

Transposable elements and human cancer: a causal relationship?

Benoît Chénais*

Université du Maine, EA2160 Mer Molécules Santé, UFR Sciences et Techniques, Avenue
Olivier Messiaen, 72085 Le Mans, France.

***Corresponding author:** Université du Maine, EA2160 Mer Molécules Santé, UFR Sciences
et Techniques, Avenue Olivier Messiaen, 72085 Le Mans, France. Tel.: +33 243 833 251;
fax: 33 243 833 917. E-mail: bchenais@univ-lemans.fr.

Abstract

Transposable elements are present in almost all genomes including that of humans. These mobile DNA sequences are capable of invading genomes and their impact on genome evolution is substantial as they contribute to the genetic diversity of organisms. The mobility of transposable elements can cause deleterious mutations, gene disruption and chromosome rearrangements that may lead to several pathologies including cancer. This mini-review aims to give a brief overview of the relationship that transposons and retrotransposons may have in the genetic cause of human cancer onset, or conversely creating protection against cancer. Finally, the cause of TE mobility may also be the cancer cell environment itself.

Keywords: cancer; chromatin methylation; chromosomal rearrangement; gene expression; retrotransposon; transposon.

Abbreviations: HERVs, human endogenous retroviruses; LINE, long interspersed nuclear element; LTR, Long terminal repeats; MITEs, miniature inverted-repeat transposable elements; NHEJ, non-homologous end joining; SINE, short interspersed nuclear element; SVA, SINE-VNTR-*Alu* element; TE, transposable element; VNTR, variable number of tandem repeats.

1. Introduction: what are transposable elements?

Transposable elements (TEs) are repeated and mobile DNA sequences, with the ability to move in and invade genomes. They are present in most genomes from bacteria to eukaryotes and generally represent a substantial fraction of the genome depending on the species [1-5]. Two classes of TE are roughly distinguished (Fig. 1). Class I elements, or retrotransposons, use reverse transcriptase to copy an RNA genome into the host DNA; they are divided into Long Terminal Repeat (LTR) and non-LTR elements [6, 7]. Classed as LTRs, the human endogenous retroviruses (HERVs) resemble retroviruses in both their structure and mobility mechanism. Most HERVs contain a nonfunctional envelope gene, which relegates them to an intracellular existence [6]. Class-I elements present in the human genome are mainly non-LTR long interspersed elements (LINEs), i.e. LINE-1 (L1), LINE-2 and LINE-3, and LTR-element HERVs (HERV-I, HERV-K, HERV-L) [8]. With class II elements, or DNA transposons, the DNA genome of the element itself serves as the template for transposition either by a "cut and paste" mechanism involving the excision and reinsertion of the DNA sequence of the element, or by using a rolling circle process or a virus-like process [7, 9]. DNA transposons present in the human genome belong to TC-1/mariner superfamily (i.e. mariner, MER2-Tigger, Tc2), hAT superfamily (i.e. MER-1-charlie, Zaphod), and some PiggyBac like elements [8]. The two classes have been subdivided into super-families and then into families on the basis of the transposition mechanism, sequence similarities and structural relationships [10].

A growing number of TEs are non-autonomous and appear dependent on another helper element. They are unable to remove and insert themselves elsewhere in the genome and need the enzyme(s) encoded by autonomous elements to move. Nevertheless, non-autonomous elements such as miniature inverted-repeat transposable elements (MITEs) and short interspersed nuclear elements (SINEs, Fig. 1B) have greatly invaded genomes [4],

notably MITEs in rice [11] and SINEs in mammals [3]. The human genome contains three main SINEs families, i.e. *Alu* repeats, MIR and MIR3 [8].

Transposable element activity and its impact on eukaryotic genome evolution is substantial and TEs are currently accepted as an evolutionary force [2, 4, 5]. Transposition activity may contribute to the genetic diversity of organisms acting as gene regulatory elements by providing their own promoters or by altering the state of chromatin [7, 12-14]. Another way for a TE to be a source of genetic innovation is its recruitment on behalf of the genome to become a new gene, a phenomenon known as TE domestication [7, 13, 14]. But the TE mobility can cause mutations, some of which can result in several diseases, including cancer [15-19]. TEs can insert themselves into genes or gene regulatory elements, potentially disrupting the gene and/or gene function. TEs can also trigger chromosome rearrangements through ectopic recombination between multiple copies of a TE [6, 12, 16, 20].

2. Non-LTR retroelement mobility: a genetic cause of cancer.

Most non-LTR retroelement families in the human genome are currently inactive, except for three. L1 elements make up nearly 17% of human genomic DNA (Figure 2), with a total of about 500,000 copies [8]. Full-length L1 elements, stretching for some 6 kb, have two open reading frames encoding proteins required for the transposition and relocation of non-autonomous elements of the SINE family. L1 elements have been active in the human genome for near 160 million years [8]. SINEs are short elements (100–400 bp) that contain a promoter for polymerase III and do not encode proteins (Figure 1B). Most known SINE elements are tRNA derivatives, except the *Alu* element that is present in over a million copies in the human genome, i.e. about 11% (Fig. 2) [8]. The *Alu* element is specific to primates and has colonized primate genomes for 65 million years [21]. Another group of non-autonomous retroelements active in humans is the SVA (for SINE-VNTR-*Alu*; with VNTR for variable number of

tandem repeats) group of TEs (Fig. 1B) [22]. Like SINEs, SVAs use the machinery of L1 elements for transposition and are specific to hominids [23]. The colonization of the human genome by SVAs is relatively recent, *i.e.* less than 25 million years, and SVAs account for nearly 3,000 copies [16].

A link between the transposition of mobile retroelements (mainly LINE or SINE/*Alu*) in the human genome and pathologies was noted many years ago. Several examples of LINE or SINE/*Alu* related cancers are listed in Table 1. For instance, in the late 1980s L1 retroelement insertion into the human proto-oncogene *c-myc* was found in human breast carcinoma cells (Table 1) [24]. In the study by Morse et al., the structure comparison of *myc* from a breast ductal adenocarcinoma and from normal breast tissue of the same patient revealed a tumor-specific rearrangement of one *myc* locus and amplification of the other *myc* locus. Within the second intron of the rearranged locus was a non-*myc* sequence with near complete homology to a LINE-1 sequence. In this case, the L1 sequence functioned as a mobile genetic element to produce a somatic mutation [24]. Another example of an L1 element somatic insertion is its integration into the tumor-suppressing gene *Apc* (adenomatous polyposis coli), found in colon cancer patients (Table 1) [25-27]. The insertion of an *Alu* element in *Apc* was also observed in association with desmoid tumors [28].

Insertions of the non-autonomous SINE *Alu* element into the intron of the NF-1 (neurofibromatosis type I) gene lead to a deletion and a reading frame shift in the downstream exon during splicing, which might be associated with neurofibromatosis[29]. The BRCA1 and BRCA2 breast/ovarian cancer associated genes are also the sites of *Alu* insertions [30, 31]. So, one of the germline mutations evidenced by Miki et al., was an insertion of an *Alu* element into exon 22, which resulted in alternative splicing that skipped exon 22. This event was caused by the retrotransposal insertion of a transcriptionally active *Alu* element [30].

3. *Alu*-mediated chromosomal recombination at the origin of cancer.

As noted above *Alu* repeats account for 11% of the human genome (Fig. 2). These repeats are thought to induce ectopic recombination (also known as non-allelic homologous recombination) at the origin of several cancer diseases. The unusually high density of *Alu* repeats in the BRCA1 gene favor ectopic recombination and numerous *Alu*-mediated deletion events have been reported at this locus [32, 33]. At least four cases of BRCA1 gene duplication were reported involving the ectopic recombination of *Alu* elements [33, 34]. Although less frequent, *Alu*-mediated deletions were also described for BRCA2 gene [33, 35].

In addition, *Alu* elements may play a role in chronic myeloid leukemia. This leukemia develops as a result of a translocation between the human chromosomes 9 and 22. The chromosome breakpoints producing this chromosome aberration contained nucleotide sequences of *Alu* elements [36]. Similarly, an internal tandem duplication of part of the *Mll* (myeloid/lymphoid or mixed-lineage leukemia) gene, resulting from ectopic recombination between those very *Alu* elements, may trigger a cascade of events frequently associated with acute myeloid leukemia [37, 38]. Duplication of the *MYB* locus, which encodes an essential transcription factor, is frequent in human cancers [39, 40]. The human *MYB* locus is flanked by *Alu* repeats and O'Neil et al., showed that its duplication is mediated somatically by homologous recombination between the flanking *Alu* elements on sister chromatids in T cell acute lymphoblastic leukemia [41].

Finally, recombination between *Alu* elements, which causes a translocation involving the Tre-2 oncogene (TRE-USP6, ubiquitin-specific protease-6), has been shown to play an important role in Ewing sarcoma development [42]. Several other examples exist such as *Alu*-mediated deletions in the von Hippel Lindau gene [43, 44], the cadherin-1 gene [45], or the caspase-activated DNase gene [46] (Table 1).

4. Transposable element domestication: an evolutionary gain sometimes related to cancer or cancer suppression.

Transposable element domestication corresponds to the diversion of these sequences and the reuse of original functions of recombinases (transposases, integrases) for another role in the cell. At least 47 human genes are known to originate from TEs [8]. In mammals, the *Syncitin* genes that play crucial roles in placenta formation are derived from HERV LTR-retroelements [47, 48]. In *D. melanogaster*, the non-LTR retroelements *Heta* and *Tart* assume the function of telomeres and telomerase [49]. TEs are also involved in the formation and epigenetic regulation of heterochromatin as described in *Drosophila* [50, 51] and several other vertebrate and invertebrate species [52-55].

The most famous example of TE domestication in mammals is the derivation of an ancient *transib*-type element that causes the V(D)J system at the base of the vertebrate immune system [56-59]. The RAG-1 and RAG-2 genes that promote recombination, retain their ability to relocate their nucleotide sequences during V(D)J recombination in the lymphocyte genome [60]. Furthermore, V(D)J recombination events are biochemically similar to the transposition of the *Hermes* transposon family, such as *hobo*, *Activator*, and *Tam3*, which relocate via the “cut-and-paste” mechanism [59]. In fact, a nucleotide sequence fragment cut out during V(D)J recombination resembled transposon DNA, except that the former is circularized [58, 59]. Fragments cut out by RAG proteins are usually degraded. However, sometimes the proteins can reinsert these fragments into other genome sites with a frequency of about 1 per 13,000-50,000 recombination events in cultured human cells [61-63]. Although this rate of illegitimate insertion cannot be directly extrapolated from cell culture to an organism, a link between these events and B- and T-cell cancerisation in the human organism can be presumed [64]. Furthermore, a link between V(D)J recombination and the onset of cancer associated with recombination-induced chromosome rearrangements

has been shown [65, 66]. RAG proteins may induce double-strand breaks in sites similar in their structure to signal sequences for V(D)J recombination. Such double-strand breaks in DNA are potential players in recombination of the genes or receptors of mature T and B cells. This, in turn, entails expression deviations of such proto-oncogenes as LMO2 and BCL2 [65, 66].

As an example of Class-II transposon domestication in the human genome, the fusion of the mariner transposon *hsmar1* with a methylase coding sequence (SET) is of particular interest with respect to cancer [67, 68]. The resulting chimeric protein, referred to as SETMAR, or Metnase, shows both methylase and nuclease activities [68-70] and promotes non-homologous end joining (NHEJ) repair in humans [71-74]. The fusion of the histone methylase SET domain and the transposase domain in the anthropoid lineage to form primate Metnase promotes accurate intrachromosomal NHEJ and thereby suppresses interchromosomal translocations. Metnase may have been selected because it has the function of opposing transposases and thus may play a key role in suppressing translocations that underlie oncogenicity [75-77]. Metnase transposase has been remarkably conserved through evolution; however, there is a clustering of substitutions located in alpha helices 4 and 5 within the putative DNA-binding site, consistent with loss of transposition-specific DNA cleavage activity and acquisition of repair-specific DNA cleavage activity [78]. Both methylase (SET) and transposase domains are necessary for Metnase function in double-strand break repair, and Metnase possesses a unique endonuclease activity that preferentially acts on ssDNA and ssDNA-overhang of a partial duplex DNA [71]. This result of Beck et al. suggests a role for Metnase's endonuclease activity in promoting the joining of non-compatible ends [71]. Metnase is widely expressed, especially in proliferating tissues, and its involvement in cancer cell resistance to Topoisomerase II- α inhibitors has been suggested. Indeed acute leukemia cells as well as breast cancer cells have an attenuated mitotic arrest due

to decatenation inhibition by Metnase [76, 77]. Therefore Metnase inhibition will restore chemotherapy sensitivity in these cells.

5. But what may cause or promote transposable element mobility?

Activation of TE mobility may occur as a result of different mechanisms. For example, L1 elements are known to actively transpose during early embryogenesis, which is believed to be triggered by total genome demethylation, or epigenetic reprogramming as shown in murine primordial cells between the E11.5 and E13.5 early embryo stages [79]. As DNA methylation is known to repress various nucleotide sequences, including L1 elements, demethylation may cause TE activation and transposition events [80, 81]. Kano et al. have demonstrated that mRNA from L1 elements transcribed in the parental organism can be passed on through oocytes or sperm cells to progeny where reverse transcription ensures further insertions of the element's copies into the genome of the developing organism during the pre-implantation stage, leading to somatic mosaicism [82]. It seems, therefore, that at least L1 activation paths exist during early stages of mammal development [83]. TE activation is rather common during embryogenesis, so retroelement insertions, including those associated with cancer development may be considered as the cost of phenotypic diversity formation.

Environmental factors may also cause TE mobility. Several chemicals containing mercury (HgS), cadmium (CdS), and nickel (NiO) have been found to increase 3-fold L1 element activity in human cell culture [84]. Meanwhile nickel chloride, which increases L1 activity 2.5-fold, has no direct effect on the sequence of the element or its proteins, but instead inhibits DNA repair systems, leading to L1 transpositions [85]. In general, active TE transposition in living organisms can be induced by a number of environmental factors, like heat shock, viral infection, various chemicals, γ irradiation, etc [86-90]. Although not clearly established, environmentally-induced TE activation could, albeit indirectly, contribute to

human carcinogenesis. This external factor effect on TE-mediated carcinogenesis is further supported by the geographic occurrence patterns of such recombination events. For example, some studies link the expression rates of breast cancer associated gene BRCA2, specific to the Portuguese population, to *Alu* activity [31, 91].

The tumor cell environment itself is favorable to TE mobility. In fact the frequent hypomethylation of chromatin in tumor cells will facilitate TE activation [80, 81, 92, 93]. L1 hypomethylation has been observed in several cancers such as breast carcinoma [94], prostate cancer [95-97], hepatocellular carcinomas [98] or colorectal cancer [99]. Moreover recent findings using second-generation DNA sequencing revealed a total of 9 *de novo* L1 retrotransposition events in 6 out of 20 examined non-small cell lung tumors [100]. In this study the new L1 insertions were accompanied by a specific genome-wide hypomethylation signature. This is consistent with the idea that epigenome alterations create a permissive environment for TE expression and/or retrotransposition [7, 100].

Recently Lee *et al.*, identified almost two hundred somatic TE insertions in 43 high-coverage whole genome sequencing datasets from five cancer types (*i.e.* colorectal, glioblastoma, multiple myeloma, ovarian and prostate) [101]. Most of the identified TE insertions are L1s (94.5%), then *Alu* (5.0%) and ERV (0.5%). All the somatic L1 and *Alu* insertions were observed in epithelial cancer whereas no L1 and *Alu* insertions were detected in blood and brain cancers. Interestingly, the somatic L1 insertions are biased away from transcriptionally active regions and conversely overrepresented in common cancer-specific hypomethylation domains [101]. This supports the view that loss of DNA methylation promotes TE integration. Lee *et al.* also showed that somatic L1 insertions tend to disrupt the expression of tumor suppressor genes commonly mutated in cancer, suggesting that some TE insertions provide a selective advantage during tumorigenesis [101].

Furthermore, tumor cells are known to contain significantly lower quantities of micro RNAs [102], so the repression of TEs through RNA interference, the mechanisms of which are detailed elsewhere [103, 104], may be bypassed in cancerous cells. Therefore, while TEs may be linked to cancer development, they themselves can become activated by the cell malignization process that promotes increased mutation and recombination rates in the genome of the transformed cells.

5. Concluding remarks

The list of cancer-linked chromosome rearrangements extends far beyond those caused by the function defects of V(D)J recombination genes. Indeed oncogenic chromosome rearrangements can be formed at fragile chromosome sites due to imperfect functioning of the NHEJ and homologous recombination repair systems. The breakpoints of such recombination events are often localized in or near the nucleotide sequences of *Alu* elements [105]. However, the presence of an *Alu* sequence itself is not sufficient. It is more the nature of this sequence that induces double-strand breaks and provides an initiation point for recombination [106, 107]. Although rare in somatic cells, the existence of ectopic recombination between *Alu* sequences leads to DNA deletions in germ cells (Table 1) [107]. As noted above, such events can cause tumors such as acute myeloid leukemia and Ewing sarcoma (Table 1) [37, 42].

Besides inducing chromosome rearrangements, TEs may also affect gene expression [5, 108]. Gene disruption is perhaps the most evident effect of TE insertion, and the inclusion of a TE often results in genetic disorders caused by deleterious premature termination of the peptide sequence. However, gene expression may be affected by TE insertion in different ways, such as the introduction of new polyadenylation sites [109, 110], the creation of new exons [111-113], or the insertion of a *cis*-element at promoter and enhancer regions [114-

116]. Such dysregulation of gene expression may be the cause of a number of human diseases including cancers [15-18].

Studies in animal models also support a causative role of TEs in cancer onset. The observation of TEs in mouse cancer cells [117-119] and the use of engineered TEs, such as *Sleeping Beauty* and *PiggyBac* as mutagen agents, demonstrate that TE insertion can lead to cancer [120-123]. In the first studies by Dupuy et al.[124] and Collier et al.[125], the reconstructed Tc-1 element *Sleeping Beauty* transposase was expressed in mice that also carried concatemers of gene trap transposons. In the work of Dupuy et al., crosses of offspring frequently died *in utero*, but mice which survived until birth succumbed to cancer rapidly [124]. In the second study by Collier *et al.*, the mice did not develop tumors spontaneously but mobilization of the transposase appeared to accelerate the onset of sarcomas in *p19 ARF* null mice [125]. In further studies a conditional *Sleeping Beauty* transposition system was used in order to target specific tissues such as gastrointestinal tract and liver [120]. Such a TE-based insertional mutagenesis provides great help in identifying cancer candidate genes in mice [120, 121] and fish [126].

Several transcription factors, such as Foxa1, GATA, LyF-1, Sp1, YY1; retinoic acid receptors (RARs) and p53 have binding sites in *Alu* elements [127-133]. In particular, putative p53 binding sites have been discovered in several LTR- and non-LTR elements of the human genome [127]. The p53 tumor suppressor gene is a transcription factor that induces activation and repression of more than a thousand human genes and is involved in about 50% of human cancers [134, 135]. Previous studies have shown that p53 can repress RNA polymerase III transcription [136]. Therefore, mutations in p53 leading to DNA binding defects would result in RNA polymerase III hyperactivity that characterizes many tumors [136]. Moreover, an increased level of *Alu* RNA has been evidenced in hepatocellular carcinoma [137]. Cui *et al.*, hypothesized that p53 binding to *Alu* elements, plus the transcriptional repression can change

the expression of the host genes through steric interaction between RNA polymerase II, which is responsible for host gene expression, and RNA polymerase III that ensures *Alu* transcription [127]. This hypothesis has to be demonstrated as well as its consequences with respect to cancer onset and/or its development.

Finally, similarities between the reverse transcriptase of human L1 elements and telomerase were described [138, 139]. The telomerase function and the telomere sequence are provided by domesticated TEs in several organisms, such as *Drosophila*, rotifers, the diatom microalga *Phaeodactylum tricornutum*, the silkworm *Bombyx mori*, etc [49, 140-142]. Due to the crucial involvement of telomerase in cancer development and the complex regulation of telomere lengthening [143, 144], this suggests that retrotransposons, by providing an alternative mechanism to maintain telomere structure, may play a yet unelucidated role in cancer biology.

Although a direct relationship between TE insertion and cancer onset is difficult to evidence, the above examples illustrate that TE-linked rearrangements in the human genome may be associated with cancers of various etiology. To demonstrate that TE insertion or TE-mediated rearrangement causes human cancer is difficult because i) sequencing repetitive elements remain a challenge and ii) experimental demonstration is impossible in humans. However, recent animal model studies with refined bioinformatic analyses of insertional mutagenesis screens [145-148], and improved human genome analysis such as the recent work of Lee *et al.* [101], offer solid arguments in favor of the causative role of TEs in cancer. Conversely, the SETMAR example illustrates the case of a positive antitumor effect due to TE domestication.

According to the available data, TEs are clearly a source of mutations and are therefore able to lead to cancer development. Cancer is not a single mutation event, but a result of the accumulation of several mutations affecting the signaling of cell cycle, apoptosis,

angiogenesis and DNA repair. Thus TEs can produce mutations causing cancer susceptibility, but are not the sole cause of cancer.

Acknowledgements: The author is grateful to Marilyn Twell for revising the English text and to Dr. Albane de Gentile for her continuous support during this project.

References

- [1] J.L. Bennetzen, Transposable elements, gene creation and genome rearrangement in flowering plants, *Current opinion in genetics & development*, 15 (2005) 621-627.
- [2] C. Biemont, C. Vieira, Genetics: junk DNA as an evolutionary force, *Nature*, 443 (2006) 521-524.
- [3] A.P. de Koning, W. Gu, T.A. Castoe, M.A. Batzer, D.D. Pollock, Repetitive elements may comprise over two-thirds of the human genome, *PLoS Genet.*, 7 (2011) e1002384.
- [4] C. Feschotte, E.J. Pritham, DNA transposons and the evolution of eukaryotic genomes, *Annu. Rev. Genet.*, 41 (2007) 331-368.
- [5] B. Chénais, A. Caruso, S. Hiard, N. Casse, The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments, *Gene*, in press (2012).
- [6] C.R. Beck, J.L. Garcia-Perez, R.M. Badge, J.V. Moran, LINE-1 elements in structural variation and disease, *Annu. Rev. Genomics Hum. Genet.*, 12 (2011) 187-215.
- [7] H.L. Levin, J.V. Moran, Dynamic interactions between transposable elements and their hosts, *Nat. Rev. Genet.*, 12 (2011) 615-627.
- [8] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczký, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H.

Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A.

Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglu, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A.

Wetterstrand, A. Patrinos, M.J. Morgan, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y.J. Chen, Initial sequencing and analysis of the human genome, *Nature*, 409 (2001) 860-921.

[9] E.J. Pritham, Transposable elements and factors influencing their success in eukaryotes, *J. Hered.*, 100 (2009) 648-655.

[10] T. Wicker, F. Sabot, A. Hua-Van, J.L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A.H. Schulman, A unified classification system for eukaryotic transposable elements, *Nat. Rev. Genet.*, 8 (2007) 973-982.

[11] C. Lu, J. Chen, Y. Zhang, Q. Hu, W. Su, H. Kuang, Miniature Inverted-repeat Transposable Elements (MITEs) Have Been Accumulated through Amplification Bursts and Play Important Roles on Gene Expression and Species Diversity in *Oryza sativa*, *Mol. Biol. Evol.*, (2011).

[12] G. Bourque, Transposable elements in gene regulation and in the evolution of vertebrate genomes, *Curr. Opin. Genet. Dev.*, 19 (2009) 607-612.

[13] C. Feschotte, Transposable elements and the evolution of regulatory networks, *Nat. Rev. Genet.*, 9 (2008) 397-405.

[14] L. Sinzelle, Z. Izsvak, Z. Ivics, Molecular domestication of transposable elements: from detrimental parasites to useful host genes, *Cell. Mol. Life Sci.*, 66 (2009) 1073-1093.

[15] V.P. Belancio, A.M. Roy-Engel, P.L. Deininger, All y'all need to know 'bout retroelements in cancer, *Semin. Cancer Biol.*, 20 (2010) 200-210.

[16] M.K. Konkel, M.A. Batzer, A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome, *Semin. Cancer Biol.*, 20 (2010) 211-221.

[17] I.A. Kozeretska, S.V. Demydov, L.I. Ostapchenko, Mobile genetic elements and cancer. From mutations to gene therapy, *Exp. Oncol.*, 33 (2011) 198-205.

- [18] E. Lerat, M. Semon, Influence of the transposable element neighborhood on human gene expression in normal and tumor tissues, *Gene*, 396 (2007) 303-311.
- [19] K.H. Burns, J.D. Boeke, Human transposon tectonics, *Cell*, 149 (2012) 740-752.
- [20] H. Zhao, G. Bourque, Recovering genome rearrangements in the mammalian phylogeny, *Genome Res.*, 19 (2009) 934-942.
- [21] M.A. Batzer, P.L. Deininger, Alu repeats and human genomic diversity, *Nat. Rev. Genet.*, 3 (2002) 370-379.
- [22] E.M. Ostertag, J.L. Goodier, Y. Zhang, H.H. Kazazian, Jr., SVA elements are nonautonomous retrotransposons that cause disease in humans, *Am. J. Hum. Genet.*, 73 (2003) 1444-1451.
- [23] H. Wang, J. Xing, D. Grover, D.J. Hedges, K. Han, J.A. Walker, M.A. Batzer, SVA elements: a hominid-specific retroposon family, *J. Mol. Biol.*, 354 (2005) 994-1007.
- [24] B. Morse, P.G. Rotherg, V.J. South, J.M. Spandorfer, S.M. Astrin, Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma, *Nature*, 333 (1988) 87-90.
- [25] Y. Miyoshi, H. Nagase, H. Ando, A. Horii, S. Ichii, S. Nakatsuru, T. Aoki, Y. Miki, T. Mori, Y. Nakamura, Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene, *Hum. Mol. Genet.*, 1 (1992) 229-233.
- [26] Y. Miyoshi, H. Ando, H. Nagase, I. Nishisho, A. Horii, Y. Miki, T. Mori, J. Utsunomiya, S. Baba, G. Petersen, S.R. Hamilton, K.W. Kinzler, B. Vogelstein, Y. Nakamura, Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients, *Proc. Natl. Acad. Sci. U. S. A.*, 89 (1992) 4452-4456.
- [27] Y. Miki, I. Nishisho, A. Horii, Y. Miyoshi, J. Utsunomiya, K.W. Kinzler, B. Vogelstein, Y. Nakamura, Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer, *Cancer Res.*, 52 (1992) 643-645.

- [28] K.C. Halling, C.R. Lazzaro, R. Honchel, J.A. Bufill, S.M. Powell, C.A. Arndt, N.M. Lindor, Hereditary desmoid disease in a family with a germline Alu I repeat mutation of the APC gene, *Hum. Hered.*, 49 (1999) 97-102.
- [29] M.R. Wallace, L.B. Andersen, A.M. Saulino, P.E. Gregory, T.W. Glover, F.S. Collins, A de novo Alu insertion results in neurofibromatosis type 1, *Nature*, 353 (1991) 864-866.
- [30] Y. Miki, T. Katagiri, F. Kasumi, T. Yoshimoto, Y. Nakamura, Mutation analysis in the BRCA2 gene in primary breast cancers, *Nat. Genet.*, 13 (1996) 245-247.
- [31] E. Teugels, S. De Brakeleer, G. Goelen, W. Lissens, E. Sermijn, J. De Greve, De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes, *Hum. Mutat.*, 26 (2005) 284.
- [32] S. Armaou, I. Konstantopoulou, T. Anagnostopoulos, E. Razis, I. Boukovinas, N. Xenidis, G. Fountzilas, D. Yannoukakos, Novel genomic rearrangements in the BRCA1 gene detected in Greek breast/ovarian cancer patients, *Eur. J. Cancer*, 43 (2007) 443-453.
- [33] S. Mazoyer, Genomic rearrangements in the BRCA1 and BRCA2 genes, *Hum. Mutat.*, 25 (2005) 415-422.
- [34] K.P. Yap, P. Ang, I.H. Lim, G.H. Ho, A.S. Lee, Detection of a novel Alu-mediated BRCA1 exon 13 duplication in Chinese breast cancer patients and implications for genetic testing, *Clin. Genet.*, 70 (2006) 80-82.
- [35] I. Tournier, B.B. Paillerets, H. Sobol, D. Stoppa-Lyonnet, R. Lidereau, M. Barrois, S. Mazoyer, F. Coulet, A. Hardouin, A. Chompret, A. Lortholary, P. Chappuis, V. Bourdon, V. Bonadona, C. Maugard, B. Gilbert, C. Nogues, T. Frebourg, M. Tosi, Significant contribution of germline BRCA2 rearrangements in male breast cancer families, *Cancer Res.*, 64 (2004) 8143-8147.

- [36] A.R. Jeffs, S.M. Benjes, T.L. Smith, S.J. Sowerby, C.M. Morris, The BCR gene recombines preferentially with Alu elements in complex BCR-ABL translocations of chronic myeloid leukaemia, *Hum. Mol. Genet.*, 7 (1998) 767-776.
- [37] S.A. Schichman, M.A. Caligiuri, M.P. Strout, S.L. Carter, Y. Gu, E. Canaani, C.D. Bloomfield, C.M. Croce, ALL-1 tandem duplication in acute myeloid leukemia with a normal karyotype involves homologous recombination between Alu elements, *Cancer Res.*, 54 (1994) 4277-4280.
- [38] M.P. Strout, G. Marcucci, C.D. Bloomfield, M.A. Caligiuri, The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia, *Proc. Natl. Acad. Sci. U. S. A.*, 95 (1998) 2390-2395.
- [39] R.G. Ramsay, T.J. Gonda, MYB function in normal and cancer cells, *Nat. Rev. Cancer*, 8 (2008) 523-534.
- [40] A. Sala, B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer, *Eur. J. Cancer*, 41 (2005) 2479-2484.
- [41] J. O'Neil, J. Tchinda, A. Gutierrez, L. Moreau, R.S. Maser, K.K. Wong, W. Li, K. McKenna, X.S. Liu, B. Feng, D. Neuberg, L. Silverman, D.J. DeAngelo, J.L. Kutok, R. Rothstein, R.A. DePinho, L. Chin, C. Lee, A.T. Look, Alu elements mediate MYB gene tandem duplication in human T-ALL, *J. Exp. Med.*, 204 (2007) 3059-3066.
- [42] M. Onno, T. Nakamura, J. Hillova, M. Hill, Rearrangement of the human *trc* oncogene by homologous recombination between Alu repeats of nucleotide sequences from two different chromosomes, *Oncogene*, 7 (1992) 2519-2523.
- [43] A. Casarin, M. Martella, R. Polli, E. Leonardi, L. Anesi, A. Murgia, Molecular characterization of large deletions in the von Hippel-Lindau (VHL) gene by quantitative real-

time PCR: the hypothesis of an alu-mediated mechanism underlying VHL gene rearrangements, *Mol. Diagn. Ther.*, 10 (2006) 243-249.

[44] G. Franke, B. Bausch, M.M. Hoffmann, M. Cybulla, C. Wilhelm, J. Kohlhase, G. Scherer, H.P. Neumann, Alu-Alu recombination underlies the vast majority of large VHL germline deletions: Molecular characterization and genotype-phenotype correlations in VHL patients, *Hum. Mutat.*, 30 (2009) 776-786.

[45] C. Oliveira, J. Senz, P. Kaurah, H. Pinheiro, R. Sanges, A. Haegert, G. Corso, J. Schouten, R. Fitzgerald, H. Vogelsang, G. Keller, S. Dwerryhouse, D. Grimmer, S.F. Chin, H.K. Yang, C.E. Jackson, R. Seruca, F. Roviello, E. Stupka, C. Caldas, D. Huntsman, Germline CDH1 deletions in hereditary diffuse gastric cancer families, *Hum. Mol. Genet.*, 18 (2009) 1545-1555.

[46] S.Y. Hsieh, W.Y. Chen, T.S. Yeh, I.S. Sheen, S.F. Huang, High-frequency Alu-mediated genomic recombination/deletion within the caspase-activated DNase gene in human hepatoma, *Oncogene*, 24 (2005) 6584-6589.

[47] S.G. Black, F. Arnaud, M. Palmarini, T.E. Spencer, Endogenous retroviruses in trophoblast differentiation and placental development, *Am. J. Reprod. Immunol.*, 64 (2010) 255-264.

[48] S. Prudhomme, B. Bonnaud, F. Mallet, Endogenous retroviruses and animal reproduction, *Cytogenet. Genome Res.*, 110 (2005) 353-364.

[49] M.L. Pardue, P. Debaryshe, Adapting to life at the end of the line: How *Drosophila* telomeric retrotransposons cope with their job, *Mob. Genet. Elements*, 1 (2011) 128-134.

[50] Z. Lippman, A.V. Gendrel, M. Black, M.W. Vaughn, N. Dedhia, W.R. McCombie, K. Lavine, V. Mittal, B. May, K.D. Kasschau, J.C. Carrington, R.W. Doerge, V. Colot, R. Martienssen, Role of transposable elements in heterochromatin and epigenetic control, *Nature*, 430 (2004) 471-476.

- [51] A.C. Spradling, H.J. Bellen, R.A. Hoskins, *Drosophila* P elements preferentially transpose to replication origins, *Proc. Natl. Acad. Sci. U. S. A.*, 108 (2011) 15948-15953.
- [52] D.C. Ferreira, C. Oliveira, F. Foresti, A new dispersed element in the genome of the catfish *Hisonotus leucofrenatus* (Teleostei: Siluriformes: Hypoptopomatinae), *Mob. Genet. Elements*, 1 (2011) 103-106.
- [53] O.G. Grushko, M.V. Sharakhova, V.N. Stegnii, I.V. Sharakhov, Molecular organization of heterochromatin in malaria mosquitoes of the *Anopheles maculipennis* subgroup, *Gene*, 448 (2009) 192-197.
- [54] T. Palomeque, J. Antonio Carrillo, M. Munoz-Lopez, P. Lorite, Detection of a mariner-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution, *Gene*, 371 (2006) 194-205.
- [55] R. Rebollo, M.M. Karimi, M. Bilenky, L. Gagnier, K. Miceli-Royer, Y. Zhang, P. Goyal, T.M. Keane, S. Jones, M. Hirst, M.C. Lorincz, D.L. Mager, Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms, *PLoS Genet.*, 7 (2011) e1002301.
- [56] V.V. Kapitonov, J. Jurka, RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons, *PLoS Biol.*, 3 (2005) e181.
- [57] D.C. van Gent, K. Mizuuchi, M. Gellert, Similarities between initiation of V(D)J recombination and retroviral integration, *Science*, 271 (1996) 1592-1594.
- [58] L. Zhou, R. Mitra, P.W. Atkinson, A.B. Hickman, F. Dyda, N.L. Craig, Transposition of hAT elements links transposable elements and V(D)J recombination, *Nature*, 432 (2004) 995-1001.
- [59] J.M. Jones, M. Gellert, The taming of a transposon: V(D)J recombination and the immune system, *Immunol. Rev.*, 200 (2004) 233-248.

- [60] Y.V. Reddy, E.J. Perkins, D.A. Ramsden, Genomic instability due to V(D)J recombination-associated transposition, *Genes Dev.*, 20 (2006) 1575-1582.
- [61] M. Chatterji, C.L. Tsai, D.G. Schatz, Mobilization of RAG-generated signal ends by transposition and insertion in vivo, *Mol. Cell. Biol.*, 26 (2006) 1558-1568.
- [62] T.L. Messier, J.P. O'Neill, S.M. Hou, J.A. Nicklas, B.A. Finette, In vivo transposition mediated by V(D)J recombinase in human T lymphocytes, *EMBO J.*, 22 (2003) 1381-1388.
- [63] A. Agrawal, Q.M. Eastman, D.G. Schatz, Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system, *Nature*, 394 (1998) 744-751.
- [64] K. Hiom, M. Melek, M. Gellert, DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations, *Cell*, 94 (1998) 463-470.
- [65] M.R. Lieber, K. Yu, S.C. Raghavan, Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations, *DNA Repair (Amst)*, 5 (2006) 1234-1245.
- [66] R. Marculescu, K. Vanura, B. Montpellier, S. Roulland, T. Le, J.M. Navarro, U. Jager, F. McBlane, B. Nadel, Recombinase, chromosomal translocations and lymphoid neoplasia: targeting mistakes and repair failures, *DNA Repair (Amst)*, 5 (2006) 1246-1258.
- [67] R. Cordaux, S. Udit, M.A. Batzer, C. Feschotte, Birth of a chimeric primate gene by capture of the transposase gene from a mobile element, *Proc. Natl. Acad. Sci. U. S. A.*, 103 (2006) 8101-8106.
- [68] S.H. Lee, M. Oshige, S.T. Durant, K.K. Rasila, E.A. Williamson, H. Ramsey, L. Kwan, J.A. Nickoloff, R. Hromas, The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair, *Proc. Natl. Acad. Sci. U. S. A.*, 102 (2005) 18075-18080.

- [69] D. Liu, J. Bischerour, A. Siddique, N. Buisine, Y. Bigot, R. Chalmers, The human SETMAR protein preserves most of the activities of the ancestral Hsmar1 transposase, *Mol. Cell. Biol.*, 27 (2007) 1125-1132.
- [70] Y. Roman, M. Oshige, Y.J. Lee, K. Goodwin, M.M. Georgiadis, R.A. Hromas, S.H. Lee, Biochemical characterization of a SET and transposase fusion protein, Metnase: its DNA binding and DNA cleavage activity, *Biochemistry*, 46 (2007) 11369-11376.
- [71] B.D. Beck, S.S. Lee, E. Williamson, R.A. Hromas, S.H. Lee, Biochemical characterization of metnase's endonuclease activity and its role in NHEJ repair, *Biochemistry*, 50 (2011) 4360-4370.
- [72] L.P. De Haro, J. Wray, E.A. Williamson, S.T. Durant, L. Corwin, A.C. Gentry, N. Osheroff, S.H. Lee, R. Hromas, J.A. Nickoloff, Metnase promotes restart and repair of stalled and collapsed replication forks, *Nucleic Acids Res.*, 38 (2010) 5681-5691.
- [73] R. Hromas, J. Wray, S.H. Lee, L. Martinez, J. Farrington, L.K. Corwin, H. Ramsey, J.A. Nickoloff, E.A. Williamson, The human set and transposase domain protein Metnase interacts with DNA Ligase IV and enhances the efficiency and accuracy of non-homologous end-joining, *DNA Repair (Amst)*, 7 (2008) 1927-1937.
- [74] M. Shaheen, E. Williamson, J. Nickoloff, S.H. Lee, R. Hromas, Metnase/SETMAR: a domesticated primate transposase that enhances DNA repair, replication, and decatenation, *Genetica*, 138 (2010) 559-566.
- [75] J. Wray, E.A. Williamson, S. Chester, J. Farrington, R. Sterk, D.M. Weinstock, M. Jasin, S.H. Lee, J.A. Nickoloff, R. Hromas, The transposase domain protein Metnase/SETMAR suppresses chromosomal translocations, *Cancer Genet. Cytogenet.*, 200 (2010) 184-190.
- [76] J. Wray, E.A. Williamson, M. Royce, M. Shaheen, B.D. Beck, S.H. Lee, J.A. Nickoloff, R. Hromas, Metnase mediates resistance to topoisomerase II inhibitors in breast cancer cells, *PLoS One*, 4 (2009) e5323.

- [77] J. Wray, E.A. Williamson, S. Sheema, S.H. Lee, E. Libby, C.L. Willman, J.A. Nickoloff, R. Hromas, Metnase mediates chromosome decatenation in acute leukemia cells, *Blood*, 114 (2009) 1852-1858.
- [78] K.D. Goodwin, H. He, T. Imasaki, S.H. Lee, M.M. Georgiadis, Crystal structure of the human Hsmar1-derived transposase domain in the DNA repair enzyme Metnase, *Biochemistry*, 49 (2010) 5705-5713.
- [79] S. Seisenberger, C. Popp, W. Reik, Retrotransposons and germ cells: reproduction, death, and diversity, *F1000 Biol. Rep.*, 2 (2010).
- [80] J.P. Ross, K.N. Rand, P.L. Molloy, Hypomethylation of repeated DNA sequences in cancer, *Epigenomics*, 2 (2010) 245-269.
- [81] Y. Watanabe, M. Maekawa, Methylation of DNA in cancer, *Adv. Clin. Chem.*, 52 (2010) 145-167.
- [82] H. Kano, I. Godoy, C. Courtney, M.R. Vetter, G.L. Gerton, E.M. Ostertag, H.H. Kazazian, Jr., L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism, *Genes Dev.*, 23 (2009) 1303-1312.
- [83] N. Lane, W. Dean, S. Erhardt, P. Hajkova, A. Surani, J. Walter, W. Reik, Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse, *Genesis*, 35 (2003) 88-93.
- [84] S.P. Kale, L. Moore, P.L. Deininger, A.M. Roy-Engel, Heavy metals stimulate human LINE-1 retrotransposition, *Int. J. Environ. Res. Public Health*, 2 (2005) 14-23.
- [85] M. El-Sawy, S.P. Kale, C. Dugan, T.Q. Nguyen, V. Belancio, H. Bruch, A.M. Roy-Engel, P.L. Deininger, Nickel stimulates L1 retrotransposition by a post-transcriptional mechanism, *J. Mol. Biol.*, 354 (2005) 246-257.

- [86] K. Fujino, S.N. Hashida, T. Ogawa, T. Natsume, T. Uchiyama, T. Mikami, Y. Kishima, Temperature controls nuclear import of Tam3 transposase in *Antirrhinum*, *Plant J.*, 65 (2011) 146-155.
- [87] M.A. Grandbastien, C. Audeon, E. Bonnivard, J.M. Casacuberta, B. Chalhoub, A.P. Costa, Q.H. Le, D. Melayah, M. Petit, C. Poncet, S.M. Tam, M.A. Van Sluys, C. Mhiri, Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae, *Cytogenet. Genome Res.*, 110 (2005) 229-241.
- [88] Y. Kimura, Y. Tosa, S. Shimada, R. Sogo, M. Kusaba, T. Sunaga, S. Betsuyaku, Y. Eto, H. Nakayashiki, S. Mayama, OARE-1, a Ty1-copia retrotransposon in oat activated by abiotic and biotic stresses, *Plant Cell Physiol.*, 42 (2001) 1345-1354.
- [89] W.J. Miller, P. Capy, Mobile genetic elements as natural tools for genome evolution, *Methods Mol. Biol.*, 260 (2004) 1-20.
- [90] A.M. Handler, S.P. Gomez, P element excision in *Drosophila* is stimulated by gamma-irradiation in transient embryonic assays, *Genet. Res.*, 70 (1997) 75-78.
- [91] M. El-Sawy, P. Deininger, Tandem insertions of Alu elements, *Cytogenet. Genome Res.*, 108 (2005) 58-62.
- [92] A.S. Wilson, B.E. Power, P.L. Molloy, DNA hypomethylation and human diseases, *Biochim. Biophys. Acta*, 1775 (2007) 138-162.
- [93] R. Chatterjee, C. Vinson, CpG methylation recruits sequence specific transcription factors essential for tissue specific gene expression, *Biochim. Biophys. Acta*, 1819 (2012) 763-770.
- [94] H.L. Asch, J.S. Winston, S.B. Edge, P.C. Stomper, B.B. Asch, Down-regulation of gelsolin expression in human breast ductal carcinoma in situ with and without invasion, *Breast Cancer Res. Treat.*, 55 (1999) 179-188.

- [95] A.R. Florl, C. Steinhoff, M. Muller, H.H. Seifert, C. Hader, R. Engers, R. Ackermann, W.A. Schulz, Coordinate hypermethylation at specific genes in prostate carcinoma precedes LINE-1 hypomethylation, *Br. J. Cancer*, 91 (2004) 985-994.
- [96] W.A. Schulz, M.J. Hoffmann, Epigenetic mechanisms in the biology of prostate cancer, *Semin. Cancer Biol.*, 19 (2009) 172-180.
- [97] L.C. Li, S.T. Okino, R. Dahiya, DNA methylation in prostate cancer, *Biochim. Biophys. Acta*, 1704 (2004) 87-102.
- [98] M.J. Kim, J.A. White-Cross, L. Shen, J.P. Issa, A. Rashid, Hypomethylation of long interspersed nuclear element-1 in hepatocellular carcinomas, *Mod. Pathol.*, 22 (2009) 442-449.
- [99] C.M. Suter, D.I. Martin, R.L. Ward, Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue, *Int. J. Colorectal Dis.*, 19 (2004) 95-101.
- [100] R.C. Iskow, M.T. McCabe, R.E. Mills, S. Torene, W.S. Pittard, A.F. Neuwald, E.G. Van Meir, P.M. Vertino, S.E. Devine, Natural mutagenesis of human genomes by endogenous retrotransposons, *Cell*, 141 (2010) 1253-1261.
- [101] E. Lee, R. Iskow, L. Yang, O. Gokcumen, P. Haseley, L.J. Luquette, 3rd, J.G. Lohr, C.C. Harris, L. Ding, R.K. Wilson, D.A. Wheeler, R.A. Gibbs, R. Kucherlapati, C. Lee, P.V. Kharchenko, P.J. Park, Landscape of Somatic Retrotransposition in Human Cancers, *Science*, (2012).
- [102] J. Lu, G. Getz, E.A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B.L. Ebert, R.H. Mak, A.A. Ferrando, J.R. Downing, T. Jacks, H.R. Horvitz, T.R. Golub, MicroRNA expression profiles classify human cancers, *Nature*, 435 (2005) 834-838.
- [103] B. Czech, G.J. Hannon, Small RNA sorting: matchmaking for Argonautes, *Nat. Rev. Genet.*, 12 (2011) 19-31.

- [104] R.K. Slotkin, R. Martienssen, Transposable elements and the epigenetic regulation of the genome, *Nat. Rev. Genet.*, 8 (2007) 272-285.
- [105] E. Kolomietz, M.S. Meyn, A. Pandita, J.A. Squire, The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors, *Genes Chromosomes Cancer*, 35 (2002) 97-112.
- [106] B. Elliott, C. Richardson, M. Jasin, Chromosomal translocation mechanisms at intronic alu elements in mammalian cells, *Mol. Cell*, 17 (2005) 885-894.
- [107] D.M. Weinstock, C.A. Richardson, B. Elliott, M. Jasin, Modeling oncogenic translocations: distinct roles for double-strand break repair pathways in translocation formation in mammalian cells, *DNA Repair (Amst)*, 5 (2006) 1065-1074.
- [108] P. Medstrand, L.N. van de Lagemaat, C.A. Dunn, J.R. Landry, D. Svenback, D.L. Mager, Impact of transposable elements on the evolution of mammalian gene regulation, *Cytogenet. Genome Res.*, 110 (2005) 342-352.
- [109] V. Perepelitsa-Belancio, P. Deininger, RNA truncation by premature polyadenylation attenuates human mobile element activity, *Nat. Genet.*, 35 (2003) 363-366.
- [110] A.M. Roy-Engel, M. El-Sawy, L. Farooq, G.L. Odom, V. Perepelitsa-Belancio, H. Bruch, O.O. Oyeniran, P.L. Deininger, Human retroelements may introduce intragenic polyadenylation signals, *Cytogenet. Genome Res.*, 110 (2005) 365-371.
- [111] N. Sela, B. Mersch, N. Gal-Mark, G. Lev-Maor, A. Hotz-Wagenblatt, G. Ast, Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome, *Genome Biol.*, 8 (2007) R127.
- [112] R. Sorek, The birth of new exons: mechanisms and evolutionary consequences, *RNA*, 13 (2007) 1603-1608.
- [113] X.H. Zhang, L.A. Chasin, Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons, *Proc. Natl. Acad. Sci. U. S. A.*, 103 (2006) 13427-13432.

- [114] G. Bourque, B. Leong, V.B. Vega, X. Chen, Y.L. Lee, K.G. Srinivasan, J.L. Chew, Y. Ruan, C.L. Wei, H.H. Ng, E.T. Liu, Evolution of the mammalian transcription factor binding repertoire via transposable elements, *Genome Res.*, 18 (2008) 1752-1762.
- [115] L. Teng, H.A. Firpi, K. Tan, Enhancers in embryonic stem cells are enriched for transposable elements and genetic variations associated with cancers, *Nucleic Acids Res.*, 39 (2011) 7371-7379.
- [116] T. Wang, J. Zeng, C.B. Lowe, R.G. Sellers, S.R. Salama, M. Yang, S.M. Burgess, R.K. Brachmann, D. Haussler, Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53, *Proceedings of the National Academy of Sciences of the United States of America*, 104 (2007) 18613-18618.
- [117] I. Tanaka, H. Ishihara, Enhanced expression of the early retrotransposon in C3H mouse-derived myeloid leukemia cells, *Virology*, 280 (2001) 107-114.
- [118] P. Brulet, M. Kaghad, Y.S. Xu, O. Croissant, F. Jacob, Early differential tissue expression of transposon-like repetitive DNA sequences of the mouse, *Proc. Natl. Acad. Sci. U. S. A.*, 80 (1983) 5641-5645.
- [119] D. Givol, Activation of oncogenes by transposable elements, *Biochemical Society symposium*, 51 (1986) 183-196.
- [120] N.G. Copeland, N.A. Jenkins, Harnessing transposons for cancer gene discovery, *Nat. Rev. Cancer*, 10 (2010) 696-706.
- [121] A.G. Uren, J. Kool, A. Berns, M. van Lohuizen, Retroviral insertional mutagenesis: past, present and future, *Oncogene*, 24 (2005) 7656-7672.
- [122] S.F. Landrette, J.C. Cornett, T.K. Ni, M.W. Bosenberg, T. Xu, piggyBac transposon somatic mutagenesis with an activated reporter and tracker (PB-SMART) for genetic screens in mice, *PLoS One*, 6 (2011) e26650.

- [123] R. Rad, L. Rad, W. Wang, J. Cadinanos, G. Vassiliou, S. Rice, L.S. Campos, K. Yusa, R. Banerjee, M.A. Li, J. de la Rosa, A. Strong, D. Lu, P. Ellis, N. Conte, F.T. Yang, P. Liu, A. Bradley, PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice, *Science*, 330 (2010) 1104-1107.
- [124] A.J. Dupuy, K. Akagi, D.A. Largaespada, N.G. Copeland, N.A. Jenkins, Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system, *Nature*, 436 (2005) 221-226.
- [125] L.S. Collier, C.M. Carlson, S. Ravimohan, A.J. Dupuy, D.A. Largaespada, Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse, *Nature*, 436 (2005) 272-276.
- [126] M. McGrail, J.M. Hatler, X. Kuang, H.K. Liao, K. Nannapaneni, K.E. Watt, J.D. Uhl, D.A. Largaespada, E. Vollbrecht, T.E. Scheetz, A.J. Dupuy, J.M. Hostetter, J.J. Essner, Somatic mutagenesis with a Sleeping Beauty transposon system leads to solid tumor formation in zebrafish, *PLoS One*, 6 (2011) e18826.
- [127] F. Cui, M.V. Sirotin, V.B. Zhurkin, Impact of Alu repeats on the evolution of human p53 binding sites, *Biol. Direct*, 6 (2011) 2.
- [128] J.E. Hambor, J. Mennone, M.E. Coon, J.H. Hanke, P. Kavathas, Identification and characterization of an Alu-containing, T-cell-specific enhancer located in the last intron of the human CD8 alpha gene, *Mol. Cell. Biol.*, 13 (1993) 7056-7070.
- [129] G.W. Humphrey, E.W. Englander, B.H. Howard, Specific binding sites for a pol III transcriptional repressor and pol II transcription factor YY1 within the internucleosomal spacer region in primate Alu repetitive elements, *Gene Expr.*, 6 (1996) 151-168.
- [130] S.L. Oei, V.S. Babich, V.I. Kazakov, N.M. Usmanova, A.V. Kropotov, N.V. Tomilin, Clusters of regulatory signals for RNA polymerase II transcription associated with Alu family repeats and CpG islands in human promoters, *Genomics*, 83 (2004) 873-882.

- [131] F.J. Piedrafita, R.B. Molander, G. Vansant, E.A. Orlova, M. Pfahl, W.F. Reynolds, An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element, *J. Biol. Chem.*, 271 (1996) 14412-14420.
- [132] G. Vansant, W.F. Reynolds, The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element, *Proc. Natl. Acad. Sci. U. S. A.*, 92 (1995) 8229-8233.
- [133] J.H. Taube, K. Allton, S.A. Duncan, L. Shen, M.C. Barton, Foxa1 functions as a pioneer transcription factor at transposable elements to activate Afp during differentiation of embryonic stem cells, *The Journal of biological chemistry*, 285 (2010) 16135-16144.
- [134] H.C. Reinhardt, B. Schumacher, The p53 network: cellular and systemic DNA damage responses in aging and cancer, *Trends Genet.*, 28 (2012) 128-136.
- [135] T. Soussi, TP53 mutations in human cancer: database reassessment and prospects for the next decade, *Adv. Cancer Res.*, 110 (2011) 107-139.
- [136] R.J. White, RNA polymerase III transcription and cancer, *Oncogene*, 23 (2004) 3208-3216.
- [137] R.B. Tang, H.Y. Wang, H.Y. Lu, J. Xiong, H.H. Li, X.H. Qiu, H.Q. Liu, Increased level of polymerase III transcribed Alu RNA in hepatocellular carcinoma tissue, *Mol. Carcinog.*, 42 (2005) 93-96.
- [138] H.C. Kopera, J.B. Moldovan, T.A. Morrish, J.L. Garcia-Perez, J.V. Moran, Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase, *Proc. Natl. Acad. Sci. U. S. A.*, 108 (2011) 20345-20350.
- [139] T.A. Morrish, J.L. Garcia-Perez, T.D. Stamato, G.E. Taccioli, J. Sekiguchi, J.V. Moran, Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres, *Nature*, 446 (2007) 208-212.

- [140] I.R. Arkhipova, H.G. Morrison, Three retrotransposon families in the genome of *Giardia lamblia*: two telomeric, one dead, *Proc. Natl. Acad. Sci. U. S. A.*, 98 (2001) 14497-14502.
- [141] H. Fujiwara, M. Osanai, T. Matsumoto, K.K. Kojima, Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*, *Chromosome Res.*, 13 (2005) 455-467.
- [142] E.A. Gladyshev, I.R. Arkhipova, Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes, *Proc. Natl. Acad. Sci. U. S. A.*, 104 (2007) 9352-9357.
- [143] J.F. Riou, D. Gomez, J.L. Mergny, L. Guittat, R. Paterski, B. Chénais, H. Morjani, C. Trentesaux, [Regulation of telomeres length: making the telomeres accessible?], *Bull. Cancer*, 92 (2005) 13-22.
- [144] J.W. Shay, W.E. Wright, Role of telomeres and telomerase in cancer, *Semin. Cancer Biol.*, 21 (2011) 349-353.
- [145] B.T. Brett, K.E. Berquam-Vrieze, K. Nannapaneni, J. Huang, T.E. Scheetz, A.J. Dupuy, Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors, *PLoS One*, 6 (2011) e24668.
- [146] J. de Jong, J. de Ridder, L. van der Weyden, N. Sun, M. van Uitert, A. Berns, M. van Lohuizen, J. Jonkers, D.J. Adams, L.F. Wessels, Computational identification of insertional mutagenesis targets for cancer gene discovery, *Nucleic Acids Res.*, 39 (2011) e105.
- [147] K.M. Mann, J.M. Ward, C.C. Yew, A. Kovochich, D.W. Dawson, M.A. Black, B.T. Brett, T.E. Sheetz, A.J. Dupuy, D.K. Chang, A.V. Biankin, N. Waddell, K.S. Kassahn, S.M. Grimmond, A.G. Rust, D.J. Adams, N.A. Jenkins, N.G. Copeland, Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma, *Proc. Natl. Acad. Sci. U. S. A.*, 109 (2012) 5934-5941.

[148] T.K. Starr, P.M. Scott, B.M. Marsh, L. Zhao, B.L. Than, M.G. O'Sullivan, A.L. Sarver, A.J. Dupuy, D.A. Largaespada, R.T. Cormier, A Sleeping Beauty transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (Apc)-dependent intestinal tumorigenesis, *Proc. Natl. Acad. Sci. U. S. A.*, 108 (2011) 5765-5770.

Figures legends

Figure 1. Schematic overview of transposable element classification. (A) The classification is mainly based on transposition mechanisms which are briefly summarized below the element drawing. For more details refer to [9, 10]. Examples of TE families in each group are given between brackets, asterisk indicate (super)families not found in human genome whereas underlined names are for (super)families present in human genome according to [8]. The presence and orientation of DNA repeats are indicated by arrows (TIR, terminal inverted repeat; DR, direct repeat). Grey arrows indicate a repeated sequence or palindrome within the DNA. (B) Schematic structure of SINE and SVA non-autonomous class-I elements. In the SINE sequence, the A and B boxes represent the internal promoter stretch; the violet box figure a LINE-1-derived sequence. SVA is a composite element containing from 5' to 3' a hexamer (CCCTCT)_n repeat, an *Alu*-homologous region of two antisense *Alu* fragments including other sequence of unknown origin, a variable number of tandem repeats (VNTR) region, a SINE-R region derived from the human endogenous retrovirus HERV-K10, and a polyadenylation site (pA) immediately before polyA tail. Both SINE and SVA are flanked by target site duplications (TSD, brown box).

Figure 2. Schematic overview of the transposable elements repartition in the human genome. Transposable element as a whole represents about 45% of the human genome. The diagram represents the percentage of each TE category compared to all transposable elements (first %) and over the entire human genome (% in brackets). Data according to [8].

Table 1. Examples of TE insertion and TE-mediated chromosomal rearrangements associated with cancer.

Locus and/or gene	Associated cancer	TE	distribution	references
<i>Insertion</i>				
APC, adenomatous polyposis coli gene	Desmoids tumors	<i>Alu</i>	germline	[28]
APC	Colon cancer	L1	germline	[26]
APC		L1	somatic	[27]
BRCA1, breast cancer 1 gene	Breast/ovarian cancer	<i>Alu</i>	germline	[31]
BRCA2, breast cancer 2 gene	Breast/ovarian cancer	<i>Alu</i>	germline	[30, 31]
MYC, c-myc proto-oncogene	Breast carcinoma	L1	somatic	[24]
NF1, neurofibromatosis 1 gene	Neurofibroma	<i>Alu</i>	germline	[29]
<i>Chromosomal deletions</i>				
VHL, von Hippel Lindau gene	von Hippel Lindau disease	<i>Alu</i>	germline	[43, 44]
BRCA1	Breast/ovarian cancers	<i>Alu</i>	germline	[32, 33]
BRCA2	Breast/ovarian cancers	<i>Alu</i>	germline	[33, 35]
CDH1, cadherin 1 gene	Hereditary diffuse gastric cancer	<i>Alu</i>	germline	[45]
CAD, caspase activated DNase gene	Hepatoma	<i>Alu</i>	somatic	[46]
<i>Chromosomal duplication</i>				
MLL1, myeloid/lymphoid mixed lineage leukemia gene	Acute myeloid leukemia	<i>Alu</i>	somatic	[37, 38]
MYB, myb transcription factor gene	T-acute lymphoblastic lymphoma	<i>Alu</i>	somatic	[41]
BRCA1	Breast/ovarian cancers	<i>Alu</i>	germline	[33, 34]
<i>Chromosomal translocation</i>				
EWSR1-ETV, t(5q23q31)(18q12)	Ewing sarcoma	<i>Alu</i>	somatic	[42]
BCR-ABL, t(9;22)(q34;q11)	Chronic myeloid leukemia	<i>Alu</i>	somatic	[36]

A

Class I: retrotransposons

Reverse Transcriptase (RT) copies genome into DNA

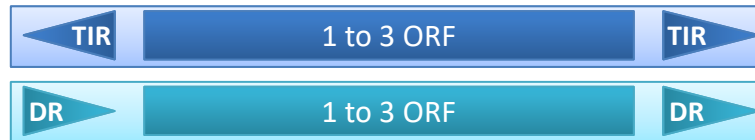
Long Terminal Repeat (LTR) elements

(*copia**, *gypsy**, *ERVs*)



Double stranded DNA is integrated using a transposase-related integrase

DIRS elements* (*DIRS**, *Ngaro**, *viper**)



DNA is integrated using a tyrosine recombinase

Non-LTR elements (*LINE*, *Penelope**)

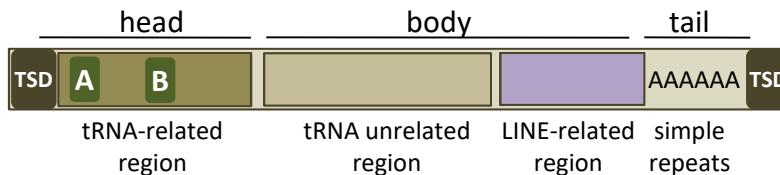


Integration and priming of RT reaction is mediated by an endonuclease

B

Non-autonomous retroelements (*SINE*, *SVA*)

Typical SINE



Class II: DNA transposons

DNA genome itself serves as the template for transposition

Terminal Inverted Repeat (TIR) elements

(*Tc1-Mariner*, *PiggyBac*, *P**, ...)



DNA is cleaved from donor site and integrated at target site using transposase

Helitron elements*



Single strand DNA intermediate is replicated from donor site by a rolling-circle mechanism

Maverick elements*



Both viral-like DNA polymerase and retroviral-like integrase are thought to be involved in double stranded DNA integration

SVA



