. 2004;32: 3846–55. https://doi.org/10.1093/nar/gkh698.
39. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, et al. Human L1 retrotransposition: Cis preference versus trans complementation. Mol Cell Biol. 2001;21:1429–39. https://doi.org/10.1128/MCB.21.4.1429-1439.2001.
40. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH. High Frequency Retrotransposition in Cultured Mammalian Cells. Cell. 1996;87: 917–27. https://doi.org/10.1016/S0092-8674(00)81998-4.
41. Cost GJ. Human L1 element target-primed reverse transcription in vitro. EMBO J. 2002;21:5899–910. https://doi.org/10.1093/emboj/cdf592.
42. Luan DD, Korman MH, Jakubczak JL, Eickbush TH, Luan DD, Korman MH, et al. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. Cell. 1993;72:595–605. https://doi.org/10.1016/0092-8674(93)90078-5.
43. Denli AM, Narvaiza I, Kerman BE, Pena M, Benner C, Marchetto MCN, et al. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell. 2015;163:583–93. https://doi.org/10.1016/j.cell.2015.09.025.
44. Piskareva O, Schmatchenko V. DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro. FEBS Lett. 2006;580:661–8. https://doi.org/10.1016/j.febslet.2005.12.077.
45. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A. 2003;100:5280–5. https://doi.org/10.1073/pnas.0831042100.

46. Mills RE, Bennett EA, Iskow RC, Devine SE. Which transposable elements are active in the human genome? Trends Genet. 2007;23:183–91. https://doi.org/10.1016/j.tig.2007.02.006.

47. Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, et al. LINE-1 retrotransposition activity in human genomes. Cell. 2010;141:1159–70. https://doi.org/10.1016/j.cell.2010.05.021.

48. Lutz SM, Vincent BJ, Kazazian HH, Batzer MA, Moran JV. Allelic heterogeneity in LINE-1 retrotransposition activity. Am J Hum Genet. 2003;73:1431–7. https://doi.org/10.1086/379744.

49. Seleme MC, Vetter MR, Cordaux R, Bastone L, Batzer MA, Kazazian HH. Extensive individual variation in L1 retrotransposition capability contributes to human genetic diversity. Proc Natl Acad Sci U S A. 2006;103:6611–6. https://doi.org/10.1073/pnas.0601324103.

50. Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, DeBerardinis RJ, et al. Many human L1 elements are capable of retrotransposition. Nat Genet. 1997;16:37–43. https://doi.org/10.1038/ng0597-37.

51. Skowronski J, Singer MF. The abundant LINE-1 family of repeated DNA sequences in mammals // The Abundant LINE-1 Family of Repeated DNA Sequences in Mammals: Genes and pseudogenes // Genes and Pseudogenes. Cold Spring Harb Symp Quant Biol. 1986;51(Pt 1 // 51):457–64. https://doi.org/10.1101/SQB.1986.051.01.055.

52. Raiz J, Damert A, Chira S, Held U, Klawitter S, Hamdorf M, et al. The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. Nucleic Acids Res. 2012;40:1666–83. https://doi.org/10.1093/nar/gkr863.

53. Hancks DC, Goodier JL, Mandal PK, Cheung LE, Kazazian HH. Retrotransposition of marked SVA elements by human L1s in cultured cells. Hum Mol Genet. 2011;20:3386–400. https://doi.org/10.1093/hmg/ddr245.

54. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet. 2003;35:41–8. https://doi.org/10.1038/ng1223.

55. Ahl V, Keller H, Schmidt S, Weichenrieder O. Retrotransposition and Crystal Structure of an Alu RNP in the Ribosome-Stalling Conformation. Mol Cell. 2015;60:715–27. https://doi.org/10.1016/j.molcel.2015.10.003.

56. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. Nat Genet. 2000;24:363–7. https://doi.org/10.1038/74184.

57. Doucet AJ, Wilusz JE, Miyoshi T, Liu Y, Moran JV. A 3′ Poly(A) Tract Is Required for LINE-1 Retrotransposition. Mol Cell. 2015;60:728–41. https://doi.org/10.1016/j.molcel.2015.10.012.

58. Kramerov DA, Vassetzky NS. Origin and evolution of SINEs in eukaryotic genomes. Heredity (Edinb). 2011;107:487–95. https://doi.org/10.1038/hdy.2011.43.

59. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. Nat Rev Genet. 2002;3:370–9. https://doi.org/10.1038/nrg798.

60. Ono M, Kawakami M, Takezawa T. A novel human nonviral retroposon derived from an endogenous retrovirus. Nucleic Acids Res. 1987;15:8725–37.

61. Shen L, Wu LC, Sanlioglu S, Chen R, Mendoza AR, Dangel AW, et al. Structure and genetics of the partially duplicated gene RP located immediately upstream of the complement C4A and the C4B genes in the HLA class III region. Molecular cloning, exon-intron structure, composite retroposon, and breakpoint of gene duplication. J Biol Chem. 1994;269:8466–76.

62. Ostertag EM, Goodier JL, Zhang Y, Kazazian HH. SVA elements are nonautonomous retrotransposons that cause disease in humans. Am J Hum Genet. 2003;73:1444–51. https://doi.org/10.1086/380207.

63. Bennett EA, Coleman LE, Tsui C, Pittard WS, Devine SE. Natural genetic variation caused by transposable elements in humans. Genetics. 2004;168:933–51. https://doi.org/10.1534/genetics.104.031757.

64. Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, Batzer MA. SVA elements: A hominid-specific retroposon family. J Mol Biol. 2005;354:994–1007. https://doi.org/10.1016/j.jmb.2005.09.085.

65. Damert A, Raiz J, Horn AV, Löwer J, Wang H, Xing J, et al. 5′-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. Genome Res. 2009;19:1992–2008. https://doi.org/10.1101/gr.093435.109.

66. Hancks DC, Kazazian HH. SVA retrotransposons: Evolution and genetic instability. Semin Cancer Biol. 2010;20:234–45. https://doi.org/10.1016/j.semcancer.2010.04.001.

67. Friedli M, Trono D. The developmental control of transposable elements and the evolution of higher species. Annu Rev Cell Dev Biol. 2015;31:429–51. https://doi.org/10.1146/annurev-cellbio-100814-125514.

68. Kaer K, Speek M. Retroelements in human disease. Gene. 2013;518:231–41. https://doi.org/10.1016/j.gene.2013.01.008.

69. Cordaux R, Hedges DJ, Herke SW, Batzer MA. Estimating the retrotransposition rate of human Alu elements. Gene. 2006;373:134–7. https://doi.org/10.1016/j.gene.2006.01.019.

70. Hancks DC, Kazazian HH. Active human retrotransposons: Variation and disease. Curr Opin Genet Dev. 2012;22:191–203. https://doi.org/10.1016/j.gde.2012.02.006.

71. Hancks DC, Kazazian HH. Roles for retrotransposon insertions in human disease. Mob DNA. 2016;7:9. https://doi.org/10.1186/s13100-016-0065-9.

72. Ewing AD, Kazazian HH. High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. Genome Res. 2010;20:1262–70. https://doi.org/10.1101/gr.106419.110.

73. Kazazian HH, Moran JV. Mobile DNA in Health and Disease. N Engl J Med. 2017;377:361–70. https://doi.org/10.1056/NEJMra1510092.

74. Chen J-M, Masson E, Macek Jr M, Raguénès O, Piskackova T, Fercot B, Fila L, Cooper DN, Audrézet M-P, Férec C. Detection of two Alu insertions in the CFTR gene. J Cyst Fibros. 2008;7:37–43. https://doi.org/10.1016/j.jcf.2007.04.001.

75. Burns KH. Transposable elements in cancer. Nat Rev Cancer. 2017;17:415–24. https://doi.org/10.1038/nrc.2017.35.

76. Scott EC, Gardner EJ, Masood A, Chuang NT, Vertino PM, Devine SE. A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. Genome Res. 2016;26:745–55. https://doi.org/10.1101/gr.201814.115.

77. Wimmer K, Callens T, Wernstedt A, Messiaen L. The NF1 gene contains hotspots for L1 endonuclease-dependent de novo insertion. PLoS Genet. 2011;7:e1002371. https://doi.org/10.1371/journal.pgen.1002371.

78. Goodier JL, Kazazian HH. Retrotransposons revisited: The restraint and rehabilitation of parasites. Cell. 2008;135:23–35. https://doi.org/10.1016/j.cell.2008.09.022.

79. Klawitter S, Fuchs NV, Upton KR, Muñoz-Lopez M, Shukla R, Wang J, et al. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat Commun. 2016;7:10286. https://doi.org/10.1038/ncomms10286.

80. Garcia-Perez JL, Widmann TJ, Adams IR. The impact of transposable elements on mammalian development. Development. 2016;143:4101–14. https://doi.org/10.1242/dev.132639.

81. Perepelitsa-Belancio V, Deininger P. RNA truncation by premature polyadenylation attenuates human mobile element activity. Nat Genet. 2003;35:363–6. https://doi.org/10.1038/ng1269.

82. Piriyapongsa J, Polavarapu N, Borodovsky M, McDonald J. Exonization of the LTR transposable elements in human genome. BMC Genomics. 2007;8:291. https://doi.org/10.1186/1471-2164-8-291.

83. Piriyapongsa J, Rutledge MT, Patel S, Borodovsky M, Jordan IK. Evaluating the protein coding potential of exonized transposable element sequences. Biol Direct. 2007;2:31. https://doi.org/10.1186/1745-6150-2-31.

84. Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, et al. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell. 2013;152:453–66. https://doi.org/10.1016/j.cell.2012.12.023.

85. Han JS, Szak ST, Boeke JD. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature. 2004;429:268–74. https://doi.org/10.1038/nature02536.

86. Liu N, Lee CH, Swigut T, Grow E, Gu B, Bassik MC, Wysocka J. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. Nature. 2018;553:228–32. https://doi.org/10.1038/nature25179.

87. Belancio VP, Hedges DJ, Deininger P. LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res. 2006;34:1512–21. https://doi.org/10.1093/nar/gkl027.

88. Larson PA, Moldovan JB, Jasti N, Kidd JM, Beck CR, Moran JV. Spliced integrated retrotransposed element (SpIRE) formation in the human genome. PLoS Biol. 2018;16:e2003067. https://doi.org/10.1371/journal.pbio.2003067.

89. Moran JV, DeBerardinis RJ, Kazazian HH. Exon shuffling by L1 retrotransposition // Exon Shuffling by L1 Retrotransposition. Science. 1999;283:1530–4. https://doi.org/10.1126/science.283.5407.1530.

90. Goodier JL. Transduction of 3′-flanking sequences is common in L1 retrotransposition. Hum Mol Genet. 2000;9:653–7. https://doi.org/10.1093/hmg/9.4.653.

91. Pickeral OK. Frequent Human Genomic DNA Transduction Driven by LINE-1 Retrotransposition. Genome Res. 2000;10:411–5. https://doi.org/10.1101/gr.10.4.411.

92. Gerdes P, Richardson SR, Faulkner GJ. TET enzymes: Double agents in the transposable element-host genome conflict. Genome Biol. 2016;17:259. https://doi.org/10.1186/s13059-016-1124-8.

93. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature. 2004;431:96–9. https://doi.org/10.1038/nature02886.

94. Georgiou I, Noutsopoulos D, Dimitriadou E, Markopoulos G, Apergi A, Lazaros L, et al. Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes. Hum Mol Genet. 2009;18:1221–8. https://doi.org/10.1093/hmg/ddp022.

95. Ergün S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, et al. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. J Biol Chem. 2004;279:27753–63. https://doi.org/10.1074/jbc.M312985200.

96. Macia A, Muñoz-Lopez M, Cortes JL, Hastings RK, Morell S, Lucena-Aguilar G, et al. Epigenetic control of retrotransposon expression in human embryonic stem cells. Mol Cell Biol. 2011;31:300–16. https://doi.org/10.1128/MCB.00561-10.

97. Garcia-Perez JL, Marchetto MCN, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV. LINE-1 retrotransposition in human embryonic stem cells. Hum Mol Genet. 2007;16:1569–77. https://doi.org/10.1093/hmg/ddm105.

98. Castro-Diaz N, Ecco G, Coluccio A, Kapopoulou A, Yazdanpanah B, Friedli M, et al. Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev. 2014;28:1397–409. https://doi.org/10.1101/gad.241661.114.

99. Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaouad M, et al. Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Res. 2017;27:335–48. https://doi.org/10.1101/gr.206805.116.

100. Wissing S, Muñoz-Lopez M, Macia A, Yang Z, Montano M, Collins W, et al. Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. Hum Mol Genet. 2012;21:208–18. https://doi.org/10.1093/hmg/ddr455.

101. Muñoz-Lopez M, Macia A, Garcia-Cañadas M, Badge RM, Garcia-Perez JL. An epi c genetic battle: LINE-1 retrotransposons and intragenomic conflict in humans. Mob Genet Elements. 2011;1:122–7. https://doi.org/10.4161/mge.1.2.16730.

102. Muñoz-Lopez M, Garcia-Cañadas M, Macia A, Morell S, Garcia-Perez JL. Analysis of LINE-1 expression in human pluripotent cells. Methods Mol Biol. 2012;873:113–25. https://doi.org/10.1007/978-1-61779-794-1_7.

103. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell. 2011;8:676–87. https://doi.org/10.1016/j.stem.2011.04.004.

104. Philippe C, Vargas-Landin DB, Doucet AJ, van Essen D, Vera-Otarola J, Kuciak M, et al. Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci; 2016.

105. Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC. Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. J Biol Chem. 2011;286:36427–37. https://doi.org/10.1074/jbc.M111.251058.

106. MacLennan M, García-Cañadas M, Reichmann J, Khazina E, Wagner G, Playfoot CJ, et al. Mobilization of LINE-1 retrotransposons is restricted by Tex19.1 in mouse embryonic stem cells. Elife. 2017. https://doi.org/10.7554/eLife.26152.

107. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev. 2009;23:1303–12. https://doi.org/10.1101/gad.1803909.

108. An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, et al. Active retrotransposition by a synthetic L1 element in mice. Proc Natl Acad Sci U S A. 2006;103:18662–7. https://doi.org/10.1073/pnas.0605300103.

109. Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH. L1 integration in a transgenic mouse model. Genome Res. 2006;16:240–50. https://doi.org/10.1101/gr.4571606.

110. Levin HL, Moran JV. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet. 2011;12:615–27. https://doi.org/10.1038/nrg3030.

111. Faulkner GJ, Garcia-Perez JL. L1 Mosaicism in Mammals: Extent, Effects, and Evolution. Trends Genet. 2017;33:802–16. https://doi.org/10.1016/j.tig.2017.07.004.

112. Muotri AR, Chu VT, Marchetto MCN, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature. 2005;435:903–10. https://doi.org/10.1038/nature03663.

113. Richardson SR, Gerdes P, Gerhardt DJ, Sanchez-Luque FJ, Bodea G-O, Muñoz-Lopez M, et al. Heritable L1 retrotransposition in the mouse primordial germline and early embryo. Genome Res. 2017;27:1395–405. https://doi.org/10.1101/gr.219022.116.

114. Fadloun A, Le Gras S, Jost B, Ziegler-Birling C, Takahashi H, Gorab E, et al. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat Struct Mol Biol. 2013;20:332–8. https://doi.org/10.1038/nsmb.2495.

115. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell. 2008;31:785–99. https://doi.org/10.1016/j.molcel.2008.09.003.

116. van den Hurk JAJM, Meij IC, Seleme MC, Kano H, Nikopoulos K, Hoefsloot LH, et al. L1 retrotransposition can occur early in human embryonic development. Hum Mol Genet. 2007;16:1587–92. https://doi.org/10.1093/hmg/ddm108.

117. Ostertag EM. Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. 2000;28:1418–23. https://doi.org/10.1093/nar/28.6.1418.

118. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet. 2009;10:691–703. https://doi.org/10.1038/nrg2640.

119. Jachowicz JW, Bing X, Pontabry J, Bošković A, Rando OJ, Torres-Padilla M-E. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. Nat Genet. 2017;49:1502–10. https://doi.org/10.1038/ng.3945.

120. Percharde M, Lin C-J, Yin Y, Guan J, Peixoto GA, Bulut-Karslioglu A, et al. A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. Cell. 2018;174:391–405.e19. https://doi.org/10.1016/j.cell.2018.05.043.

121. Nichols J, Smith A. Naive and primed pluripotent states. Cell Stem Cell. 2009;4:487–92. https://doi.org/10.1016/j.stem.2009.05.015.

122. de LAA, Ferrari F, Xi R, Fujiwara Y, Benvenisty N, Deng H, et al. Hallmarks of pluripotency. Nature. 2015;525:469–78. https://doi.org/10.1038/nature15515.

123. Wu J, Izpisua Belmonte JC. Dynamic Pluripotent Stem Cell States and Their Applications. Cell Stem Cell. 2015;17:509–25. https://doi.org/10.1016/j.stem.2015.10.009.

124. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol. 2013;20:1131–9. https://doi.org/10.1038/nsmb.2660.

125. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. The DNA methylation landscape of human early embryos. Nature. 2014;511:606–10. https://doi.org/10.1038/nature13544.

126. Okamoto I, Patrat C, Thépot D, Peynot N, Fauque P, Daniel N, et al. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature. 2011;472:370–4. https://doi.org/10.1038/nature09872.

127. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, et al. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell. 2016;165:1012–26. https://doi.org/10.1016/j.cell.2016.03.023.

128. Theunissen TW, Friedli M, He Y, Planet E, O'Neil RC, Markoulaki S, et al. Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell. 2016;19:502–15. https://doi.org/10.1016/j.stem.2016.06.011.

129. Yamanaka S. Induced pluripotent stem cells: Past, present, and future. Cell Stem Cell. 2012;10:678–84. https://doi.org/10.1016/j.stem.2012.05.005.

130. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature. 2011;471:68–73. https://doi.org/10.1038/nature09798.

131. Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat Rev Cancer. 2011;11:268–77. https://doi.org/10.1038/nrc3034.

132. Friedli M, Turelli P, Kapopoulou A, Rauwel B, Castro-Díaz N, Rowe HM, et al. Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. Genome Res. 2014;24:1251–9. https://doi.org/10.1101/gr.172809.114.

133. Arokium H, Kamata M, Kim S, Kim N, Liang M, Presson AP, Chen IS. Deep sequencing reveals low incidence of endogenous LINE-1 retrotransposition in human induced pluripotent stem cells. PLoS One. 2014;9:e108682. https://doi.org/10.1371/journal.pone.0108682.

134. Quinlan AR, Boland MJ, Leibowitz ML, Shumilina S, Pehrson SM, Baldwin KK, Hall IM. Genome sequencing of mouse induced pluripotent stem cells reveals retroelement stability and infrequent DNA rearrangement during reprogramming. Cell Stem Cell. 2011;9:366–73. https://doi.org/10.1016/j.stem.2011.07.018.

135. Bhutani K, Nazor KL, Williams R, Tran H, Dai H, Džakula Ž, et al. Whole-genome mutational burden analysis of three pluripotency induction methods. Nat Commun. 2016;7:10536. https://doi.org/10.1038/ncomms10536.

136. Gilbert N, Lutz S, Morrish TA, Moran JV. Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol Cell Biol. 2005;25:7780–95. https://doi.org/10.1128/MCB.25.17.7780-7795.2005.

137. Gerdes P, Richardson SR, Mager DL, Faulkner GJ. Transposable elements in the mammalian embryo: Pioneers surviving through stealth and service. Genome Biol. 2016;17:100. https://doi.org/10.1186/s13059-016-0965-5.

138. Walter M, Teissandier A, Pérez-Palacios R, Bourc'his D. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells; 2016.

139. Mätlik K, Redik K, Speek M. L1 antisense promoter drives tissue-specific transcription of human genes. J Biomed Biotechnol. 2006;2006:71753. https://doi.org/10.1155/JBB/2006/71753.

140. Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell. 2012;151:1617–32. https://doi.org/10.1016/j.cell.2012.11.039.

141. Koyanagi-Aoi M, Ohnuki M, Takahashi K, Okita K, Noma H, Sawamura Y, et al. Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. Proc Natl Acad Sci U S A. 2013;110:20569–74. https://doi.org/10.1073/pnas.1319061110.

142. McClintock B. Controlling elements and the gene. Cold Spring Harb Symp Quant Biol. 1956;21:197–216.

143. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, et al. L1 retrotransposition in human neural progenitor cells. Nature. 2009;460:1127–31. https://doi.org/10.1038/nature08248.

144. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, de SF, et al. Somatic retrotransposition alters the genetic landscape of the human brain. Nature. 2011;479:534–7. https://doi.org/10.1038/nature10531.

145. Soifer H, Higo C, Kazazian HH, Moran JV, Mitani K, Kasahara N. Stable integration of transgenes delivered by a retrotransposon-adenovirus hybrid vector. Hum Gene Ther. 2001;12:1417–28. https://doi.org/10.1089/104303401750298571.

146. Richardson SR, Morell S, Faulkner GJ. L1 retrotransposons and somatic mosaicism in the brain. Annu Rev Genet. 2014;48:1–27. https://doi.org/10.1146/annurev-genet-120213-092412.

147. Faulkner GJ, Billon V. L1 retrotransposition in the soma: A field jumping ahead. Mob DNA. 2018;9:22. https://doi.org/10.1186/s13100-018-0128-1.

148. Bestor TH, Bourc'his D. Transposon silencing and imprint establishment in mammalian germ cells. New York: Cold Spring Harbor Laboratory Press; 2004.

149. Yu F, Zingler N, Schumann G, Strätling WH. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. 2001;29:4493–501. https://doi.org/10.1093/nar/29.21.4493.

150. Muotri AR, Marchetto MCN, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. L1 retrotransposition in neurons is modulated by MeCP2. Nature. 2010;468:443–6. https://doi.org/10.1038/nature09544.

151. Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, Lie DC, et al. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat Neurosci. 2009;12:1097–105. https://doi.org/10.1038/nn.2360.

152. Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sánchez-Luque FJ, Bodea GO, et al. Ubiquitous L1 mosaicism in hippocampal neurons. Cell. 2015;161:228–39. https://doi.org/10.1016/j.cell.2015.03.026.

153. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell. 2012;151:483–96. https://doi.org/10.1016/j.cell.2012.09.035.

154. Erwin JA, Paquola ACM, Singer T, Gallina I, Novotny M, Quayle C, et al. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. Nat Neurosci. 2016;19:1583–91. https://doi.org/10.1038/nn.4388.

155. Goodier JL. Restricting retrotransposons: A review. Mob DNA. 2016;7:16. https://doi.org/10.1186/s13100-016-0070-z.

156. Sanchez-Luque FJ, Richardson SR, Faulkner GJ. Retrotransposon Capture Sequencing (RC-Seq): A Targeted, High-Throughput Approach to Resolve Somatic L1 Retrotransposition in Humans. Methods Mol Biol. 2016;1400:47–77. https://doi.org/10.1007/978-1-4939-3372-3_4.

157. Evrony GD, Lee E, Park PJ, Walsh CA. Resolving rates of mutation in the brain using single-neuron genomics; 2016.

158. Belancio VP, Roy-Engel AM, Pochampally RR, Deininger P. Somatic expression of LINE-1 elements in human tissues. Nucleic Acids Res. 2010;38:3909–22. https://doi.org/10.1093/nar/gkq132.

159. Bestor TH. Cytosine methylation mediates sexual conflict. Trends Genet. 2003;19:185–90. https://doi.org/10.1016/S0168-9525(03)00049-0.

160. Kubo S, Seleme MC, Soifer HS, Perez JLG, Moran JV, Kazazian HH, Kasahara N. L1 retrotransposition in nondividing and primary human somatic cells. Proc Natl Acad Sci U S A. 2006;103:8036–41. https://doi.org/10.1073/pnas.0601954103.

161. Barbieri D, Elvira-Matelot E, Pelinski Y, Genève L, de LB, Yogarajah G, et al. Thrombopoietin protects hematopoietic stem cells from retrotransposon-mediated damage by promoting an antiviral response. J Exp Med. 2018;215:1463–80. https://doi.org/10.1084/jem.20170997.

162. Shukla R, Upton KR, Muñoz-Lopez M, Gerhardt DJ, Fisher ME, Nguyen T, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell. 2013;153:101–11. https://doi.org/10.1016/j.cell.2013.02.032.

163. Ewing AD, Gacita A, Wood LD, Ma F, Xing D, Kim M-S, et al. Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. Genome Res. 2015;25:1536–45. https://doi.org/10.1101/gr.196238.115.

164. Doucet-O'Hare TT, Rodić N, Sharma R, Darbari I, Abril G, Choi JA, et al. LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. Proc Natl Acad Sci U S A. 2015;112:E4894–900. https://doi.org/10.1073/pnas.1502474112.

165. Nikolic A, Volarevic V, Armstrong L, Lako M, Stojkovic M. Primordial Germ Cells: Current Knowledge and Perspectives. Stem Cells Int. 2016;2016:1741072. https://doi.org/10.1155/2016/1741072.

166. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ, et al. Landscape of somatic retrotransposition in human cancers. Science. 2012;337:967–71. https://doi.org/10.1126/science.1222077.

167. Crichton JH, Dunican DS, MacLennan M, Meehan RR, Adams IR. Defending the genome from the enemy within: Mechanisms of retrotransposon suppression in the mouse germline. Cell Mol Life Sci. 2014;71:1581–605. https://doi.org/10.1007/s00018-013-1468-0.

168. Molaro A, Malik HS. Hide and seek: How chromatin-based pathways silence retroelements in the mammalian germline. Curr Opin Genet Dev. 2016;37:51–8. https://doi.org/10.1016/j.gde.2015.12.001.

169. Heras SR, Macias S, Cáceres JF, Garcia-Perez JL. Control of mammalian retrotransposons by cellular RNA processing activities. Mob Genet Elements. 2014;4:e28439. https://doi.org/10.4161/mge.28439.

170. de La RL, Deniz Ö, Cheng KCL, Todd CD, Cruz C, Houseley J, Branco MR. TET-dependent regulation of retrotransposable elements in mouse embryonic stem cells. Genome Biol. 2016;17:234. https://doi.org/10.1186/s13059-016-1096-8.

171. Rowe HM, Trono D. Dynamic control of endogenous retroviruses during development. Virology. 2011;411:273–87. https://doi.org/10.1016/j.virol.2010.12.007.

172. Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, et al. Evolution of the mammalian transcription factor binding repertoire via transposable elements. Genome Res. 2008;18:1752–62. https://doi.org/10.1101/gr.080663.108.

173. Kunarso G, Chia N-Y, Jeyakani J, Hwang C, Lu X, Chan Y-S, et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat Genet. 2010;42:631–4. https://doi.org/10.1038/ng.600.

174. Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K, Wang J, et al. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. Genes Dev. 2011;25:594–607. https://doi.org/10.1101/gad.2008511.

175. Rebollo R, Miceli-Royer K, Zhang Y, Farivar S, Gagnier L, Mager DL. Epigenetic interplay between mouse endogenous retroviruses and host genes. Genome Biol. 2012;13:R89. https://doi.org/10.1186/gb-2012-13-10-r89.

176. Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C, et al. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell. 2012;148:335–48. https://doi.org/10.1016/j.cell.2011.11.058.

177. Chuong EB, Rumi MAK, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet. 2013;45:325–9. https://doi.org/10.1038/ng.2553.

178. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, et al. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature. 2010;463:237–40. https://doi.org/10.1038/nature08674.

179. Bulut-Karslioglu A, de La R-VIA, Ramirez F, Barenboim M, Onishi-Seebacher M, Arand J, et al. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. Mol Cell. 2014;55:277–90. https://doi.org/10.1016/j.molcel.2014.05.029.

180. Kondo Y, J-PJ I. Enrichment for histone H3 lysine 9 methylation at Alu repeats in human cells. J Biol Chem. 2003;278:27658–62. https://doi.org/10.1074/jbc.M304072200.

181. Martens JHA, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J. 2005;24:800–12. https://doi.org/10.1038/sj.emboj.7600545.

182. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007;448:553–60. https://doi.org/10.1038/nature06008.

183. Pezic D, Manakov SA, Sachidanandam R, Aravin AA. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev. 2014;28:1410–28. https://doi.org/10.1101/gad.240895.114.

184. Turelli P, Castro-Diaz N, Marzetta F, Kapopoulou A, Raclot C, Duc J, et al. Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements. Genome Res. 2014;24:1260–70. https://doi.org/10.1101/gr.172833.114.

185. Ecco G, Imbeault M, Trono D. KRAB zinc finger proteins. Development. 2017;144:2719–29. https://doi.org/10.1242/dev.132605.

186. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature. 2010;464:927–31. https://doi.org/10.1038/nature08858.

187. Quenneville S, Turelli P, Bojkowska K, Raclot C, Offner S, Kapopoulou A, Trono D. The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. Cell Rep. 2012;2:766–73. https://doi.org/10.1016/j.celrep.2012.08.043.

188. Rowe HM, Friedli M, Offner S, Verp S, Mesnard D, Marquis J, et al. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. Development. 2013;140:519–29. https://doi.org/10.1242/dev.087585.

189. Jacobs FMJ, Greenberg D, Nguyen N, Haeussler M, Ewing AD, Katzman S, et al. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature. 2014;516:242–5. https://doi.org/10.1038/nature13760.

190. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet. 1998;20:116–7. https://doi.org/10.1038/2413.

191. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 1997;13:335–40. https://doi.org/10.1016/S0168-9525(97)01181-5.

192. Chedin F, Lieber MR, Hsieh C-L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc Natl Acad Sci U S A. 2002;99:16916–21. https://doi.org/10.1073/pnas.262443999.

193. Okano M, Bell DW, Haber DA, Li E. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell. 1999;99:247–57. https://doi.org/10.1016/S0092-8674(00)81656-6.

194. Suetake I, Shinozaki F, Miyagawa J, Takeshima H, Tajima S. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem. 2004;279:27816–23. https://doi.org/10.1074/jbc.M400181200.

195. Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Hérault Y, et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science. 2016;354:909–12. https://doi.org/10.1126/science.aah5143.

196. Thayer RE, Singer MF, Fanning TG. Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1-encoded protein. Gene. 1993;133:273–7. https://doi.org/10.1016/0378-1119(93)90651-I.

197. Shen Y, Chow J, Wang Z, Fan G. Abnormal CpG island methylation occurs during in vitro differentiation of human embryonic stem cells. Hum Mol Genet. 2006;15:2623–35. https://doi.org/10.1093/hmg/ddl188.

198. Hamdorf M, Idica A, Zisoulis DG, Gamelin L, Martin C, Sanders KJ, Pedersen IM. miR-128 represses L1 retrotransposition by binding directly to L1 RNA. Nat Struct Mol Biol. 2015;22:824–31. https://doi.org/10.1038/nsmb.3090.

199. von Meyenn F, Iurlaro M, Habibi E, Liu NQ, Salehzadeh-Yazdi A, Santos F, et al. Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. Mol Cell. 2016;62:848–61. https://doi.org/10.1016/j.molcel.2016.04.025.

200. Arand J, Wossidlo M, Lepikhov K, Peat JR, Reik W, Walter J. Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. Epigenetics Chromatin. 2015;8(1). https://doi.org/10.1186/1756-8935-8-1.

201. Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell. 2014;15:459–71. https://doi.org/10.1016/j.stem.2014.09.002.

202. Peat JR, Dean W, Clark SJ, Krueger F, Smallwood SA, Ficz G, et al. Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. Cell Rep. 2014;9:1990–2000. https://doi.org/10.1016/j.celrep.2014.11.034.

203. Guo F, Li X, Liang D, Li T, Zhu P, Guo H, et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell. 2014;15:447–59. https://doi.org/10.1016/j.stem.2014.08.003.

204. Ito S, Kuraoka I. Epigenetic modifications in DNA could mimic oxidative DNA damage: A double-edged sword. DNA Repair (Amst). 2015;32:52–7. https://doi.org/10.1016/j.dnarep.2015.04.013.

205. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502:472–9. https://doi.org/10.1038/nature12750.

206. Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y, et al. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res. 2011;21:1332–42. https://doi.org/10.1038/cr.2011.113.

207. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473:398–402. https://doi.org/10.1038/nature10008.

208. Zhang P, Ludwig AK, Hastert FD, Rausch C, Lehmkuhl A, Hellmann I, et al. L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins. Nucleus. 2017;8:548–62. https://doi.org/10.1080/19491034.2017.1330238.

209. Coluccio A, Ecco G, Duc J, Offner S, Turelli P, Trono D. Individual retrotransposon integrants are differentially controlled by KZFP/KAP1-dependent histone methylation, DNA methylation and TET-mediated hydroxymethylation in naive embryonic stem cells. Epigenetics Chromatin. 2018;11:7. https://doi.org/10.1186/s13072-018-0177-1.

210. Heras SR, Macias S, Plass M, Fernandez N, Cano D, Eyras E, et al. The Microprocessor controls the activity of mammalian retrotransposons. Nat Struct Mol Biol. 2013;20:1173–81. https://doi.org/10.1038/nsmb.2658.

211. Marchetto MCN, Narvaiza I, Denli AM, Benner C, Lazzarini TA, Nathanson JL, et al. Differential L1 regulation in pluripotent stem cells of humans and apes. Nature. 2013;503:525–9. https://doi.org/10.1038/nature12686.

212. Yang N, Kazazian HH. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol. 2006;13:763–71. https://doi.org/10.1038/nsmb1141.

213. Zhang A, Dong B, Doucet AJ, Moldovan JB, Moran JV, Silverman RH. RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic Acids Res. 2014;42:3803–20. https://doi.org/10.1093/nar/gkt1308.

214. Ku H-Y, Lin H. PIWI proteins and their interactors in piRNA biogenesis, germline development and gene expression. Natl Sci Rev. 2014;1:205–18. https://doi.org/10.1093/nsr/nwu014.

215. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: The vanguard of genome defence. Nat Rev Mol Cell Biol. 2011;12:246–58. https://doi.org/10.1038/nrm3089.

216. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science. 2007;318:761–4. https://doi.org/10.1126/science.1146484.

217. Carmell MA, Girard A, van de Kant HJG, Bourc'his D, Bestor TH, de RDG, Hannon GJ. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell. 2007;12:503–14. https://doi.org/10.1016/j.devcel.2007.03.001.

218. de FS, Bartonicek N, Di Giacomo M, Abreu-Goodger C, Sankar A, Funaya C, et al. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature. 2011;480:259–63. https://doi.org/10.1038/nature10547.

219. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S-i, Sakaue M, Matsuoka C, et al. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells. 2006;11:805–14. https://doi.org/10.1111/j.1365-2443.2006.00984.x.

220. Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature. 2004;432:231–5. https://doi.org/10.1038/nature03049.

221. Han J, Lee Y, Yeom K-H, Kim Y-K, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. 2004;18:3016–27. https://doi.org/10.1101/gad.1262504.

222. Chekulaeva M, Filipowicz W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. Curr Opin Cell Biol. 2009;21:452–60. https://doi.org/10.1016/j.ceb.2009.04.009.

223. Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell. 2004;116:281–97.

224. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics. 2002;79:285–96. https://doi.org/10.1006/geno.2002.6718.

225. Conticello SG, Harris RS, Neuberger MS. The Vif Protein of HIV Triggers Degradation of the Human Antiretroviral DNA Deaminase APOBEC3G. Curr Biol. 2003;13:2009–13. https://doi.org/10.1016/j.cub.2003.10.034.

226. Conticello SG, Thomas CJF, Petersen-Mahrt SK, Neuberger MS. Evolution of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases. Mol Biol Evol. 2005;22:367–77. https://doi.org/10.1093/molbev/msi026.

227. Salter JD, Bennett RP, Smith HC. The APOBEC Protein Family: United by Structure, Divergent in Function. Trends Biochem Sci. 2016;41:578–94. https://doi.org/10.1016/j.tibs.2016.05.001.

228. Schumann GG, Gogvadze EV, Osanai-Futahashi M, Kuroki A, Münk C, Fujiwara H, et al. Unique functions of repetitive transcriptomes. Int Rev Cell Mol Biol. 2010;285:115–88. https://doi.org/10.1016/B978-0-12-381047-2.00003-7.

229. Schumann GG. APOBEC3 proteins: Major players in intracellular defence against LINE-1-mediated retrotransposition. Biochem Soc Trans. 2007;35: 637–42. https://doi.org/10.1042/BST0350637.

230. Orecchini E, Frassinelli L, Galardi S, Ciafrè SA, Michienzi A. Post-transcriptional regulation of LINE-1 retrotransposition by AID/APOBEC and ADAR deaminases. Chromosom Res. 2018. https://doi.org/10.1007/s10577-018-9572-5.

231. Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, Kurata T, et al. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Res. 2007;35:2955–64. https://doi.org/10.1093/nar/gkm181.

232. Muckenfuss H, Hamdorf M, Held U, Perkovic M, Löwer J, Cichutek K, et al. APOBEC3 proteins inhibit human LINE-1 retrotransposition. J Biol Chem. 2006;281:22161–72. https://doi.org/10.1074/jbc.M601716200.

233. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition; 2014.

234. Horn AV, Klawitter S, Held U, Berger A, Vasudevan AAJ, Bock A, et al. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. Nucleic Acids Res. 2014;42:396–416. https://doi.org/10.1093/nar/gkt898.

235. Liang W, Xu J, Yuan W, Song X, Zhang J, Wei W, et al. APOBEC3DE Inhibits LINE-1 Retrotransposition by Interacting with ORF1p and Influencing LINE Reverse Transcriptase Activity. PLoS One. 2016;11:e0157220. https://doi.org/10.1371/journal.pone.0157220.

236. Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, et al. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutières syndrome-related SAMHD1. Cell Rep. 2013;4:1108–15. https://doi.org/10.1016/j.celrep.2013.08.019.

237. Hu S, Li J, Xu F, Mei S, Le Duff Y, Yin L, et al. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. PLoS Genet. 2015;11:e1005367. https://doi.org/10.1371/journal.pgen.1005367.

238. Kuntz S, Kieffer E, Bianchetti L, Lamoureux N, Fuhrmann G, Viville S. Tex19, a mammalian-specific protein with a restricted expression in pluripotent stem cells and germ line. Stem Cells. 2008;26:734–44. https://doi.org/10.1634/stemcells.2007-0772.

239. Goodier JL, Cheung LE, Kazazian HH. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. PLoS Genet. 2012;8:e1002941. https://doi.org/10.1371/journal.pgen.1002941.

240. Li X, Zhang J, Jia R, Cheng V, Xu X, Qiao W, et al. The MOV10 helicase inhibits LINE-1 mobility. J Biol Chem. 2013;288:21148–60. https://doi.org/10.1074/jbc.M113.465856.

241. Goodier JL, Pereira GC, Cheung LE, Rose RJ, Kazazian HH. The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition. PLoS Genet. 2015;11:e1005252. https://doi.org/10.1371/journal.pgen.1005252.

242. Moldovan JB, Moran JV. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS Genet. 2015;11:e1005121. https://doi.org/10.1371/journal.pgen.1005121.

243. Li P, Du J, Goodier JL, Hou J, Kang J, Kazazian HH, et al. Aicardi-Goutières syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion. Nucleic Acids Res. 2017;45:4619–31. https://doi.org/10.1093/nar/gkx178.

244. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell. 2008;134:587–98. https://doi.org/10.1016/j.cell.2008.06.032.

245. Ivics Z, Li MA, Mátés L, Boeke JD, Nagy A, Bradley A, Izsvák Z. Transposon-mediated genome manipulation in vertebrates. Nat Methods. 2009;6:415–22. https://doi.org/10.1038/nmeth.1332.

246. Kawakami K, Largaespada DA, Ivics Z. Transposons As Tools for Functional Genomics in Vertebrate Models. Trends Genet. 2017;33:784–801. https://doi.org/10.1016/j.tig.2017.07.006.

247. Kebriaei P, Izsvák Z, Narayanavari SA, Singh H, Ivics Z. Gene Therapy with the Sleeping Beauty Transposon System. Trends Genet. 2017;33:852–70. https://doi.org/10.1016/j.tig.2017.08.008.

248. Wilber A, Linehan JL, Tian X, Woll PS, Morris JK, Belur LR, et al. Efficient and stable transgene expression in human embryonic stem cells using transposon-mediated gene transfer. Stem Cells. 2007;25:2919–27. https://doi.org/10.1634/stemcells.2007-0026.

249. Chen Y-T, Furushima K, Hou P-S, Ku AT, Deng JM, Jang C-W, et al. PiggyBac transposon-mediated, reversible gene transfer in human embryonic stem cells. Stem Cells Dev. 2010;19:763–71. https://doi.org/10.1089/scd.2009.0118.

250. Orbán TI, Apáti A, Németh A, Varga N, Krizsik V, Schamberger A, et al. Applying a "double-feature" promoter to identify cardiomyocytes differentiated from human embryonic stem cells following transposon-based gene delivery. Stem Cells. 2009;27:1077–87. https://doi.org/10.1002/stem.45.

251. Apáti Á, Pászty K, Hegedűs L, Kolacsek O, Orbán TI, Erdei Z, et al. Characterization of calcium signals in human embryonic stem cells and in their differentiated offspring by a stably integrated calcium indicator protein. Cell Signal. 2013;25:752–9. https://doi.org/10.1016/j.cellsig.2012.12.024.

252. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hämäläinen R, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature. 2009;458:766–70. https://doi.org/10.1038/nature07863.

253. Grabundzija I, Wang J, Sebe A, Erdei Z, Kajdi R, Devaraj A, et al. Sleeping Beauty transposon-based system for cellular reprogramming and targeted gene insertion in induced pluripotent stem cells. Nucleic Acids Res. 2013;41: 1829–47. https://doi.org/10.1093/nar/gks1305.

254. Sebe A, Ivics Z. Reprogramming of Human Fibroblasts to Induced Pluripotent Stem Cells with Sleeping Beauty Transposon-Based Stable Gene Delivery. Methods Mol Biol. 2016;1400:419–27. https://doi.org/10.1007/978-1-4939-3372-3_26.

255. Filareto A, Parker S, Darabi R, Borges L, Iacovino M, Schaaf T, et al. An ex vivo gene therapy approach to treat muscular dystrophy using inducible pluripotent stem cells. Nat Commun. 2013;4:1549. https://doi.org/10.1038/ncomms2550.

256. Tanaka A, Woltjen K, Miyake K, Hotta A, Ikeya M, Yamamoto T, et al. Efficient and reproducible myogenic differentiation from human iPS cells: Prospects for modeling Miyoshi Myopathy in vitro. PLoS One. 2013;8:e61540. https://doi.org/10.1371/journal.pone.0061540.

257. Belay E, Mátrai J, Acosta-Sanchez A, Ma L, Quattrocelli M, Mátés L, et al. Novel hyperactive transposons for genetic modification of induced pluripotent and adult stem cells: A nonviral paradigm for coaxed differentiation. Stem Cells. 2010;28:1760–71. https://doi.org/10.1002/stem.501.

258. Yusa K. Seamless genome editing in human pluripotent stem cells using custom endonuclease-based gene targeting and the piggyBac transposon. Nat Protoc. 2013;8:2061–78. https://doi.org/10.1038/nprot.2013.126.

259. Xie F, Ye L, Chang JC, Beyer AI, Wang J, Muench MO, Kan YW. Seamless gene correction of β-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. Genome Res. 2014;24:1526–33. https://doi.org/10.1101/gr.173427.114.

260. Xu X, Tay Y, Sim B, Yoon S-I, Huang Y, Ooi J, et al. Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. Stem Cell Reports. 2017;8:619–33. https://doi.org/10.1016/j.stemcr.2017.01.022.

261. Wang G, Yang L, Grishin D, Rios X, Ye LY, Hu Y, et al. Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies. Nat Protoc. 2017;12:88–103. https://doi.org/10.1038/nprot.2016.152.

262. Hacein-Bey-Abina S, von KC, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science. 2003;302:415–9. https://doi.org/10.1126/science.1088547.

263. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest. 2008;118:3132–42. https://doi.org/10.1172/JCI35700.

264. Grabundzija I, Irgang M, Mátés L, Belay E, Matrai J, Gogol-Döring A, et al. Comparative analysis of transposable element vector systems in human cells. Mol Ther. 2010;18:1200–9. https://doi.org/10.1038/mt.2010.47.

265. Mátés L, Chuah MKL, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet. 2009;41:753–61. https://doi.org/10.1038/ng.343.

266. Sjeklocha LM, Wong PY-P, Belcher JD, Vercellotti GM, Steer CJ. β-Globin sleeping beauty transposon reduces red blood cell sickling in a patient-derived CD34(+)-based in vitro model. PLoS One. 2013;8:e80403. https://doi.org/10.1371/journal.pone.0080403.

267. Holstein M, Mesa-Nuñez C, Miskey C, Almarza E, Poletti V, Schmeer M, et al. Efficient Non-viral Gene Delivery into Human Hematopoietic Stem Cells by Minicircle Sleeping Beauty Transposon Vectors. Mol Ther. 2018;26:1137–53. https://doi.org/10.1016/j.ymthe.2018.01.012.

268. Richter M, Saydaminova K, Yumul R, Krishnan R, Liu J, Nagy E-E, et al. In vivo transduction of primitive mobilized hematopoietic stem cells after intravenous injection of integrating adenovirus vectors. Blood. 2016;128:2206–17. https://doi.org/10.1182/blood-2016-04-711580.

269. Wen S, Zhang H, Li Y, Wang N, Zhang W, Yang K, et al. Characterization of constitutive promoters for piggyBac transposon-mediated stable transgene expression in mesenchymal stem cells (MSCs). PLoS One. 2014;9:e94397. https://doi.org/10.1371/journal.pone.0094397.

270. Bahrambeigi V, Ahmadi N, Moisyadi S, Urschitz J, Salehi R, Haghjooy Javanmard S. PhiC31/PiggyBac modified stromal stem cells: Effect of interferon γ and/or tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) on murine melanoma. Mol Cancer. 2014;13:255. https://doi.org/10.1186/1476-4598-13-255.

271. Escobar H, Schöwel V, Spuler S, Marg A, Izsvák Z. Full-length Dysferlin Transfer by the Hyperactive Sleeping Beauty Transposase Restores Dysferlin-deficient Muscle. Mol Ther Nucleic Acids. 2016;5:e277. https://doi.org/10.1038/mtna.2015.52.

272. Loperfido M, Jarmin S, Dastidar S, Di Matteo M, Perini I, Moore M, et al. piggyBac transposons expressing full-length human dystrophin enable genetic correction of dystrophic mesoangioblasts. Nucleic Acids Res. 2016;44:744–60. https://doi.org/10.1093/nar/gkv1464.

273. Albieri I, Onorati M, Calabrese G, Moiana A, Biasci D, Badaloni A, et al. A DNA transposon-based approach to functional screening in neural stem cells. J Biotechnol. 2010;150:11–21. https://doi.org/10.1016/j.jbiotec.2010.07.027.

274. Mayasari NI, Mukougawa K, Shigeoka T, Kawakami K, Kawaichi M, Ishida Y. Mixture of differentially tagged Tol2 transposons accelerates conditional disruption of a broad spectrum of genes in mouse embryonic stem cells. Nucleic Acids Res. 2012;40:e97. https://doi.org/10.1093/nar/gks262.

275. Pettitt SJ, Tan E-P, Yusa K. piggyBac transposon-based insertional mutagenesis in mouse haploid embryonic stem cells. Methods Mol Biol. 2015;1239:15–28. https://doi.org/10.1007/978-1-4939-1862-1_2.

276. Gayle S, Pan Y, Landrette S, Xu T. piggyBac insertional mutagenesis screen identifies a role for nuclear RHOA in human ES cell differentiation. Stem Cell Reports. 2015;4:926–38. https://doi.org/10.1016/j.stemcr.2015.03.001.

277. Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho Y-J, et al. Clonal dynamics of native haematopoiesis. Nature. 2014;514:322–7. https://doi.org/10.1038/nature13824.

278. Henssen AG, Henaff E, Jiang E, Eisenberg AR, Carson JR, Villasante CM, et al. Genomic DNA transposition induced by human PGBD5; 2015.

279. Ivics Z. Endogenous Transposase Source in Human Cells Mobilizes piggyBac Transposons. Mol Ther. 2016;24:851–4. https://doi.org/10.1038/mt.2016.76.

280. Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet. 2011;12:301–15. https://doi.org/10.1038/nrg2985.

281. Gogol-Döring A, Ammar I, Gupta S, Bunse M, Miskey C, Chen W, et al. Genome-wide Profiling Reveals Remarkable Parallels Between Insertion Site Selection Properties of the MLV Retrovirus and the piggyBac Transposon in Primary Human CD4(+) T Cells. Mol Ther. 2016;24:592–606. https://doi.org/10.1038/mt.2016.11.

282. Ammar I, Gogol-Döring A, Miskey C, Chen W, Cathomen T, Izsvák Z, Ivics Z. Retargeting transposon insertions by the adeno-associated virus Rep protein. Nucleic Acids Res. 2012;40:6693–712. https://doi.org/10.1093/nar/gks317.

283. Li MA, Pettitt SJ, Eckert S, Ning Z, Rice S, Cadiñanos J, et al. The piggyBac transposon displays local and distant reintegration preferences and can cause mutations at noncanonical integration sites. Mol Cell Biol. 2013;33:1317–30. https://doi.org/10.1128/MCB.00670-12.

284. Huang X, Guo H, Tammana S, Jung Y-C, Mellgren E, Bassi P, et al. Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. Mol Ther. 2010;18:1803–13. https://doi.org/10.1038/mt.2010.141.

285. Wilson MH, Coates CJ, George AL. PiggyBac transposon-mediated gene transfer in human cells. Mol Ther. 2007;15:139–45. https://doi.org/10.1038/sj.mt.6300028.

286. Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. Genesis. 2009;47:404–8. https://doi.org/10.1002/dvg.20508.

287. Wang H, Mayhew D, Chen X, Johnston M, Mitra RD. "Calling cards" for DNA-binding proteins in mammalian cells. Genetics. 2012;190:941–9. https://doi.org/10.1534/genetics.111.137315.

288. de JJ, Akhtar W, Badhai J, Rust AG, Rad R, Hilkens J, et al. Chromatin landscapes of retroviral and transposon integration profiles. PLoS Genet. 2014;10:e1004250. https://doi.org/10.1371/journal.pgen.1004250.