

# 1 **Cancer immunotherapy: how to exploit transposable elements?**

2

3 Sophie Lanciano\* and Gael Cristofari\*

4

5 University Cote d'Azur, Inserm, CNRS, Institute for Research on Cancer and Aging of Nice (IRCAN), Nice, France

6 \* Co-corresponding authors. Correspondence should be addressed to:

7 [Sophie.Lanciano@univ-cotedazur.fr](mailto:Sophie.Lanciano@univ-cotedazur.fr) or [Gael.Cristofari@univ-cotedazur.fr](mailto:Gael.Cristofari@univ-cotedazur.fr)

8

9 Immunotherapy is one of the most promising advances in the fight against cancer and is beginning to  
10 join more conventional treatment approaches such as surgery, radiation therapy or chemotherapy. A  
11 set of strategies consists of ensuring that the immune cells recognize tumor-specific antigens (TSAs)  
12 exposed at the surface of tumor cells and eliminate the tumor. TSAs can be processed within tumor  
13 cells and presented by human leukocyte antigen (HLA) molecules at the cell membrane, where they  
14 can be recognized by cytotoxic T lymphocytes. Additionally, the external region of membrane proteins  
15 can act as TSAs independently of HLA presentation. They can be recognized by antibodies, including  
16 therapeutic monoclonal antibodies, and thus mark tumor cells for elimination by another set of  
17 immune cells. Membrane proteins are also potential targets for chimeric antigen receptor (CAR) T-cells,  
18 which are patient-derived T-cells engineered *ex vivo* to recognize a specific antigen in an HLA-  
19 independent manner, then expanded and reinjected into the patient. Discovering TSAs with the  
20 strongest immunogenicity and specificity is essential for the development of efficient and safe  
21 immunotherapies. Moreover, the identification of TSAs that are shared by many tumor types and cases  
22 could enable the development of universal and off-the-shelf cancer vaccines or therapies, as opposed  
23 to the current state of more personalized but expensive and time-consuming therapies.

24 Currently, the prevailing approach to immunotherapy is to target neoantigens resulting from somatic  
25 mutations in canonical coding genomic regions, with a preference for recurrent mutations. However,  
26 recent findings suggest that the presumed non-coding part of the genome, particularly transposable

27 element (TE)-derived sequences, may represent a predominant source of highly immunogenic and  
28 recurrent TSAs. TEs are repeated sequences that can disperse and replicate through a cut-and-paste or  
29 copy-and-paste mechanism and comprise nearly half of the human genome (1). Some of the TEs from  
30 the L1, Alu or SVA families are able to replicate in modern humans and are major drivers of genome  
31 evolution, active in the germ line or the early embryo but silenced in most somatic tissues.  
32 Nevertheless, they are reactivated in many cancers, presumably because of the profound epigenetic  
33 remodeling that occurs in tumors, such as the global alteration of DNA methylation. Consequently, cis-  
34 regulatory sequences carried over by TEs, particularly their embedded promoters, become  
35 derepressed. As TEs are frequently inserted in introns or close to genes and often contain splice sites,  
36 TEs thereby reactivated can function as alternative promoters and first exons for neighboring genes,  
37 resulting in the synthesis of chimeric transcripts subsequently translated into noncanonical proteins.  
38 Of note, this phenomenon not only involves replication-competent TEs, but also fossil sequences that  
39 have lost their capacity for mobilization for millions of years.

40 A first hint that TSAs could originate from TEs came from an unbiased proteomic approach based on  
41 the identification of HLA-bound peptides by mass spectrometry in murine cell lines and in a handful of  
42 human primary lung tumors or leukemic cells (2). However, a recent study by Ting Wang and colleagues  
43 has now considerably extended the breadth of this initial observation (3). They developed a  
44 computational tool to systematically identify tumor-specific chimeric transcripts and applied it to more  
45 than 10,000 tumor samples across 33 cancer types from The Cancer Genome Atlas (TCGA) datasets,  
46 and to 675 cancer cell lines (**Figure 1, a**). To ensure tumor-specificity, candidate chimeric transcripts  
47 were filtered out using normal samples from TCGA, or from other comprehensive catalogues of normal  
48 tissue expression such as GTEx or FANTOM. Ultimately, they identified ~2,300 tumor-specific TE-  
49 chimeric transcripts for which nearly half were predicted to encode chimeric or out-of-frame proteins.  
50 Going further, they examined tumor mass spectrometry data available from the Clinical Proteomic  
51 Tumor Analysis Consortium and show that some of the predicted peptides are indeed translated in  
52 tumors. They also performed HLA-precipitation mass spectrometry experiments in 4 cancer cell lines,

53 validating between 1 and 4 TE-derived tumor-specific peptides as TSAs with strong predicted affinities  
54 for HLA molecules, per cell line (**Figure 1, b**).

55 With the same idea in mind, Sebastian Amigorena, Joshua Waterfall, Marianne Burbage, and colleagues  
56 have recently focused on peptides derived from TE-exons junctions (JETs) as potential TSAs in humans  
57 and mouse models (4,5). While their study was focused on a particular type of cancer – non-small cell  
58 lung cancer – they too have shown the expression of TE-derived TSAs that are presented by HLA at the  
59 surface of primary lung tumors and can be shared by different patients. Notably, they also provided  
60 functional evidence of immunogenicity *ex vivo*, as well as the presence of specific T cells in tumors and  
61 lymph-nodes of patients, indicative of an immune response directed against the tumor and mediated  
62 by TE-derived TSAs. In mouse models, both prophylactic and therapeutic immunization with TE-derived  
63 TSAs could also delay tumor growth and improve survival (5).

64 Interestingly, most analyzed tumors (94%) express at least one such chimeric transcript, independently  
65 of the tumor type (3). Some tumor-specific TE-chimeric transcripts are recurrently detected in a wide  
66 range of tumors. For example, the *L1P2\_STK31* chimeric transcript is present in 15% of the samples  
67 across all cancer types (3). Others are found in more than 30% of lung adenocarcinoma or squamous  
68 cell carcinoma cases (4). Altogether, combining 20 TE-derived TSAs was predicted to be found in up to  
69 two-thirds of the TCGA samples across all cancer types (3), opening the possibility of large-spectrum  
70 and even universal pan-cancer vaccines.

71 Finally, in addition to HLA-bound peptides, Ting Wang and colleagues also spotted tumor-specific TE  
72 chimeric transcripts encoding membrane proteins that could be targeted by CAR-T cells or monoclonal  
73 antibodies (**Figure 1, c-d**) (3). This finding parallels the recent discovery that some TEs (endogenous  
74 retroviruses or ERV) encode envelope glycoproteins exposed at the membrane of tumor cells, which  
75 can elicit a B-cell response with antitumoral activity and are particularly important in achieving immune  
76 checkpoint blockade (see Supplemental Information).

77 A key aspect of these efforts is to ensure the tumor specificity of the discovered epitopes. Given the  
78 repeated, interspersed – and sometimes polymorphic – nature of TEs, some TE-derived epitopes may  
79 originate from multiple loci and be redundant (1), which could complicate detection in healthy tissues.  
80 With respect to this issue, the focus on peptides corresponding to TE-exon junctions (4,5) represents a  
81 clear benefit, but also narrows the range of possibilities. Recent work showing the expression of some  
82 TEs in normal somatic tissues, particularly epithelia from the gastro-intestinal tract (see Supplemental  
83 Information), will also help with defining the most tumor-specific antigens to limit toxicity in normal  
84 tissues. Different classes and many families of TEs were found capable to drive TSA synthesis. Despite  
85 their repeated nature, individual copies have acquired unique reactivation capacities in tumors, with  
86 modalities depending on the cancer type. In mouse models, ablation of Setdb1, a histone 3-lysine 9  
87 (H3K9me3) methyltransferase which controls heterochromatin formation, broadens the spectrum of  
88 immunogenic TE-exon junctions expressed in tumor cells, suggesting that carefully designed epigenetic  
89 therapies could be combined with TE-based immunotherapies (**Figure 1, g**) (5). Therefore, a deeper  
90 understanding of the mechanisms behind the silencing of individual TE loci is essential to fully deliver  
91 the promises of TE-based biomarkers and immunotherapies, while maintaining a high-level of  
92 specificity and safety. Dedicated approaches are emerging to identify expressed TE loci, to explore the  
93 full diversity of TE transcripts (unit-length TE transcripts, chimeric gene-TE transcripts, or pervasive  
94 transcription), and to study their regulation (1).

95 Beyond providing an unsuspected reservoir of recurrent and immunogenic tumor-specific antigens,  
96 with multiple applications towards the development of specific immunotherapies (**Figure 1, h**;  
97 universal cancer vaccines, new CAR-T cells, response to immune checkpoint blockade), TE expression  
98 in tumors can also trigger a viral mimicry state mediated by innate immune responses through double-  
99 stranded RNA and cytosolic DNA sensing (see Supplemental Information, and **Figure 1, e-f**). Thus,  
100 altogether, TE activity in tumors is becoming a promising and actionable Trojan horse for antitumoral  
101 therapies, acting both through innate and acquired immunity. Combining expertise in cancer biology,

102 immunology, genomics, and bioinformatics will be pivotal to understand the role of these elements in  
103 cancer and fully uncover how they can be leveraged to create new vulnerabilities in tumors.

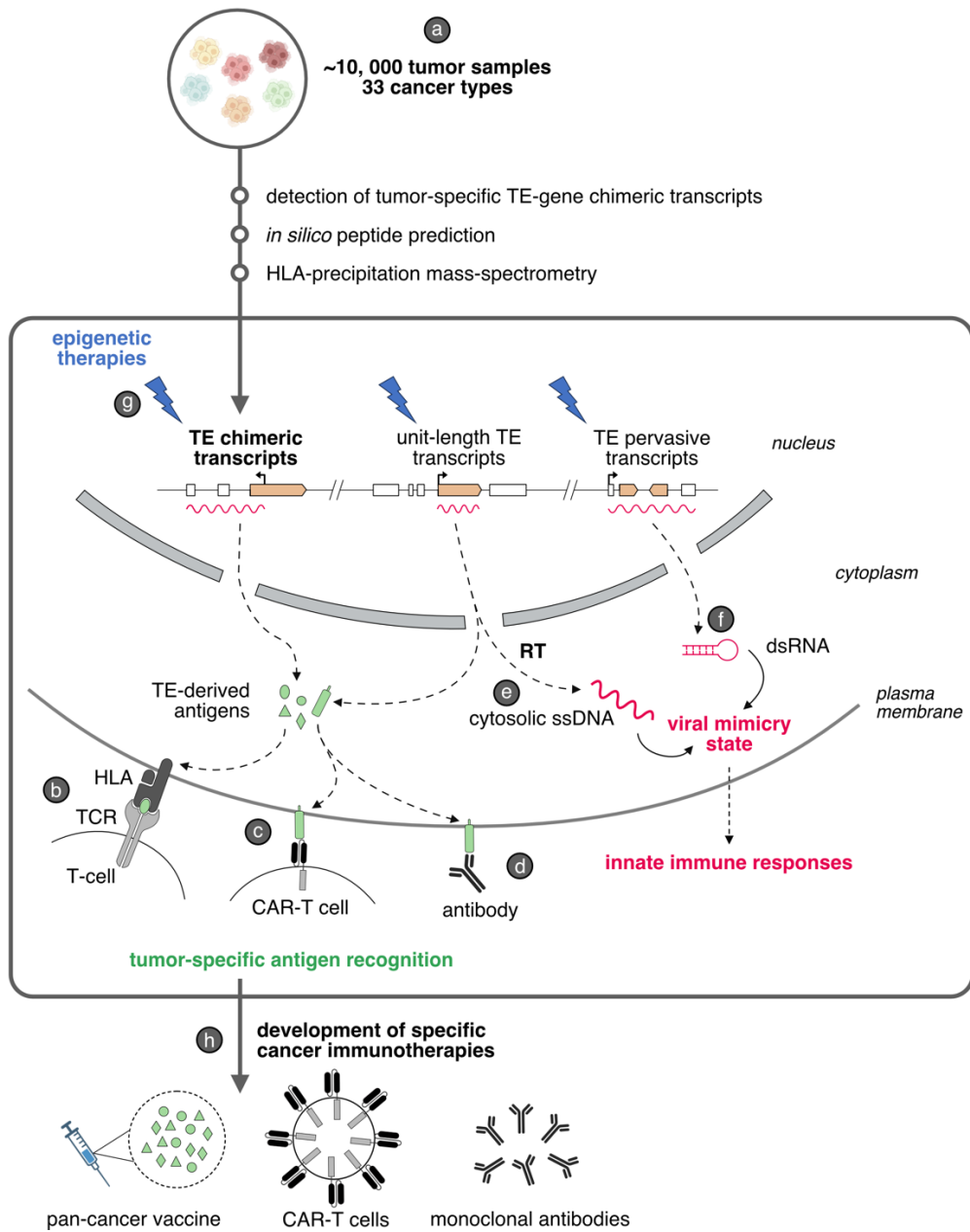
## 104 **ACKNOWLEDGMENTS**

105 This work was supported by the Agence Nationale de la Recherche (Labex SIGNALIFE, ANR-11-LABX-  
106 0028; Idex UCAJEDI, ANR-15-IDEX-01; ImpacTE, ANR-19-CE12-0032; ActiveLINE, ANR-21-CE12-0001),  
107 the Institut National du Cancer (INCa PLBIO 2020-095), the Canceropôle PACA, INCa and the Région  
108 Sud (Projet Emergence), Inserm (GOLD cross-cutting programme on genomic variability), CNRS (GDR  
109 3546). This work was funded, in whole or in part, by the Agence Nationale de la Recherche, Grant  
110 ANR-21-CE12-0001. A CC-BY public copyright license has been applied by the authors to the present  
111 document and will be applied to all subsequent versions up to the Author Accepted Manuscript  
112 arising from this submission, in accordance with the grant's open access conditions.

## 113 **REFERENCES**

- 114 1. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression. *Nat Rev Genet*  
115 2020;21:721–36
- 116 2. Laumont CM, Vincent K, Hesnard L, Audemard É, Bonneil É, Laverdure J-P, et al. Noncoding regions are the  
117 main source of targetable tumor-specific antigens. *Sci Transl Med* 2018;10:eaau5516
- 118 3. Shah NM, Jang HJ, Liang Y, Maeng JH, Tzeng S-C, Wu A, et al. Pan-cancer analysis identifies tumor-specific  
119 antigens derived from transposable elements. *Nat Genet* 2023;55:631–9
- 120 4. Merlotti A, Sadacca B, Arribas YA, Ngoma M, Burbage M, Goudot C, et al. Noncanonical splicing junctions  
121 between exons and transposable elements represent a source of immunogenic recurrent neo-antigens in  
122 patients with lung cancer. *Sci Immunol* 2023;8:eabm6359
- 123 5. Burbage M, Rocañín-Arjó A, Baudon B, Arribas YA, Merlotti A, Rookhuizen DC, et al. Epigenetically  
124 controlled tumor antigens derived from splice junctions between exons and transposable elements. *Sci*  
125 *Immunol* 2023;8:eabm6360

126



128

129

130 **Figure 1. Transposable elements can be harnessed for anticancer immunotherapy through different**

131 **pathways, which encompass both acquired and innate immunity. (a)** A recent study has achieved a pan-cancer

132 screen for tumor-specific antigens (TSAs) derived from transposable element (TE, light brown) chimeric

133 transcripts (3). TE-derived transcripts (including TE chimeric transcripts and unit-length TE transcripts) can code

134 for: (b) tumor-specific antigens (TSAs) either bound by HLAs and recognized by cytotoxic T-cells, or (c-d) directly

135 exposed at the cell surface as membrane proteins. The latter can be target of CAR-T cells (c) or marked by

136 antibodies for subsequent elimination by the immune system **(d)**. In addition, TE-derived nucleic acids can  
137 induce a viral mimicry state. TE-derived single-stranded DNA (ssDNA) molecules are produced by endogenous  
138 reverse transcriptases (RT), presumably encoded by LINE-1 retrotransposons **(e)**. Such ssDNA species can  
139 accumulate in the cytosol and activate the cGAS-STING pathway (not shown). Moreover, double-stranded RNA  
140 (dsRNA) species, predominantly formed by inverted TEs (i.e. Alu retrotransposons) inserted in intronic regions  
141 can activate RNA sensors such as MDA5 or RIG-I **(f)**. Epigenetic therapies could reverse TE silencing, thereby  
142 promoting TE expression and TE-mediated immunotherapy efficacy **(g)** (4,5). In the near future, TE-derived TSAs  
143 could serve for the development of specific antitumoral immunotherapies such as pan-cancer vaccines, new  
144 CAR-T cells, or monoclonal antibodies **(h)**.