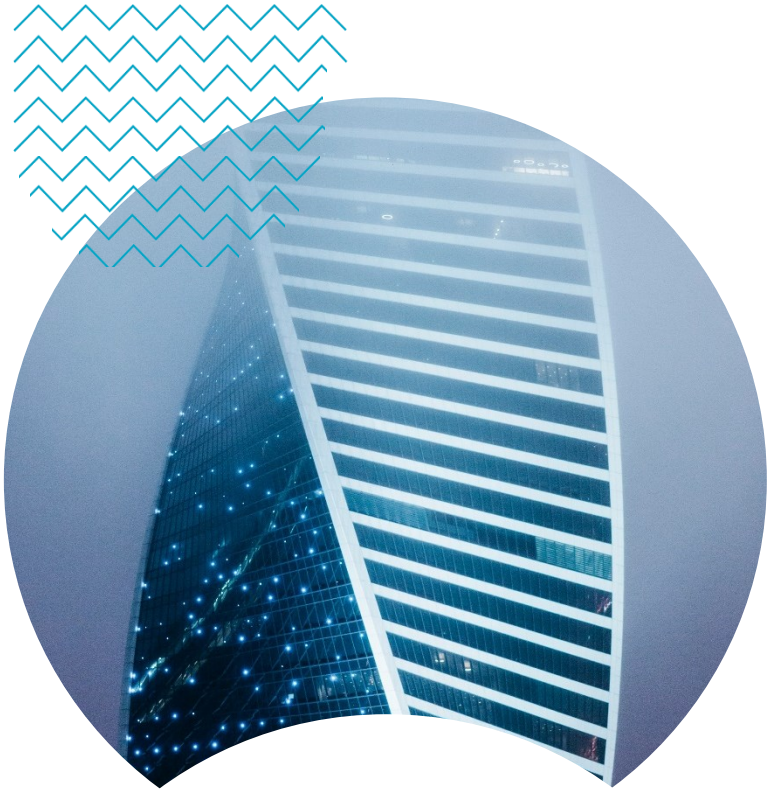
myNEO Therapeutics



Transposable Elements in Cancer

myNEO NV
Ottergemsesteenweg-Zuid 808
9000 Ghent, Belgium

www.myneotx.com
contact@myneotx.com

Abstract

Transposable elements, once considered genomic junk, have recently been recognized as key players in cancer development. Normally silenced in terminally differentiated cells, transposable elements can be reactivated by epigenetic dysregulation during cancer transformation. This whitepaper discusses how their reactivation contributes to cancer progression and how it can result in biomarkers and neoantigens for cancer diagnosis and treatment.

Introduction

Transposable elements (TEs), which account for almost 50% of the human genome, are **repetitive DNA sequences that have the ability to change their position** within a genome. By moving through the genome, they can alter gene expression or modify gene products, leading to genetic instability [1,2].

TEs are broadly divided into two major classes based on their moving mechanism and transposition intermediate. The first class is called **retrotransposons or RNA transposons** and these move through a "copy-and-paste" mechanism via reverse-transcribed RNA intermediates (Figure 1A), while class II TEs, also known as **DNA transposons**, utilize a "cut-and-paste" mechanism, excising themselves from one location and reinserting into another within the genome (Figure 1B)[2,3].

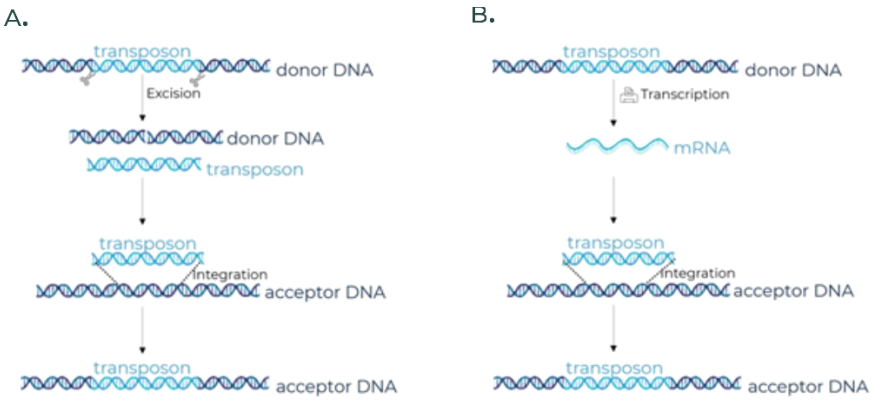


Figure 1: Mobilization mechanisms of A. Retrotransposons versus B. transposons.

Furthermore, class I TEs can be categorized into long terminal repeat (LTR) and non-LTR retrotransposons based on the presence or absence of LTRs at both ends of the transposon respectively (Figure 2).

Although generally silenced through epigenetic mechanisms, TEs can become dysregulated in certain conditions, such as cancer, where their transcriptional silencing is lost[4]. Understanding the regulation and activity of TEs is therefore crucial to determine their impact on genome evolution and disease progression.

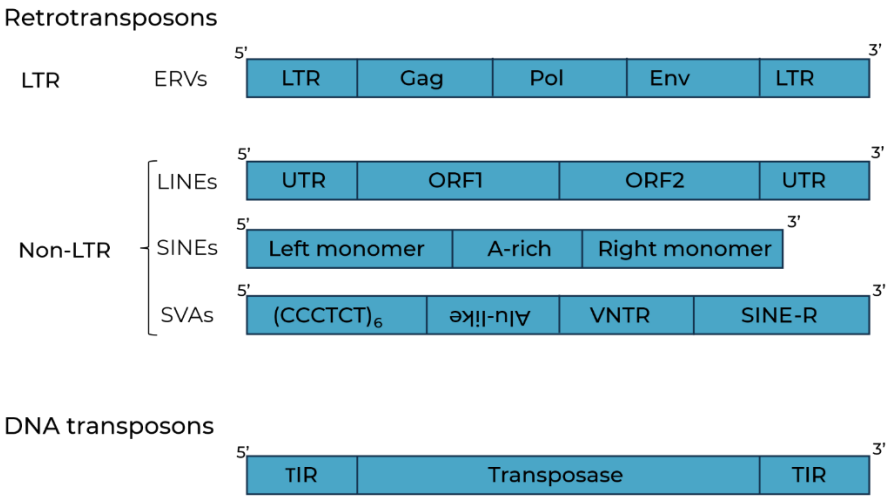


Figure 2: Non-exhaustive classification of the different types of transposable elements. LTR (long terminal repeat), Gag Pol Env (3 genes related to retroviral genes), LINE (long interspersed nuclear element), UTR (untranslated region), ORF (open reading frame), SINE (short interspersed nuclear element), SVA (SINE-VNTR-Alu), VNTR (variable number tandem repeat), TIR (terminal inverted repeat)

Transposable elements in cancer

TEs are increasingly recognized as contributors to tumorigenesis, particularly due to their ability to **disrupt gene regulation and promote genomic instability**. Indeed, cancer cells – which arise from their normal counterparts through the sequential acquisition of genetic and epigenetic changes – demonstrate widespread TE reactivation[1,5].

One of the most common mechanisms of TE reactivation in cancer involves the **loss of DNA methylation**[5]. Hypomethylation of the Long Interspersed Nuclear Element 1 (LINE-1, a non-LTR retrotransposon) promoter results in its reactivation, leading to DNA damage, chromosomal rearrangements, and **gene disruption** (Figure 3A). For example, in colorectal cancer, somatic L1 insertions have been shown to cause genetic disruption in the tumor suppressor gene Adenomatous polyposis coli or APC, leading to its inactivation[1,5–7].

Besides causing the TEs to transpose, hypomethylation can also result in onco-exaptation. **Onco-exaptation** is the process where hypomethylation exposes normally inaccessible regulatory regions in TEs which can serve as cryptic promotors for oncogenes (Figure 3B)[7–9]. TEs in general, but more specifically endogenous retroviruses (ERVs, a type of LTR-retrotransposons), form a significant reservoir of promoter sequences in the human genome which makes them an important target of onco-exaptation. For example, in Hodgkin's lymphoma, hypomethylation of LTR elements of ERVs has been found to drive the expression of oncogene Interferon regulatory factor 5 (IRF5), and these LTR-IRF5 chimeric transcripts were not found in healthy B-cell controls[1,6,10,11].

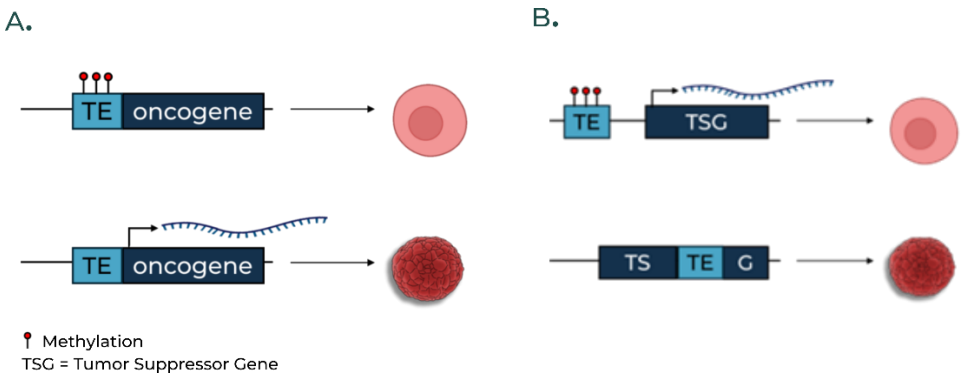


Figure 3: Different effects of TE reactivation by demethylation on cancer development. A. Demethylation reactivates the transposing mechanism of the TE, resulting in insertion and genetic disruption of the TSG. B. Demethylation makes the promotor located in the TE accessible, leading to transcription of the adjacent oncogene.

Detection of cancer-specific TE insertions

One of the consequences of their repetitive nature is that TE-derived short-sequencing reads commonly map to different positions in the genome during the alignment step. Discarding these ambiguously mapped reads leads to a loss of the signal generated by TEs that are still able to transpose, including the TEs reactivated during cancer development, while keeping them increases the risks for False Positives[15–17]. Various alignment tools including BWA[18] and bowtie[19], allow **multi-mapping** by assigning mapping quality based on the likelihood of sequencing errors or alignment probability and by randomly selecting a subset of positions for reads that map equally well to multiple locations. This approach **balances the need to capture potential alignments while managing ambiguity**, which is common in repetitive genomic regions.

Following alignment, the detection of novel TE insertions is typically based on one, or an ensemble approach, of the following three strategies: split-read (SR) analysis, discordant read pair (DRP) alignment, or searching for TE-specific sequence motifs. **Split-Read-based methods** identify reads spanning insertion breakpoints, where part of the read aligns to the reference genome while the remaining portion corresponds to the start or end of a TE. **Discordant Read Pair analysis**, on the other hand, examines paired-end reads and flags cases where the two read ends are unusually distant, have opposite orientations, or where only one read aligns to the reference genome. A third approach focuses on **identifying known TE-specific motifs** within sequencing reads, rather than relying on alignment patterns. Such common TE insertion signatures are LTRs associated with LTR retrotransposons or terminal inverted repeats (TIRs) found in DNA transposons.

Several specialized tools have been developed to detect tumor-specific TE insertions which typically analyze both tumor and normal samples[16]. For instance, TranspoSeq[20] and TranspoSeq-Exome[21] combine motif, SR, and DRP strategies to identify cancer-specific TE insertions, while Jitterbug[22] employs SR and DRP techniques to detect somatic L1 insertions in cancer.

Targeting TEs in Immunotherapy

T-cells are the main driver of anti-tumor immunity and as such vital to the clinical success of immune checkpoint inhibitors (ICIs) such as anti-PD-1 and anti-CTLA-4. Despite promising results, many malignancies still do not respond well to checkpoint immunotherapy mainly in patients having no or very limited spontaneous T-cell activity. However, by combining ICIs with neoantigen-specific strategies, such as neoantigen vaccines, immune responses can be improved by immune (re)activation prior to the ICI therapy

In general, single-nucleotide variants (SNV) and small insertion/deletion mutations (indels) are the most common sources for neoantigens. However, in cancers with a low tumor mutational burden, it is often impossible to find a sufficient number of actionable SNVs and indels to elicit an effective anti-tumor response[23].

TE- derived neoantigens could serve as an additional source of neoantigens that can play a valuable role. The latter targets can involve **tumor specific antigens (TSAs)** as well as **tumor-associated antigens (TAAs)**. In the case of TSAs, a TE sequence is inserted within the coding region of a regular gene of cancer cells, leading to an altered protein that is only present in cancer cells. This type of neoantigens is also called non-classical splice-derived neoantigens, describing the conjunction of TEs with canonical exons of protein-coding genes[24]. On the other hand, TE-derived TAAs are TE proteins that are exclusively expressed in cancer cells caused by epigenetic dysregulation in cancer cells, naturally silenced in healthy cells.

Shah et al. analyzed TCGA tumors for chimeric RNA products, meaning fusion products where a TE sequence was fused with a gene sequence. They reported high levels of sharedness of these variants within and between cancer types, with some chimeric proteins reaching up to 50% coverage in single tumor types[24]. Similarly, Bonté et al. identified HLA class I-bound peptides originating from TEs in glioblastoma (GBM)patients and found a subset of highly recurrent, GBM-specific peptides which they propose as potential targets for cancer therapy[25]. These findings further underscore the potential of TE-derived sequences in identifying novel immunotherapeutic targets.

TE-derived neoantigens are known for their **inherent potential for immunogenicity**, offering a unique advantage for immunotherapeutic strategies. Studies have shown that expression of transposable elements in tumors is linked to higher immune infiltration and recognition by immune cells[26]. One of the reasons for the increased immunogenicity of TE-derived neoantigens is their **similarity to viral elements** due to their ancient viral origin.

TE Load as a Diagnostic Biomarker

Due to the fact that patients with tumors exhibiting higher levels of TE activity often show better responses to ICI therapies, using TEs as a **biomarker for response prediction** is being explored. Particularly in low TMB tumors, TMB has been an unreliable diagnostic biomarker for ICI therapy, and other biomarkers are needed[27]. Therefore, measuring TE activity in tumors could help to improve predictions about which patients are more likely to benefit from immunotherapy. For instance, Smith *et al.* were able to link the expression of ERVs, with the response to programmed death receptor 1 (PD-1) inhibition therapy[28].

TEs, particularly LINE1 (long interspersed nuclear elements), have also gained attention as potential biomarkers for **early cancer detection**, especially through liquid biopsies. Hypomethylation of LINE1 is commonly observed in cancer, leading to increased TE activity and genomic instability[29]. Research indicates that L1 hypomethylation occurs very early in cancer progression. This epigenetic alteration can be detected in circulating tumor DNA, offering a minimally invasive approach to identifying early-stage tumors[30]. Monitoring LINE1 hypomethylation in liquid biopsies could provide a valuable tool for early detection, complementing other biomarkers and enabling earlier interventions.

Outlook

Transposable elements offer novel opportunities to target the often-overlooked "dark genome" for therapeutic purposes. Their unique ability to generate neoantigens positions them as ideal candidates for personalized immunotherapies, particularly in tumors with low actionable neoantigen loads. Moreover, the quantification of TE load and activity can complement existing biomarkers, aiding in the stratification of patients and potentially predicting response to treatment.

Furthermore, recent advances in epigenetic therapies highlight the potential of inducing TE expression through epigenetic drugs[31]. In combination with other immunotherapies this can be a powerful strategy for poorly immunogenic tumors to increase their mutational load.

It should be noted that while TE-gene chimeric transcripts are one of the easiest to detect sources of TE-derived neoantigens, they are not necessarily the most prevalent. Proteins originating from intergenic TEs can be more widespread than chimeric transcripts and even have the potential to be shared within cancer indications. To accurately identify these non-canonical cancer peptides, an integrated approach between sequencing datasets and mass-spec ligandome sets is essential. However, confirming the presence or absence of these peptides with high enough sensitivity is challenging.[29]

About myNEO Therapeutics

myNEO Therapeutics is a distinguished biopharmaceutical powerhouse, dedicated to pioneer breakthrough immunotherapies to fight cancer. We are leveraging our discovery ImmunoEngine platform to tap into novel promising tumor targets found in the dark genome – named camyotopes™ – which have the potential to unlock immunotherapy for large patient populations who do currently not respond. myNEO Therapeutics is one of the companies that emerged from the Novalis biotech incubator fund at the end of 2018, founded by two leading entrepreneurs already known for several successes in the biotech industry: Wim Van Criekinge, professor of computational biology at Ghent University, and childhood friend Jan Van den Berghe.

Interested in more information about myNEO Therapeutics? Contact us!

Lien Lybaert, PhD.
Chief Development Officer
lien.lybaert@myneotx.com

References

1. Grundy, E. E., Diab, N. & Chiappinelli, K. B. Transposable element regulation and expression in cancer. *FEBS J* **289**, 1160–1179 (2022).
2. Wang, Z. Y., Ge, L. P., Ouyang, Y., Jin, X. & Jiang, Y. Z. Targeting transposable elements in cancer: developments and opportunities. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1879**, 189143 (2024).
3. Lerat, E. & Sémon, M. Influence of the transposable element neighborhood on human gene expression in normal and tumor tissues. *Gene* **396**, 303–311 (2007).
4. Jang, H. S. *et al.* Transposable elements drive widespread expression of oncogenes in human cancers. *Nature Genetics 2019 51:4* **51**, 611–617 (2019).
5. Lee, M., Ahmad, S. F. & Xu, J. Regulation and function of transposable elements in cancer genomes. *Cellular and Molecular Life Sciences 2024 81:1* **81**, 1–27 (2024).
6. Pradhan, R. K. & Ramakrishna, W. Transposons: Unexpected players in cancer. *Gene* **808**, 145975 (2022).
7. Wang, Z. Y., Ge, L. P., Ouyang, Y., Jin, X. & Jiang, Y. Z. Targeting transposable elements in cancer: developments and opportunities. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1879**, 189143 (2024).
8. Karttunen, K., Patel, D. & Sahu, B. Transposable elements as drivers of dedifferentiation: Connections between enhancers in embryonic stem cells, placenta, and cancer. *BioEssays* **46**, 2400059 (2024).
9. Grundy, E. E., Diab, N. & Chiappinelli, K. B. Transposable element regulation and expression in cancer. *FEBS J* **289**, 1160–1179 (2022).
10. Lynch-Sutherland, C. F., Chatterjee, A., Stockwell, P. A., Eccles, M. R. & Macaulay, E. C. Reawakening the Developmental Origins of Cancer Through Transposable Elements. *Front Oncol* **10**, 521163 (2020).
11. Babaian, A. & Mager, D. L. Endogenous retroviral promoter exaptation in human cancer. *Mob DNA* **7**, 1–21 (2016).
12. Lynch-Sutherland, C. F., Chatterjee, A., Stockwell, P. A., Eccles, M. R. & Macaulay, E. C. Reawakening the Developmental Origins of Cancer Through Transposable Elements. *Front Oncol* **10**, 521163 (2020).
13. Kunarso, G. *et al.* Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature Genetics 2010 42:7* **42**, 631–634 (2010).
14. Liu, A., Yu, X. & Liu, S. Pluripotency transcription factors and cancer stem cells: small genes make a big difference. *Chin J Cancer* **32**, 483 (2013).

15. Lanciano, S. & Cristofari, G. Measuring and interpreting transposable element expression. *Nat Rev Genet* **21**, 721–736 (2020).
16. Goerner-Potvin, P. & Bourque, G. Computational tools to unmask transposable elements. *Nat Rev Genet* **19**, 688–704 (2018).
17. O'Neill, K., Brocks, D. & Hammell, M. G. Mobile genomics: tools and techniques for tackling transposons. *Philosophical Transactions of the Royal Society B* **375**, (2020).
18. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754 (2009).
19. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, 1–10 (2009).
20. Elena Helman ScB, by, Meyerson, M. & Brown, E. N. Somatic retrotransposition in the cancer genome. (2013).
21. Helman, E. *et al.* Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res* **24**, 1053 (2014).
22. Hénaff, E., Zapata, L., Casacuberta, J. M. & Ossowski, S. Jitterbug: Somatic and germline transposon insertion detection at single-nucleotide resolution. *BMC Genomics* **16**, 1–16 (2015).
23. Smith, C. C. *et al.* Alternative tumour-specific antigens. *Nature Reviews Cancer 2019 19:8* **19**, 465–478 (2019).
24. Shah, N. M. *et al.* Pan-cancer analysis identifies tumor-specific antigens derived from transposable elements. *Nature Genetics 2023 55:4* **55**, 631–639 (2023).
25. Bonté, P. E. *et al.* Single-cell RNA-seq-based proteogenomics identifies glioblastoma-specific transposable elements encoding HLA-I-presented peptides. *Cell Rep* **39**, (2022).
26. Kong, Y. *et al.* Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nature Communications 2019 10:1* **10**, 1–14 (2019).
27. Jardim, D. L., Goodman, A., de Melo Gagliato, D. & Kurzrock, R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer Cell* **39**, 154–173 (2021).
28. Smith, C. C. *et al.* Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J Clin Invest* **128**, 4804–4820 (2019).
29. Zheng, Y. *et al.* Prediction of genome-wide DNA methylation in repetitive elements. *Nucleic Acids Res* **45**, 8697–8711 (2017).
30. Sunami, E., de Maat, M., Vu, A., Turner, R. R. & Hoon, D. S. B. LINE-1 Hypomethylation During Primary Colon Cancer Progression. *PLoS One* **6**, e18884 (2011).
31. Goyal, A. *et al.* DNMT and HDAC inhibition induces immunogenic neoantigens from human endogenous retroviral element-derived transcripts. doi:10.1038/s41467-023-42417-w.