frontiers in Immunology

Check for updates

Systems Biology to Understand and Regulate Human Retroviral Proinflammatory Response

Mohamed Helmy[1,2] and Kumar Selvarajoo[1,3,4*]

[1] Bioinformatics Institute (BII), Agency for Science, Technology and Research (ASTAR), Singapore, Singapore, [2] Department of Computer Science, Lakehead University, Thunder Bay, ON, Canada, [3] Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (ASTAR), Singapore, Singapore, [4] Synthetic Biology Translational Research Program & SynCTI, Yong Loo Lin School of Medicine, National University of Singapore (NUS), Kent Ridge, Singapore

The majority of human genome are non-coding genes. Recent research have revealed that about half of these genome sequences make up of transposable elements (TEs). A branch of these belong to the endogenous retroviruses (ERVs), which are germline viral infection that occurred over millions of years ago. They are generally harmless as evolutionary mutations have made them unable to produce viral agents and are mostly epigenetically silenced. Nevertheless, ERVs are able to express by still unknown mechanisms and recent evidences have shown links between ERVs and major proinflammatory diseases and cancers. The major challenge is to elucidate a detailed mechanistic understanding between them, so that novel therapeutic approaches can be explored. Here, we provide a brief overview of TEs, human ERVs and their links to microbiome, innate immune response, proinflammatory diseases and cancer. Finally, we recommend the employment of systems biology approaches for future HERV research.

Keywords: human retroviral, systems biology, proinflammatory response, computational modeling, cancer

INTRODUCTION

After the human genome sequencing project, it is now well-known that only a very minor component of the whole genome, that is, only about 1-2% constitute of protein-coding genes (1). The remaining sequences make up the numerous transcriptional and translational regulatory components, such as ribosomal DNA genes, transfer RNA genes, and non-coding DNA sequences. Notably, in mammalian cells, about half of the non-coding DNA sequences are transposable elements (TEs) (2).

TEs, also referred as transposons, were originally discovered in the 1940s by Barbara McClintock while studying the maize genetics. It was shown that within the genome, TEs were able to transiently

move across their location through chromosomal breakage mechanisms (3, 4). Although there were some interests in TEs after their discovery, only during the early 2000s that the topic obtained notable recognition through ENCODE, FANTOM and the roadmap epigenomics projects (5–7). These initiatives helped to identify functional elements of the human genome, especially the non-coding regions.

The dynamic property of large-scale TEs found in eukaryotic genome has possibly provided them the advantage to evolve over environmental changes, disease perturbation or evolutionary pressure. However, at the same time, make them prone to viral infection and disease evolution (8, 9). Thus, research into finding the mechanisms and the role of individual TE is an intense current research domain. However, despite the immense efforts to study and characterize these non-coding elements, the bulk of them still remain poorly understood (10).

TEs, whose size range anywhere from a few hundred to several thousand pairs of nucleotides, are currently classified into two major categories: i) DNA transposons (or TE Class II) and ii) retrotransposons (or TE Class I) (**Figure 1A**) (12).

DNA TRANSPOSONS

DNA transposons are transposase genes that are flanked between two Terminal Inverted Repeat (TIR). They are excised by transposase enzymes and move through a DNA intermediate which recognizes TIRs and become inserted into a new genomic
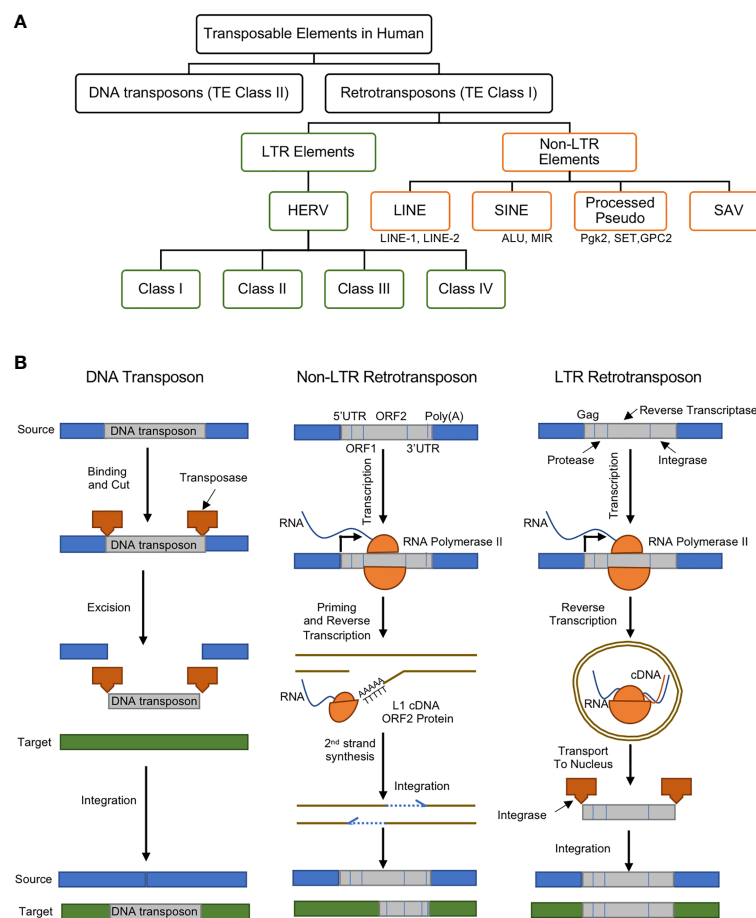
FIGURE 1 | (A) Classes and types of transposable elements (TEs) in the human genome, (B) gene regulatory mechanisms of DNA transposon (left), non-LTR retrotransposon (middle), and LTR retrotransposon (right). DNA transposons are mobilized through a cut and paste machinery that is based on a transposase that is encoded at their flanked sequences. The transposase cuts the DNA transposon from the source and integrates it into a new genomic location. Retrotransposons require an intermediate step of the reverse transcription of an RNA to be mobilized. Non-LTR retrotransposons are transcribed to produce a full-length mRNA, but its transposition uses target-site primed reverse transcription (TPRT). This produces a single strand of cDNA, which is used to create the complementary strand of the ds-DNA that is integrated in the genomic DNA. This process can lead to target-site duplications (TSDs) and small deletions in the target. Non-LTR retrotransposons encode one or two open reading frames (ORFs). The retrotransposition of the LTR-retrotransposons require the encoding of the Gag protein, protease, reverse transcriptase, and integrase enzymes. The RNA polymerase II of the host recognizes the promoter of the LTR-retrotransposons and creates its mRNA. The Gag makes a virus-like particle that contains the mRNA, the reverse transcriptase and the integrase then uses them to create the full-length double-stranded DNA, which is later integrated into the host's DNA using the integrase enzyme. Panel (B) was modified from (11).

location. Upon insertion and integration, the target site DNA is duplicated (Target Site Duplications), which becomes a unique signature for each DNA transposon (**Figure 1B**, left). The major focus of this manuscript is LTR retrotransposons, particularly the endogenous retroviruses, as will be discussed in better details below.

RETROTRANSPOSONS

Retrotransposons comprise of long terminal repeats (LTRs) and non-long terminal repeats (non-LTRs) sequences (**Figure 1B**, middle and right). Retrotransposons are elements that are reverse transcribed into a cDNA and integrated back into a different region of the genome. DNA transposons, on the other hand, are able to insert themselves into different locations with the aid of a circular DNA intermediate (**Figure 1B**) (11). Thus, retrotransposons are considered replicative, while DNA transposons are termed non-replicative. Both types of TEs can be further divided into subclasses, families or superfamilies that share a common genetic organization or a recent ancestral origin (12). Nevertheless, the classification of TEs may most likely evolve as new data or knowledge is generated, and although both retrotransposons and DNA transposons play important gene regulatory roles, there are mounting recent reports showing their associations to about 100 diseases (13).
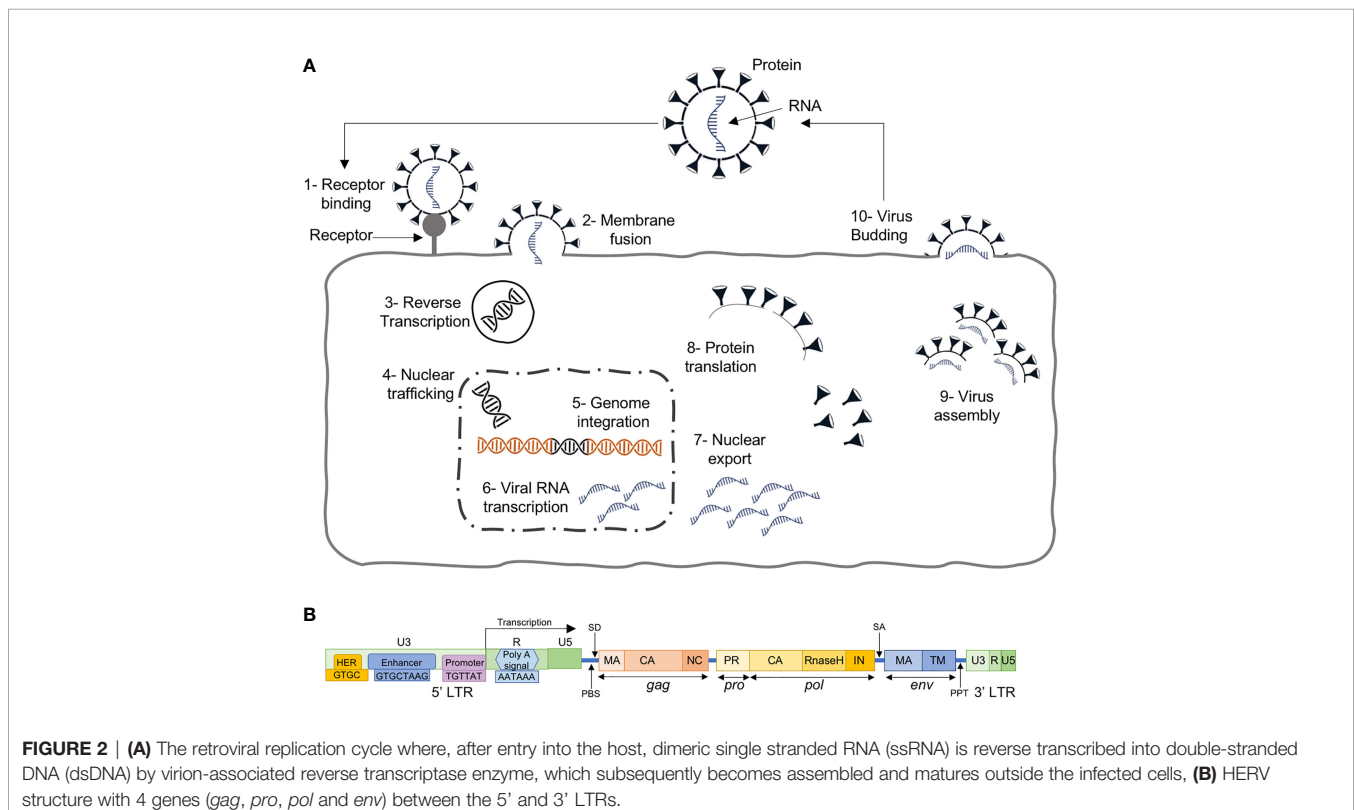
Non-LTR Retrotransposons

Non-LTR retrotransposons (also known as polyA retrotransposons or target-primed (TP) retrotransposons), comprises of long and short interspersed nuclear elements (LINEs and SINEs) and they are more like an integrated mRNA (**Figure 1B**, middle) (14). Non-LTR retrotransposons are one of the most abundant retrotransposons in all eukaryotic genomes (15). The estimations show that about 30% of the human genome is directly or indirectly driven by the Non-LTR retrotransposons can due to their excessive copy number (13, 16). LINEs consist of an internal promoter for RNA polymerase II, a 5′ untranslated region (UTR), two open reading frames (ORFs), and a 3′ terminal polyadenylation site (17). SINEs possess an RNA polymerase III internal promoter and a 3′ A-rich tract (17). SINEs partner with LINEs for reverse transcriptase and endonuclease functions. SINE and LINE mutations and polymorphisms have been linked to numerous human diseases (9). For example, increased LINE-1 expression led to the induction of type I interferon in systemic autoimmune disease patients (18).

LTR Retrotransposons: Retroviruses

The LTR retrotransposons, in eukaryotic genomes, are pairs of identical sequences of DNA that are found on either end of a series of genes, pseudogenes, or endogenous viral elements (**Figure 1B**, right). They are known to comprise roughly a tenth of the human genome (19). In jawed vertebrates and human, the LTR retrotransposons are largely endogenous retroviruses (ERVs).

The ERVs are known to be present in vertebrates for more than 400 million years, and are inherited viral elements from ancient retroviruses that infected germ cells or their progenitors (19, 20). They are able to insert a copy of their RNA genome into the DNA of its host cells' genome to become a provirus (20). This is done through a few regulatory steps (**Figure 2A**): i) the virus

FIGURE 2 | **(A)** The retroviral replication cycle where, after entry into the host, dimeric single stranded RNA (ssRNA) is reverse transcribed into double-stranded DNA (dsDNA) by virion-associated reverse transcriptase enzyme, which subsequently becomes assembled and matures outside the infected cells, **(B)** HERV structure with 4 genes (*gag*, *pro*, *pol* and *env*) between the 5' and 3' LTRs.

uses "retro" technique in the host cell's cytoplasm to produce DNA from its RNA genome using its own reverse transcriptase enzyme, ii) transported to the nucleus, the resultant DNA, through an integrase enzyme, is inserted into the host cell's genome at random, iii) the modified host cell genome transcribes the viral genes, together with its own genes, which eventually leads to multiplication of the original virus.

In their host, these have been inactivated by the accumulation of coding sequence mutations and hence, have reached a so-called genetic fixation through various transcriptional/translational repression and histone hypermethylation. Since the insertions are random, and their functions were initially unknown, these sequences were referred to as junk DNA (21).

HERVs, the ERVs in human, are similar to exogenous retroviruses, consisting of four genes (*gag*, *pro*, *pol* and *env*) between the LTRs (**Figure 2B**). The LTRs serve as promoters with typical RNA regulatory sequences and transcription factor binding sites for HERV protein expression. HERVs are permanently integrated genome elements that are mostly located in the heterochromatin and have undergone methylation silencing (22). That is, after million years of evolution and multiple cycles of integration and reintegration, HERVs have acquired many mutations silencing their activity and remain harmless to the host.

More recently, several HERVs' roles have been identified, especially for developmental and immune responses in the host (23, 24). The high frequency of HERVs within LTRs indicates that viral adaptations can take advantage of immune signaling pathways that promote viral transcription and replication. They are, thus, still able to regulate and shape the host immune system and may indicate a crucial ancestral role in the evolution and functioning of the mammalian or human immune system. Notably, recent works have shown links between increased HERV protein expressions with human diseases through the proinflammatory pathways, notably in the triggering of interferon responses indicating an involvement in multiple human diseases (25).

HERVS IN HUMAN DISEASES

The ability of HERVs to stably integrate into the genome and to modify the expression of nearby genes drew attention for the investigation of their role in different human diseases (26). It was shown that HERVs are highly expressed in neoplastic cells, which indicated a relationship between them and the neoplastic transformation. To date, the relationship between HERVs and several types of cancer is established including soft tissue sarcoma, chronic myeloid leukemia, breast cancer, hepatocellular carcinoma, prostate carcinoma, ovarian carcinoma, and pancreatic cancer (23, 27). HERVs were also linked to neurological disorders (28), mental illness (29), neuropsychological diseases (30), neurodevelopmental disorders and neurodegenerative diseases (31, 32). In addition, they are also found to be involved in Type 1 diabetes (33), autoimmune disorders and multiple sclerosis (34, 35). Here, we will briefly review the relationships between HERVs and

microbiome, their involvement in immunity, proinflammatory diseases and cancer.

HERVs and Microbiome

Commensal microbiota, communities of thousands of microorganisms, are known to play key roles in modulating host immune systems and their homeostasis maintenance (36). Nevertheless, the cellular processes by which microbiota benefits the host cells in still largely unknown.

One recent work studied the effect of skin microbiota on ERVs. Keratinocytes promoted the induction of commensal-specific T cells, attributed by the discrete expression of defined ERVs, which activated cyclic GMP-AMP synthase (cGAS), stimulator of interferon genes protein (STING) signaling. Subsequently, ERV reverse transcription inhibition resulted in aberrated immunity to the microbiota and its associated tissue repair function (37).

Large-scale ERVs expression was investigated by high-throughput microarray in murine and human immune cells (38). Wildtype and MYD88 knock out (KO) mice revealed that certain ERVs expression were dependent on gut microbiota. Notably, ERVs expression were largely decreased in MYD88 KO mice, indicating that microbiota is necessary to induce ERVs expression through MYD88 signaling in mice's gut.

In human, one study demonstrated a strong correlation between the expression of HERV-H, -K and -W, and the concentration of *Bifidobacterium* spp., indicating that the colonization with commensal microbes made up of *Bifidobacterium* spp. causes global modulation of HERVs in the gut (39).
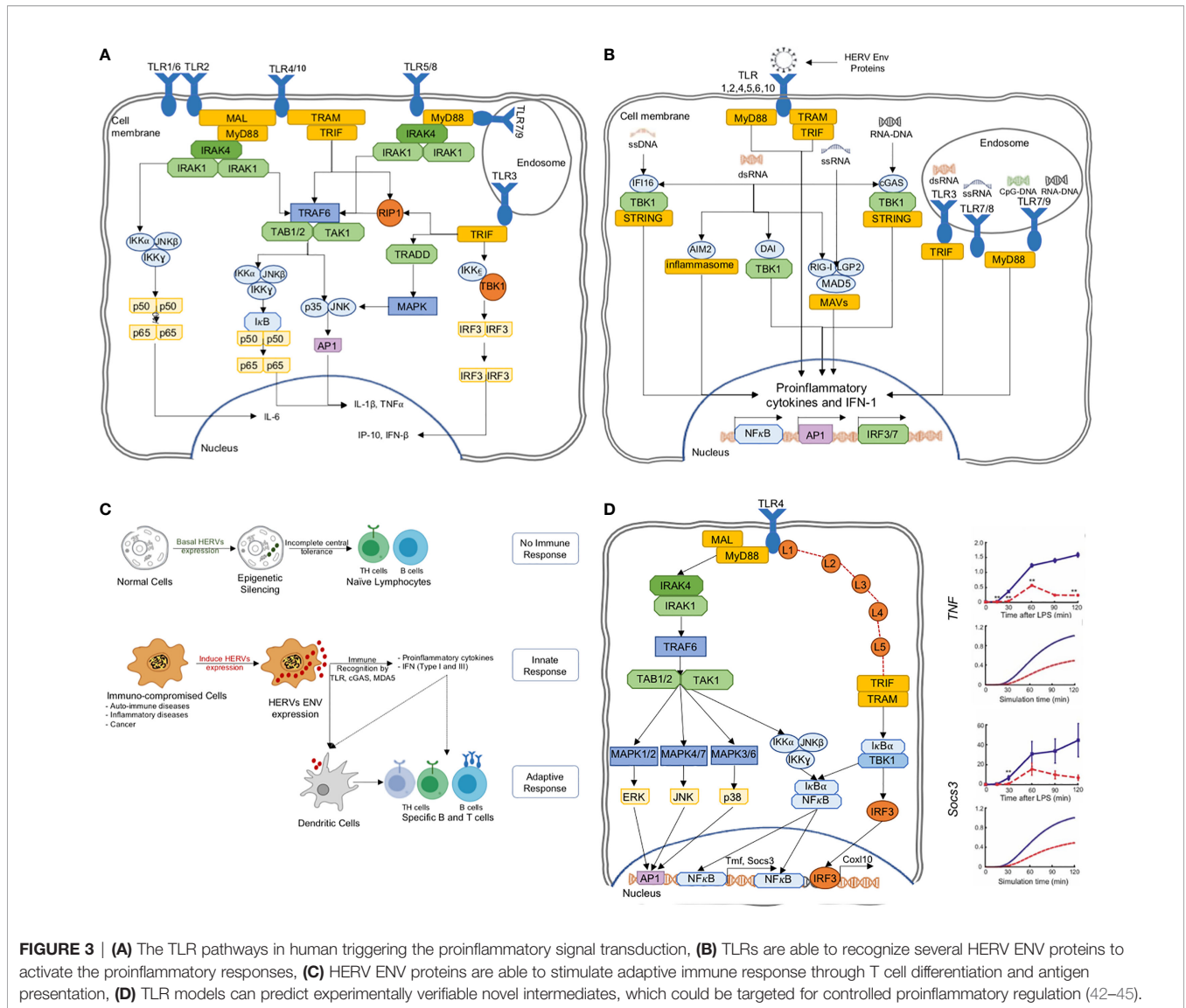
Thus, these works provide evidence that ERVs can regulate immune response through microbiome, and, collectively, they may be better controlled for health benefits.

HERVs in Innate Immunity

HERVs have become an integral part of host immunity and are able to protect the host from exogenous retroviral infections. HERV components can be detected by several innate immunity sensors, such as the pattern recognition receptors (PRRs), which play pivotal roles in antiviral protections.

A major class of PRRs are the Toll-like receptors (TLRs). There are ten types found in humans, and they play the first line of defense in detecting a wide variety of pathogenic interventions (40, 41) (**Figure 3A**). In particular, TLR3,7,8 and 9 are endosomal and together they sense dsRNA, ssRNA and CpG DNA. At plasma membrane, the TLR2/6/10 are known to detect HIV viral proteins such as GAG (p17, p24) and ENV (gp41, gp120). On the other hand, TLR4 senses ENV, HIV-TAT and other virion-associated bacterial components such as lipopolysaccharide (LPS) (46). These diverse TLR engagements with viral proteins or components lead to the activation of MAP kinases and NF-κB and the resultant proinflammatory cytokine production for neutralization of the infections (23) (**Figure 3B**). Moreover, ENV stimulates dendritic cells to promote T helper cell differentiation (**Figure 3C**).

The HERV-W Env was shown to interact with TLR4 to induce cytokines IL-1, IL-6 and TNF-α (47). In another example, HERV-K(HML-2) Env activated TLR7/8 in neurons and microglia, and strongly correlates with the onset and causing neurodegeneration *via* proinflammatory response (28, 48).

FIGURE 3 | (A) The TLR pathways in human triggering the proinflammatory signal transduction, **(B)** TLRs are able to recognize several HERV ENV proteins to activate the proinflammatory responses, **(C)** HERV ENV proteins are able to stimulate adaptive immune response through T cell differentiation and antigen presentation, **(D)** TLR models can predict experimentally verifiable novel intermediates, which could be targeted for controlled proinflammatory regulation (42–45).

Similarly, HERV-H was shown to be associated with multiple sclerosis (49). HERV-derived peptides, involving ENV conserved subunit ISD, have been implicated in immune-suppressive mechanisms (35). On the other hand, during pregnancy, the HERV protein, Syncytin-2, immunosuppresses T cells in an exosome-mediated manner (50). Thus, there are mounting evidence pointing to the importance of HERVs in triggering the human innate immune response. While some of the HERV sequences activate the immune system, others suppress it. Thus, a deeper knowledge of TLR and HERV interactions are necessary for the positive regulation of proinflammatory response for therapeutic agents.

In a transcriptome-wide RNA-Seq analysis of murine macrophages, full-length ERV-derived lncRNA (lnc-EPAV) expression was rapidly upregulated by viral RNA mimics or RNA viruses (2). Silencing lnc-EPAV showed that it facilitated the expression of RELA, an NF-κB subunit that plays a pivotal role in innate immunity (51). lnc-EPAV-deficiency also reduced

the expression of type I interferons (IFNs). Following lethal RNA virus infection, this resulted in increased viral loads and mortality.

Another transcriptome-wide analysis of B cells revealed a small number of highly distinct ERVs that are strongly and consistently induced during murine B cell activation (52). Notably, a single endogenous MLV provirus, Xmv45, was found to be expressed at significantly higher levels.

Another research area for exploiting HERV is to regulate the tumor suppressor p53 in cancer. By DNA damage and cellular stress, the p53 pathway leads to cellular apoptosis. It was previously found that 30% of binding sites were located within copies of a few HERV families (23, 53). This benefits retroviruses as p53 pathway leads to rapid induction of transcriptional processes that neutralizes viral RNA from the host cell.

Having mentioned that HERV can be positively regulated, it must also be noted that the random insertion mechanisms of retroviruses could unintentionally activate oncogenes or

deactivate tumor suppressing genes (54). In other words, HERVs have the potential to transform normal cells into cancerous cells. Moreover, proviruses are known to remain latent for extended period of time within the host and could likely become activated by key changes to host cell environment.

HERVs in Cancer

Back in the early 70s, there has been studies on reverse transcriptase activity and oncoviruses in human cancers (55, 56). Subsequent works eventually led to the link between HERVs activations and myriad cancers (57). Elevated expression of the oncogene ETV1 through HERV-K_22q11.23 5′-LTR-UTR sequence has been shown in prostate cancers (58). In melanoma, the activation of MEK- ERK and p16INK4A-CDK4 pathways increased the expression of HERV-K (59). There are numerous other links that show the connection or causality between cancer and elevated HERV expressions (60). Such mounting evidences have led to the development of immunotherapeutic strategies that targets HERVs (61, 62). Nevertheless, these methods may not be effective as the detailed mechanistic understanding, such as how the upstream signaling process that activates the downstream distinct HERVs in diverse cancers are still largely unknown. For example, a dynamic computational model describing the detailed signaling reactions, from TRAIL ligand-receptor binding, was used to identify a crucial intracellular target, protein kinase C, whose downregulation sensitized cancer cell death from 60% to 95% *via* increased apoptosis signaling (63). Detailed mechanistic understanding such as tracking intracellular signaling dynamics, thus, can better help to elucidate crucial specific targets that selectively control specific set of downstream proinflammatory or apoptosis responses.

It is noteworthy to mention that TLRs' expressions are also elevated in many cancers (64, 65). Hence, it is conceivable that HERVs can be expressed *via* TLR signaling, especially that there are evidences for increased NF-κB, MAPK and IRF1 activity with increased HERV expressions (59, 66, 67). Thus, further research into TLR-induced HERV expressions will provide better clues into finding the missing mechanistic links.

HERVs in Neuro and Autoimmune Diseases

HERVs (also LINEs and SINEs) have been identified in numerous neurodegenerative diseases, including schizophrenia, Alzheimer's disease, stroke, multiple sclerosis (MS), amyotrophic lateral sclerosis, accelerated neurological decline in aging, and neuropathogenesis of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) (9, 68)

In MS, HERV-W env binds with TLR4 on oligodendroglial precursor cells and induces pro-inflammatory response leading to the suppression of myelin expression in MS lesions (69). HERV-W env is also known to be a superantigen linked with demyelination in MS and, consequently, HERV-W env antibody have been shown to effectively rescue myelin expression (47, 70).

High-throughput RNA-Seq was performed to investigate the expression and role of HERVs in systemic lupus erythematosus (SLE) (71). Between control and test case datasets, 481 HERV

encoding regions were identified, and significant variations between differentially expressed HERVs and non-differentially expressed HERVs were determined. Notably, HERV-K, HERV-H, and MER4 were over-expressed in SLE patients. Subsequently, utilizing locus specific resolution of HERV mapping, key innate immune pathways regulated by the differential HERV expression in SLE were identified.

So far, we have introduced our basic understanding of TEs, HERVs, the innate immune signaling through TLRs, and the causal relationship between HERVs and proinflammatory diseases and cancers. Although we are gaining a better understanding on the retrovirus effects on host proinflammatory response and their role in cancer, there is no simple way to generalize the behaviors under different conditions and cell types. As noted, in some cases, the HERVs are able to activate immunity while in others it de-activates. To tackle the conundrum, we can explore computational models that are able to aid in the distinction of differing proinflammatory kinetics.

SYSTEMS BIOLOGY AND MACHINE LEARNING

Computational modelling is becoming an integral part of biological research, especially when investigating cellular network responses. There are a wide range of computational modeling techniques and algorithms today that provide novel mechanistic insights (72). The methods can be broadly classified into i) parametric approaches such as dynamic modeling using ordinary differential equations (42, 63, 73–75), and ii) non-parametric models using Boolean logics, stoichiometric matrix and Bayesian inference algorithms (76, 77).

Dynamic models are mostly built using differential equations to represent each step in a cellular network, such as signaling or metabolic reaction. Using the reaction topology and kinetics, the models predict dynamic outcomes to different *in silico* perturbations, or to understand the key regulatory mechanisms, such as bottlenecks, and flux distributions (78, 79). In other words, the dynamic models utilize *a priori* knowledge of cellular networks and temporal experimental data to simulate the molecular activation or concentration profiles over time.

Although dynamic or kinetic models have been widely used and have proven their benefits, stoichiometric constraint-based modeling and Boolean logic-based approaches are also popular. These models are used when reactions kinetics are not feasible to obtain and will largely provide qualitative information on the outcome of any perturbation that need to be modeled or tested (80).

In a previous study, using a perturbation-response kinetic modeling approach, we simulated the differential dynamic activation kinetics of NF-κB and MAP kinases in MyD88 and TRIF mutant cells to TLR4 activation (43). Key predictions of the model include the enhancement of alternative pathways at pathway junctions through signaling flux redistribution, and the presence of novel intermediates on TRIF-dependent pathways, which were subsequently verified experimentally and

characterized (42, 44, 45) (**Figure 3D**). Similar successes were also noted for TLR3, TNF and TRAIL signaling, where dynamic models were able to elucidate novel mechanistic features of signaling processes (63, 73–75).

In another study, Hoffmann and colleagues investigated the cell type differential activation mechanisms of IKK–IκB–NF-κB signaling module to TNF and TLR4 stimulation in murine embryonic fibroblasts (81). Their ordinary differential equations model predicted the rapid termination of IKK activity in TNF stimulation is by the negative feedback control through post induction of A20, whereas the prolonged IKK activation in LPS stimulation is caused by positive autocrine feedback of TNF signaling.

For cancer, mathematical modeling of signaling pathways was also used to investigate different aspects of tumor growth and progression. Hendrata and Sudiono presented a model based on partial differential equations to study apoptosis in HeLa cancer cells (82). The model simulations were fitted with experimentally profiles observed tumor growth rate and phenotypic patterns, and this was subsequently used to predict the growth of tumours over time for various mesenchymal stem secretions. There are also numerous other cancer and proinflammatory signaling models that have provided indispensable data for guiding experimental elucidation of key regulatory properties (83–86). Thus, computational and systems biology research could aid in elucidating the regulatory properties of diverse HERVs provided time-series experiments are performed in wildtype and mutant experiments. However, traditional computational modeling approaches, both parametric and non-parametric, have limitations in their capabilities, scale or required data (87).

Machine learning approaches were used in combination with systems biology to make prediction models where traditional modeling approaches failed due to their limitations or the unavailability of sufficient data. This was demonstrated in different fields including cancer, metabolic engineering and proinflammatory disease (87–89). In metabolic engineering, systems biology approaches combined with machine learning models were able to overcome several limitations in data quality and modeling scale (87). Machine learning models in proinflammatory disease helped to establish a relationship between the peptides amino acid sequence (the epitope of the antigen) and the induced proinflammatory response, an approach that can be adopted in the HERV research to investigate the immune response to HERVs (88). Machine learning also solved the challenge of predicting cancer survival and its associated pathways from gene expression data and predicted the survival and association in four types of cancer (89).

Unlike current signaling or metabolic network modeling developed and used, HERV modeling is faced with limited or sparse information that make it a big challenge to even start constructing a workable model. Different computational models require different types of data and levels of quality that help those models perform to the best of their capabilities. For instance, dynamic models require detailed information on the signaling pathway, its reactions, the rate of the reactions, gene expressions and multiple time points (87). Furthermore, other types of data

(such as genomic and protein sequences and 3-D structure data) are important for predicting molecular interactions and signaling pathways associations (90). Such data is not readily available for HERVs as they are understudied (10). To overcome this limitation, systems biology coupled with machine or deep learning models provide hope for further understanding the roles of HERVs on immune response or cancer progression. This approach could provide valuable clues for predicting causal functions using heuristics methods (91).

For example, most mammalian retroviruses envelope glycoproteins are linked to group-specific antigen (gag) genes, which codes for matrix proteins (MAs) involved in virus component assembly, transport and budding. Notably, HERV gag protein have been shown to correlate with prostate cancer progression (92). Since, the initiation sites and termination sites of gag gene transcription are poorly understood, Ma *et al.* developed a computational method using support vector machine and random forest to identify MAs in HERVs and predict the initiation sites and termination sites. Their model scanned 94,671 HERV sequences from 118 families to predict 104 new putative MAs in human chromosomes (93).

Using a similar approach, Dey et al. investigated the protein-protein interactions between human and SARS-CoV-2 virus. Using different sequence-based features of human proteins, e.g. amino acid composition, pseudo amino acid composition, and conjoint triad, they predicted 1,326 human target proteins of SARS-CoV-2 and compared them with gene ontology and KEGG pathway enrichment (94). In a more recent work, Kojima et al. used machine learning methods, using k-mer occurrence of ancient RNA viral sequences, to identify novel viruses in the human genome (95). Thus, although new in application, machine learning techniques will play a bigger role in elucidating the sequence and regulatory causality between HERVs and major diseases, and will aid in better understanding of the complex interplay between HERVs, human proinflammatory response and diseases like cancer, multiple sclerosis (MS) and AIDS.

CONCLUSION

HERVs are TEs that make up a large portion of the human genome. Although mostly underexplored, many of them are known to be activated during embryogenesis and are epigenetically silenced subsequently (96), while others are known to cause host genome instability leading to major diseases such as inflammation and cancer (23, 25). During evolution, HERV genome sequences have acquired large scale mutations and deletions and, therefore, are unable to produce infectious viral agents.

HERV can be transcriptionally activated and are recently known to regulate proinflammatory responses, nevertheless, the details of the regulation remain largely a mystery. Understanding the role of each HERV in general cellular functions such as growth, cell division, differentiation, and immune responses are

vital to reveal the connection of HERVs in human diseases. For example, with the ability to regulate innate response through TLRs and the resultant adaptive immune response through antigens recognition, they could be explored as future agents for cancer immunotherapy (62, 97). Furthermore, HERV expressions are mostly silenced by DNA methylation and histone modification. Epigenetic regulatory drugs have, thus, been indicated to reactivate HERVs for immune related clearance (22, 98).

Overall, recent works on HERVs suggest that they could be positively tuned to fight major human diseases. In the near future, we will likely witness the extension of systems biology, through computational modeling of regulatory pathways, and machine learning techniques, through heuristic modeling and data analytics, to help support in overcoming the difficulty of putting a wealth of HERV information together into a hypothesis driven solution to tackle and treat complex diseases.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MH wrote the manuscript and prepared the figures. KS wrote the manuscript and led the project. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

1. Sana J, Faltejskova P, Svoboda M, Slaby O. Novel Classes of Non-Coding Rnas and Cancer. *J Transl Med 2012 101* (2012) 10:1–21. doi: 10.1186/1479-5876-10-103

2. Makałowski W, Gotea V, Pande A, Makałowska I. . "Transposable Elements: Classification, Identification, and Their Use as a Tool for Comparative Genomics,". In: *Methods in Molecular Biology*. New York, NY: Humana Press Inc. (2019) p. 177–207. doi: 10.1007/978-1-4939-9074-0_6

3. McClintock B. The Origin and Behavior of Mutable Loci in Maize. *Proc Natl Acad Sci USA* (1950) 36:344–55. doi: 10.1073/pnas.36.6.344

4. Ravindran S. Barbara Mcclintock and the Discovery of Jumping Genes. *Proc Natl Acad Sci U.S.A.* (2012) 109:20198–9. doi: 10.1073/pnas.1219372109

5. Bono H, Kasukawa T, Furuno M, Hayashizaki Y, Okazaki Y, fantom DB. Database of Functional Annotation of RIKEN Mouse Cdna Clones. *Nucleic Acids Res* (2002) 30:116–8. doi: 10.1093/nar/30.1.116

6. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative Analysis of 111 Reference Human Epigenomes. *Nature* (2015) 518:317–29. doi: 10.1038/nature14248

7. Feingold EA, Good PJ, Guyer MS, Kamholz S, Liefer L, Wetterstrand K, et al. The ENCODE (Encyclopedia of DNA Elements) Project. *Science (80)* (2004) 306:636–40. doi: 10.1126/science.1105136

8. Broecker F, Moelling K. Evolution of Immune Systems From Viruses and Transposable Elements. *Front Microbiol* (2019) 10:51. doi: 10.3389/FMICB.2019.00051

9. Payer LM, Burns KH. Transposable Elements in Human Genetic Disease. *Nat Rev Genet* (2019) 20:760–72. doi: 10.1038/S41576-019-0165-8

10. Farh KKH, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and Epigenetic Fine Mapping of Causal Autoimmune Disease Variants. *Nature* (2015) 518:337–43. doi: 10.1038/nature13835

11. Levin HL, Moran JV. Dynamic Interactions Between Transposable Elements and Their Hosts. *Nat Rev Genet 2011 129* (2011) 12:615–27. doi: 10.1038/nrg3030

12. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten Things You Should Know About Transposable Elements 06 Biological Sciences 0604 Genetics. *Genome Biol* (2018) 19:199. doi: 10.1186/s13059-018-1577-z

13. Hancks DC, Kazazian HH Jr. Active Human Retrotransposons: Variation and Disease. *Curr Opin Genet Dev* (2012) 22:191–203. doi: 10.1016/J.GDE.2012.02.006

14. Han JS. Non-Long Terminal Repeat (Non-LTR) Retrotransposons: Mechanisms, Recent Developments, and Unanswered Questions. *Mob DNA 2010 11* (2010) 1:1–12. doi: 10.1186/1759-8753-1-15

15. Eickbush TH, Malik HS. Origins and Evolution of Retrotransposons. *Mob DNA II* Washington, DC: ASM Press. (2014). pp. 1111–44. doi: 10.1128/9781555817954.CH49

16. Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as Regulators of Gene Expression. *Science* (2016) 351:aac7247. doi: 10.1126/SCIENCE.AAC7247

17. Platt RN 2nd, Vandewege MW, Ray DA. Mammalian Transposable Elements and Their Impacts on Genome Evolution. *Chromosome Res* (2018) 26:25–43. doi: 10.1007/S10577-017-9570-Z

18. Mavragani CP, Sagalovskiy I, Guo Q, Nezos A, Kapsogeorgou EK, Lu P, et al. Expression of Long Interspersed Nuclear Element 1 Retroelements and Induction of Type I Interferon in Patients With Systemic Autoimmune Disease. *Arthritis Rheumatol (Hoboken NJ)* (2016) 68:2686–96. doi: 10.1002/ART.39795

19. Cordaux R, Batzer MA. The Impact of Retrotransposons on Human Genome Evolution. *Nat Rev Genet* (2009) 10:691–703. doi: 10.1038/nrg2640

20. Johnson WE. Origins and Evolutionary Consequences of Ancient Endogenous Retroviruses. *Nat Rev Microbiol* (2019) 17:355–70. doi: 10.1038/s41579-019-0189-2

21. Palazzo AF, Gregory TR. The Case for Junk DNA. *PloS Genet* (2014) 10: e1004351. doi: 10.1371/journal.pgen.1004351

22. Hurst TP, Magiorkinis G. Epigenetic Control of Human Endogenous Retrovirus Expression: Focus on Regulation of Long-Terminal Repeats (Ltrs). *Viruses* (2017) 9(6):130. doi: 10.3390/v9060130

23. Grandi N, Tramontano E. Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses. *Front Immunol* (2018) 9:2039. doi: 10.3389/fimmu.2018.02039

24. Rowe HM, Trono D. Dynamic Control of Endogenous Retroviruses During Development. *Virology* (2011) 411:273–87. doi: 10.1016/j.virol.2010.12.007

25. Alcazer V, Bonaventura P, Depil S. Human Endogenous Retroviruses (Hervs): Shaping the Innate Immune Response in Cancers. *Cancers (Basel)* (2020) 12 (3):610. doi: 10.3390/cancers12030610

26. Le Dantec C, Vallet S, Brooks WH, Renaudineau Y. Human Endogenous Retrovirus Group E and Its Involvement in Diseases. *Viruses* (2015) 7:1238–57. doi: 10.3390/V7031238

27. Giebler M, Staege MS, Blauschmidt S, Ohm LI, Kraus M, Würl P, et al. Elevated HERV-K Expression in Soft Tissue Sarcoma Is Associated With Worsened Relapse-Free Survival. *Front Microbiol* (2018) 0:211. doi: 10.3389/FMICB.2018.00211

28. Küry P, Nath A, Créange A, Dolei A, Marche P, Gold J, et al. Human Endogenous Retroviruses in Neurological Diseases. *Trends Mol Med* (2018) 24:379–94. doi: 10.1016/j.molmed.2018.02.007

29. Canli T. A Model of Human Endogenous Retrovirus (HERV) Activation in Mental Health and Illness. *Med Hypotheses* (2019) 133:109404. doi: 10.1016/J.MEHY.2019.109404

30. Wang DS, Zhong B, Zhang MS, Gao Y. Upregulation of Serum Mir-103 Predicts Unfavorable Prognosis in Patients With Colorectal Cancer. *Eur Rev Med Pharmacol Sci* (2018) 22:4518–23. doi: 10.26355/eurrev_201807_15506

31. Joel G, Kremer D, Küry P. Human Endogenous Retroviruses: Ammunition for Myeloid Cells in Neurodegenerative Diseases? *Neural Regener Res* (2020) 15:1043–4. doi: 10.4103/1673-5374.270311

32. Emanuela B, Matteucci C, Cipriani C, Grelli S, Ricceri L, Calamandrei G, et al. Endogenous Retroviruses Activity as a Molecular Signature of Neurodevelopmental Disorders. *Int J Mol Sci* (2019) 20:6050. doi: 10.3390/IJMS20236050

33. Levet S, Charvet B, Bertin A, Deschaumes A, Perron H, Hober D. Human Endogenous Retroviruses and Type 1 Diabetes. *Curr Diabetes Rep* (2019) 19. doi: 10.1007/S11892-019-1256-9

34. Tarlinton RE, Ekaterina M, Rizvanov AA, Svetlana K, Subhash V. Role of Viruses in the Pathogenesis of Multiple Sclerosis. *Viruses* (2020) 12. doi: 10.3390/V12060643

35. Gröger V, Cynis H. Human Endogenous Retroviruses and Their Putative Role in the Development of Autoimmune Disorders Such as Multiple Sclerosis. *Front Microbiol* (2018) 9:265. doi: 10.3389/fmicb.2018.00265

36. Leshem A, Liwinski T, Elinav E. Immune-Microbiota Interplay and Colonization Resistance in Infection. *Mol Cell* (2020) 78:597–613. doi: 10.1016/J.MOLCEL.2020.03.001

37. Lima-Junior DS, Krishnamurthy SR, Bouladoux N, Collins N, Han SJ, Chen EY, et al. Endogenous Retroviruses Promote Homeostatic and Inflammatory Responses to the Microbiota. *Cell* (2021) 184:3794–3811.e19. doi: 10.1016/J.CELL.2021.05.020

38. Young GR, Mavrommatis B, Kassiotis G. Microarray Analysis Reveals Global Modulation of Endogenous Retroelement Transcription by Microbes. *Retrovirology* (2014) 11:59. doi: 10.1186/1742-4690-11-59

39. Bergallo M, Galliano I, Montanari P, Zaniol E, Graziano E, Calvi C, et al. Modulation of Human Endogenous Retroviruses -H, -W and -K Transcription by Microbes. *Microbes Infect* (2020) 22:366–70. doi: 10.1016/J.MICINF.2020.01.006

40. Krishnan J, Selvarajoo K, Tsuchiya M, Lee G, Choi S. Toll-Like Receptor Signal Transduction. *Exp Mol Med* (2007) 39:421–38. doi: 10.1038/emm.2007.47

41. O'Neill LAJ, Golenbock D, Bowie AG. The History of Toll-Like Receptors-Redefining Innate Immunity. *Nat Rev Immunol* (2013) 13:453–60. doi: 10.1038/nri3446

42. Selvarajoo K. *Immuno Systems Biology*. New York, NY: Springer New York (2013). doi: 10.1007/978-1-4614-7690-0

43. Selvarajoo K. Discovering Differential Activation Machinery of the Toll-Like Receptor 4 Signaling Pathways in Myd88 Knockouts. *FEBS Lett* (2006) 580:1457–64. doi: 10.1016/j.febslet.2006.01.046

44. Selvarajoo K, Takada Y, Gohda J, Helmy M, Akira S, Tomita M, et al. Signaling Flux Redistribution at Toll-Like Receptor Pathway Junctions. *PloS One* (2008) 3:e3430. doi: 10.1371/journal.pone.0003430

45. Kagan JC, Su T, Horng T, Chow A, Akira S, Medzhitov R. TRAM Couples Endocytosis of Toll-Like Receptor 4 to the Induction of Interferon-β. *Nat Immunol* (2008) 9:361–8. doi: 10.1038/ni1569

46. Browne EP. The Role of Toll-Like Receptors in Retroviral Infection. *Microorganisms* (2020) 8:1–21. doi: 10.3390/microorganisms8111787

47. Rolland A, Jouvin-Marche E, Viret C, Faure M, Perron H, Marche PN. The Envelope Protein of a Human Endogenous Retrovirus-W Family Activates Innate Immunity Through CD14/TLR4 and Promotes Th1-Like Responses. *J Immunol* (2006) 176:7636–44. doi: 10.4049/jimmunol.176.12.7636

48. Dembny P, Newman AG, Singh M, Hinz M, Szczepek M, Krüger C, et al. Human Endogenous Retrovirus Hervk(Hml-2) Rna Causes Neurodegeneration Through Toll-Like Receptors. *JCI Insight* (2020) 5. doi: 10.1172/JCI.INSIGHT.131093

49. Morandi E, Tanasescu R, Tarlinton RE, Constantinescu CS, Zhang W, Tench C, et al. The Association Between Human Endogenous Retroviruses and Multiple Sclerosis: A Systematic Review and Meta-Analysis. *PloS One* (2017) 12(2):e0172415. doi: 10.1371/journal.pone.0172415

50. Lokossou AG, Toudic C, Nguyen PT, Elisseeff X, Vargas A, Rassart É, et al. Endogenous Retrovirus-Encoded Syncytin-2 Contributes to Exosome-Mediated Immunosuppression of T Cells. *Biol Reprod* (2020) 102:185–98. doi: 10.1093/biolre/ioz124

51. DiDonato JA, Mercurio F, Karin M. NF-κB and the Link Between Inflammation and Cancer. *Immunol Rev* (2012) 246:379–400. doi: 10.1111/J.1600-065X.2012.01099.X

52. Attig J, Young GR, Stoye JP, Kassiotis G. Physiological and Pathological Transcriptional Activation of Endogenous Retroelements Assessed by RNA-Sequencing of B Lymphocytes. *Front Microbiol* (2017) 8:2489. doi: 10.3389/FMICB.2017.02489

53. Levine AJ, Ting DT, Greenbaum BD. P53 and the Defenses Against Genome Instability Caused by Transposons and Repetitive Elements. *BioEssays* (2016) 38:508–13. doi: 10.1002/bies.201600031

54. Cegolon L, Salata C, Weiderpass E, Vineis P, Palù G, Mastrangelo G. Human Endogenous Retroviruses and Cancer Prevention: Evidence and Prospects. *BMC Cancer* (2013) 13(1):1-10. doi: 10.1186/1471-2407-13-4

55. Zhdanov VM, Soloviev VD, Bektemirov TA, Ilyin KV, Bykovsky AF, Mazurenko NP, et al. Isolation of Oncornaviruses From Continuous Human Cell Cultures. *Intervirology* (1973) 1:19–26. doi: 10.1159/000148828

56. Sarngadharan MG, Sarin PS, Reitz MS, Gallo RC. Reverse Transcriptase Activity of Human Acute Leukaemic Cells: Purification of the Enzyme, Response to AMV 70S RNA, and Characterization of the DNA Product. *Nat New Biol* (1972) 240:67–72. doi: 10.1038/newbio240067a0

57. Bannert N, Hofmann H, Block A, Hohn O. Hervs New Role in Cancer: From Accused Perpetrators to Cheerful Protectors. *Front Microbiol* (2018) 9:178. doi: 10.3389/fmicb.2018.00178

58. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, et al. Distinct Classes of Chromosomal Rearrangements Create Oncogenic ETS Gene Fusions in Prostate Cancer. *Nature* (2007) 448:595–9. doi: 10.1038/nature06024

59. Li Z, Sheng T, Wan X, Liu T, Wu H, Dong J. Expression of HERV-K Correlates With Status of MEKERK and P16ink4acdk4 Pathways in Melanoma Cells. *Cancer Invest* (2010) 28:1031–7. doi: 10.3109/07357907.2010.512604

60. Gonzalez-Cao M, Iduma P, Karachaliou N, Santarpia M, Blanco J, Rosell R. Human Endogenous Retroviruses and Cancer. *Cancer Biol Med* (2016) 13:483–8. doi: 10.20892/j.issn.2095-3941.2016.0080

61. Curty G, Marston JL, De Mulder Rougvie M, Leal FE, Nixon DF, Soares MA. Human Endogenous Retrovirus K in Cancer: A Potential Biomarker and Immunotherapeutic Target. *Viruses* (2020) 12(7):726. doi: 10.3390/v12070726

62. Bermejo AV, Ragonnaud E, Daradoumis J, Holst P. Cancer Associated Endogenous Retroviruses: Ideal Immune Targets for Adenovirus-Based Immunotherapy. *Int J Mol Sci* (2020) 21:1–21. doi: 10.3390/ijms21144843

63. Hayashi K, Tabata S, Piras V, Tomita M, Selvarajoo K. Systems Biology Strategy Reveals Pkcd Is Key for Sensitizing TRAIL-Resistant Human Fibrosarcoma. *Front Immunol* (2015) 6:659. doi: 10.3389/fimmu.2014.00659

64. Messaritakis I, Stogiannitsi M, Koulouridi A, Sfakianaki M, Voutsina A, Sotiriou, et al. Evaluation of the Detection of Toll-Like Receptors (Tlrs) in Cancer Development and Progression in Patients With Colorectal Cancer. *PloS One* (2018) 13:e0197327. doi: 10.1371/journal.pone.0197327

65. Kasurinen A, Hagström J, Laitinen A, Kokkola A, Böckelman C, Haglund C. Evaluation of Toll-Like Receptors as Prognostic Biomarkers in Gastric Cancer: High Tissue TLR5 Predicts a Better Outcome. *Sci Rep* (2019) 9:1–10. doi: 10.1038/s41598-019-49111-2

66. Urban-Wojciuk Z, Khan MM, Oyler BL, Fåhraeus R, Marek-Trzonkowska N, Nita-Lazar A, et al. The Role of Tlrs in Anti-Cancer Immunity and Tumor Rejection. *Front Immunol* (2019) 10:2388. doi: 10.3389/fimmu.2019.02388

67. Manghera M, Ferguson-Parry J, Lin R, Douville RN. NF-κB and IRF1 Induce Endogenous Retrovirus K Expression *via* Interferon-Stimulated Response Elements in Its 5′ Long Terminal Repeat. *J Virol* (2016) 90:9338–49. doi: 10.1128/jvi.01503-16

68. Römer C. Viruses and Endogenous Retroviruses as Roots for Neuroinflammation and Neurodegenerative Diseases. *Front Neurosci* (2021) 0:648629. doi: 10.3389/FNINS.2021.648629

69. Kremer D, Schichel T, Förster M, Tzekova N, Bernard C, van der Valk P, et al. Human Endogenous Retrovirus Type W Envelope Protein Inhibits Oligodendroglial Precursor Cell Differentiation. *Ann Neurol* (2013) 74:721–32. doi: 10.1002/ANA.23970

70. Kremer D, Förster M, Schichel T, Göttle P, Hartung HP, Perron P, et al. The Neutralizing Antibody Gnbac1 Abrogates Herv-W Envelope Protein-Mediated Oligodendroglial Maturation Blockade. *Mult Scler* (2015) 21:1200–3. doi: 10.1177/1352458514560926

71. Stearrett N, Dawson T, Rahnavard A, Bachali P, Bendall ML, Zeng C, et al. Expression of Human Endogenous Retroviruses in Systemic Lupus Erythematosus: Multiomic Integration With Gene Expression. *Front Immunol* (2021) 12:661437. doi: 10.3389/FIMMU.2021.661437

72. Saa PA, Nielsen LK. Formulation, Construction and Analysis of Kinetic Models of Metabolism: A Review of Modelling Frameworks. *Biotechnol Adv* (2017) 35:981–1003. doi: 10.1016/j.biotechadv.2017.09.005

73. Helmy M, Gohda J, Inoue J, Tomita M, Tsuchiya M, Selvarajoo K. Predicting Novel Features of Toll-Like Receptor 3 Signaling in Macrophages. *PloS One* (2009) 4:e4661. doi: 10.1371/journal.pone.0004661

74. Hayashi K, Piras V, Tabata S, Tomita M, Selvarajoo K. A Systems Biology Approach to Suppress TNF-Induced Proinflammatory Gene Expressions. *Cell Commun Signal* (2013) 11:84. doi: 10.1186/1478-811X-11-84

75. Piras V, Hayashi K, Tomita M, Selvarajoo K. Enhancing Apoptosis in TRAIL-Resistant Cancer Cells Using Fundamental Response Rules. *Sci Rep* (2011) 1:1–12. doi: 10.1038/srep00144

76. Toya Y, Shimizu H. Flux Analysis and Metabolomics for Systematic Metabolic Engineering of Microorganisms. *Biotechnol Adv* (2013) 31:818–26. doi: 10.1016/j.biotechadv.2013.05.002

77. Checcoli A, Pol JG, Naldi A, Noel V, Barillot E, Kroemer G, et al. Dynamical Boolean Modeling of Immunogenic Cell Death. *Front Physiol* (2020) 11:590479. doi: 10.3389/FPHYS.2020.590479

78. Selvarajoo K. A Systems Biology Approach to Overcome TRAIL Resistance in Cancer Treatment. *Prog Biophys Mol Biol* (2017) 128:142–54. doi: 10.1016/j.pbiomolbio.2017.02.009

79. Selvarajoo K, Tomita M, Tsuchiya M. Can Complex Cellular Processes be Governed by Simple Linear Rules? *J Bioinform Comput Biol* (2009) 7:243–68. doi: 10.1142/S0219720009003947

80. Zañudo GTJ, Steinway SN, Albert R. Discrete Dynamic Network Modeling of Oncogenic Signaling: Mechanistic Insights for Personalized Treatment of Cancer. *Curr Opin Syst Biol* (2018) 9:1–10. doi: 10.1016/J.COISB.2018. 02.002

81. Werner SL, Barken D, Hoffmann A. Stimulus Specificity of Gene Expression Programs Determined by Temporal Control of IKK Activity. *Sci (80-)* (2005) 309:1857–61. doi: 10.1126/science.1113319

82. Hendrata M, Sudiono J. A Computational Model for Investigating Tumor Apoptosis Induced by Mesenchymal Stem Cell-Derived Secretome. *Comput Math Methods Med* (2016) 9:2016. doi: 10.1155/2016/4910603

83. Cheong R, Hoffmann A, Levchenko A. Understanding NF-κB Signaling *via* Mathematical Modeling. *Mol Syst Biol* (2008) 4:192. doi: 10.1038/msb.2008.30

84. Santos SDM, Verveer PJ, Bastiaens PIH. Growth Factor-Induced Mapk Network Topology Shapes Erk Response Determining Pc-12 Cell Fate. *Nat Cell Biol* (2007) 9:324–30. doi: 10.1038/ncb1543

85. Bruggeman FJ, Westerhoff HV, Hoek JB, Kholodenko BN. Modular Response Analysis of Cellular Regulatory Networks. *J Theor Biol* (2002) 218:507–20. doi: 10.1006/jtbi.2002.3096

86. Bekisz S, Geris L. Cancer Modeling: From Mechanistic to Data-Driven Approaches, and From Fundamental Insights to Clinical Applications. *J Comput Sci* (2020) 46:101198. doi: 10.1016/J.JOCS.2020.101198

87. Helmy M, Smith D, Selvarajoo K. Systems Biology Approaches Integrated With Artificial Intelligence for Optimized Metabolic Engineering. *Elsevier BV* (2020) 11:e00149. doi: 10.1016/j.mec.2020.e00149

88. Gupta S, Madhu MK, Sharma AK, Sharma VK. Proinflam: A Webserver for the Prediction of Proinflammatory Antigenicity of Peptides and Proteins. *J Transl Med* (2016) 14:1–10. doi: 10.1186/S12967-016-0928-3

89. Feng J, Zhang H, Li F. Investigating the Relevance of Major Signaling Pathways in Cancer Survival Using a Biologically Meaningful Deep Learning Model. *BMC Bioinforma* (2021) 22:1–13. doi: 10.1186/S12859-020-03850-6

90. Teyra J, Kelil A, Jain S, Helmy M, Jajodia R, Hooda Y, et al. Large-Scale Survey and Database of High Affinity Ligands for Peptide Recognition Modules. *Mol Syst Biol* (2020) 16:e9310. doi: 10.15252/MSB.20199310

91. Xu C, Jackson SA. Machine Learning and Complex Biological Data. *Genome Biol* (2019) 20:1–4. doi: 10.1186/s13059-019-1689-0

92. Reis BS, Jungbluth AA, Frosina D, Holz M, Ritter E, Nakayama E, et al. Prostate Cancer Progression Correlates With Increased Humoral Immune Response to a Human Endogenous Retrovirus GAG Protein. *Clin Cancer Res* (2013) 19:6112–25. doi: 10.1158/1078-0432.CCR-12-3580

93. Ma Y, Liu R, Lv H, Han J, Zhong D, Zhang X. A Computational Method for Prediction of Matrix Proteins in Endogenous Retroviruses. *PloS One* (2017) 12 (5):e0176909. doi: 10.1371/journal.pone.0176909

94. Dey L, Chakraborty S, Mukhopadhyay A. Machine Learning Techniques for Sequence-Based Prediction of Viral–Host Interactions Between SARS-Cov-2 and Human Proteins. *BioMed J* (2020) 43:438–50. doi: 10.1016/j.bj.2020.08.003

95. Kojima S, Yoshikawa K, Ito J, Nakagawa S, Parrish NF, Horie M, et al. Virus-Like Insertions With Sequence Signatures Similar to Those of Endogenous Nonretroviral RNA Viruses in the Human Genome. *Proc Natl Acad Sci U.S.A.* (2021) 118(5). doi: 10.1073/pnas.2010758118

96. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, et al. Intrinsic Retroviral Reactivation in Human Preimplantation Embryos and Pluripotent Cells. *Nature* (2015) 522:221–46. doi: 10.1038/nature14308

97. Pyeon D, Vu L, Giacobbi NS, Westrich JA. The Antiviral Immune Forces Awaken in the Cancer Wars. *PloS Pathog* (2020) 16:e1008814. doi: 10.1371/journal.ppat.1008814

98. Daskalakis M, Brocks D, Sheng YH, Islam MS, Ressnerova A, Assenov Y, et al. Reactivation of Endogenous Retroviral Elements *via* Treatment With DNMT- and HDAC-Inhibitors. *Cell Cycle* (2018) 17:811–22. doi: 10.1080/15384101.2018.1442623

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.