

Published in final edited form as:

Annu Rev Immunol. 2023 April 26; 41: 99–125. doi:10.1146/annurev-immunol-101721-033341.

The Immunological Conundrum of Endogenous Retroelements

George Kassiotis^{1,2}

George Kassiotis: george.kassiotis@crick.ac.uk

¹Retroviral Immunology Laboratory, The Francis Crick Institute, London, United Kingdom

²Department of Infectious Disease, Faculty of Medicine, Imperial College London, London, United Kingdom

Abstract

Our defenses against infection rely on the ability of the immune system to distinguish invading pathogens from self. This task is exceptionally challenging, if not seemingly impossible, in the case of retroviruses that have integrated almost seamlessly into the host. This review examines the limits of innate and adaptive immune responses elicited by endogenous retroviruses and other retroelements, the targets of immune recognition, and the consequences for host health and disease. Contrary to theoretical expectation, endogenous retroelements retain substantial immunogenicity, which manifests most profoundly when their epigenetic repression is compromised, contributing to autoinflammatory and autoimmune disease and age-related inflammation. Nevertheless, recent evidence suggests that regulated immune reactivity to endogenous retroelements is integral to immune system development and function, underpinning cancer immunosurveillance, resistance to infection, and responses to the microbiota. Elucidation of the interaction points with endogenous retroelements will therefore deepen our understanding of immune system function and contribution to disease.

Keywords

endogenous retrovirus; endogenous retroelement; interferon response; adaptive immunity

Introduction

To mount an appropriate response against infectious pathogens, the immune system must gauge the threat each pathogen poses, by analyzing its molecular components and their dissimilarity to the host. This task is hindered by enormous variability in pathogen persistence and immunogenicity (1). These two properties are intricately linked, as pathogen persistence in the host requires evasion of host immunity. Consequently, many pathogens have evolved mechanisms for stealth or active suppression of the immune response (1). Thereby, life-long infection can be achieved, in spite of or at an equilibrium with

This work is licensed under a [BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

Disclosure Statement

G.K. is a scientific co-founder of EnaraBio and a member of its scientific advisory board.

host immunity. An example of this strategy is mimicry, whereby pathogen components molecularly or functionally resemble the host.¹ With all their components made by the host cell, viruses are particularly successful at mimicry, which they use at multiple levels.

Molecular mimicry has long been proposed as a strategy that allows persistence of pathogens whose antigens are similar to host antigens by exploiting self-tolerance (2). However, as self-tolerance may be incomplete, molecular mimicry of T cell and B cell antigens is also a potential trigger of autoimmunity (3). Viral mimicry of key host immune mediators also allows viruses to suppress or subvert host immunity, and several viruses express functional host protein mimics, such as cytokines, chemokines, and their receptors (4), or antagonists of antiviral host signaling cascades, such as the interferon response (5, 6).

Blurring the boundaries between virus and host is an evolutionary strategy of many virus families establishing chronic infection, but none has taken it to the extreme form that retroviruses have. By virtue of their ability to integrate functional copies of their genome into the host germline DNA, they can become a true part of the host. Indeed, our DNA is host to hundreds of thousands of retrovirus integrations, acquired over successive waves of retrovirus infection and further amplification in the germline. It is also host to numerous other types of retrotransposable elements, collectively referred to here as endogenous retroelements (EREs): genomic parasites that also rely on reverse transcription and integration of their genomes into the host DNA.

Although the vast majority of ERE integrations are mutated, incomplete, or defective genomic copies, many have retained the ability to complete some or all of the steps in the retroelement replication cycle, including transcription, translation or reverse transcription, and reintegration. Retention of this ability, combined with the sheer number of EREs in the genome, provides an extensive interface of interaction with host physiology and pathology. The influence of regulatory sequences provided by EREs on host gene function has long been recognized (7, 8), and immune genes are no exception. The full extent to which EREs shape immune gene function and evolution is likely underestimated owing to the incomplete annotation of ERE transcripts, but appreciation is growing and comprehensive efforts are underway; these are not the focus of this review.

This extreme form of parasitism—where EREs become part of self—may be considered the endgame for host immunity against them. Nevertheless, both innate and adaptive host immune responses against EREs are observed, and the supporting evidence is reviewed here. Also reviewed is the potential association of retained ERE immunogenicity not only with the control of EREs themselves, but also with overall immune reactivity to foreign or other self-targets, whereby they might tune resistance to infection, equilibrium with commensals, immunosurveillance of cancer, or development of autoimmunity.

¹The term viral mimicry is used here to describe a virus strategy whereby virus components molecularly or functionally resemble respective host components. In its simplest form, this strategy is analogous to those employed by many organisms that avoid detection by blending into the background, such as the octopus mimicking the seabed, as Aristotle noted. The term is also used elsewhere in the literature to describe the reverse phenomenon, whereby the host produces molecular patterns that resemble those usually produced during virus infection and that induce an antiviral state, characterized by a type I interferon response. Where interferon responses are triggered by endogenous retroviruses—which are, rather than mimic, viruses—we refer to this phenomenon simply as induction of interferon responses.

Endogenous Retroelement Life Cycle and Evolution

With nucleic acids being the cornerstone of all life, it is perhaps expected that genetic material will be exchanged between exogenous viruses, EREs, and their hosts, providing the substrate for coevolution. The focus in this review is on EREs (class I transposable elements), defined as transposable elements with reverse transcription and host germline integration of their genomes as obligatory steps in their replication cycle (9). This definition excludes DNA (class II) transposable elements, which lack an RNA intermediate (9). It also excludes endogenous Borna-like N(EBLN) elements that are fixed in the human population (10) and human herpesvirus 6 copies, which are found in the germline of ~1% of humans, the integration of which was likely accidental.

EREs in the human genome comprise phylogenetically diverse families, but a major distinction relates to their genomic structure and presumed origin (Figure 1). Deriving from germline infection by exogenous retroviruses, endogenous retroviruses (ERVs) exhibit the typical structure of retroviral genomes with *gag*, *pro-pol*, and *env* open reading frames (ORFs) flanked by directly repeated long terminal repeats (LTRs). Whereas *gag*, encoding capsid proteins, and *pro-pol*, encoding enzymatic activities, are necessary for autonomous replication of ERVs, *env*, encoding the envelope glycoprotein that mediates entry, is necessary only for infection of new target cells. Consequently, once in the germline, even ERVs that have lost the *env* ORF can still amplify their copies, sometimes more successfully, by retrotransposition (11).

The ability of ERVs to complete the replication cycle is ultimately disrupted by accumulated mutation or loss of ORFs or LTR sequences over evolutionary time. Indeed, the vast majority of ERVs in mammalian genomes are replication defective, with only the most recently acquired copies retaining functional ORFs or replication capacity, which varies between species. This also varies within species, as recent ERV integrations may also exhibit insertional polymorphism between individuals.

Likely owing to selection of phenotypic traits, such as cancer susceptibility, caused by ERV insertional mutagenesis during their establishment, laboratory mouse strains carry an unusually high burden of complete or near-complete ERVs (12). Germline integrations of replication-competent mouse mammary tumor viruses (MMTVs) and murine leukemia viruses (MLVs) are present in several laboratory mouse strains, able to transmit as both endogenous and exogenous retroviruses (13). This duality confounded early investigations into the genetics of cancer but ultimately led to the discovery of ERVs (12). Laboratory mouse strains lacking endogenous MLVs that are able to replicate in murine cells still carry multiple related defective integrations with most ORFs still intact. When expressed in the same cell, these defective copies collectively produce all the components required to form a transducing retrovirus particle, with each functional component donated by a different provirus. Such trans-complementation permits the mobilization of individually defective endogenous MLVs, generating new integrations. Moreover, recombination between defective endogenous MLV genomes during such mobilization can also restore the replication defects, giving rise to fully infectious MLVs, transmitted as exogenous retroviruses. This process is exemplified by the restoration of infectivity, through

recombination, of *Emv2*, a single-copy defective endogenous MLV with an ecotropic *env* in commonly used C57BL/6 mice (14–18). Infectious *Emv2* recombinant MLVs have been discovered in murine cancer cell lines, as well as in immunodeficient mouse strains, highlighting the relatively high frequency of infectivity restoration (14–18).

Human ERVs (HERVs) are thought to be generally older and, therefore, more defective than murine ERVs. Consequently, no HERV is thought to have retained full replication capacity. Nevertheless, proviruses with seemingly intact ORFs are found within the HERV-K(HML-2) family (19). The human MMTV-like (HML) families of HERV-K share similarities with MMTVs, as the name suggests, but represent diverse groups of proviruses that entered our ancestors' germlines at different times during evolution (19). The HERV-K(HML-2) family in particular includes the most recently acquired integrations, several of which are human specific and insertionally polymorphic (19). It also includes the most complete proviruses compared to any other family, leading to speculation of a naturally infectious HERV-K(HML-2) provirus. However, evidence for replication of a HERV-K(HML-2) virus in humans is still lacking. HERV-K(HML-2) proviruses appear unable to complete the replication cycle in humans and chimpanzees, but recent analysis of the gorilla genome identified HERV-K(HML-2) viruses that bear the hallmarks of very recent integration, raising the possibility of an infectious HERV-K(HML-2) virus still present in gorillas (20). Moreover, retrotransposition of a HERV-K(HML-2)-based reporter construct and mobilization of endogenous HERV-K(HML-2) have recently been reported in cells expressing the pluripotency transcription factor SOX2 in vitro (21). Thus, depending on the mouse strain, murine ERVs are able to complete the replication cycle in part or in full, and HERVs may be able to complete at least some of the steps.

Grouped with ERVs are also the mammalian apparent LTR retrotransposons (MaLRs), whose genomes are also flanked by LTRs (Figure 1). These, however, are nonautonomous ancient retro-transposons that lack the canonical retrovirus ORFs and have relied on ERVs for transposition (22). In addition to these LTR retroelements (ERVs and MaLRs), a larger proportion of the genome comprises EREs lacking LTRs (Figure 1). Phylogenetic analyses of the reverse transcriptase (RT) suggest common ancestry of LTR and non-LTR retroelements (9, 23). However, the replication cycles of these two types of ERE do exhibit notable differences. Long interspersed nuclear elements 1 (LINE-1) are a group of abundant non-LTR retroelements that contain autonomous retrotransposition-competent copies (24). These carry two functional ORFs, encoding an RNA-binding protein (ORF1p) and an enzyme with endonuclease and RT activities (ORF2p). Importantly, reverse transcription of LINE-1 RNA is primed by the integration target site and, therefore, takes place in the nucleus. Although ORF2p exhibits *cis* preference for LINE-1 mRNA binding, it is able to bind unrelated RNA transcribed from other EREs or cellular genes. Indeed, nonautonomous retrotransposition of other groups of non-LTR retroelements, including the prolific primate-specific *Alu* elements and other short interspersed nuclear elements (SINEs), and the hominid-specific composite SINE-VNTR-*Alu* (SVA) elements, relies on LINE-1 ORF2p (24).

The distinct life styles of LTR and non-LTR retroelements (Figure 1), together with their relative copy number and transcriptional patterns, can greatly influence the creation

of ligands for innate and adaptive immune sensors and receptors and, hence, ERE immunogenicity.

Endogenous Retroelements and Innate Immunity

All cellular organisms possess cell-intrinsic defenses against parasitism, but the evolution of multi-cellular organisms provided the opportunity to couple such cell-intrinsic immunity to organismal immunity through cellular communication. One prime example is the evolution of the interferon response system, initiated in some cells to warn other cells of infection (5, 6). Given the vast array of viruses and other parasites the immune system has to recognize and defend against, numerous sensors or receptors for different molecular patterns have evolved and are strategically placed in different subcellular locations (25–27). While some sensors have specificity for molecular patterns that are exclusively expressed by microbes, others recognize patterns that are also produced by the host but are usually shielded under physiological conditions (25–29). Consequently, discrimination between self-expressed and non-self-expressed ligands by the latter groups of sensors is not always absolute. Such discrimination may be particularly challenging for ligands expressed by EREs, which straddle host and virus.

Immune homeostasis would require the establishment of equilibrium between physiologically present ligands, including those potentially produced by EREs and tonic or regulated innate immune signaling. Disruption of this equilibrium by the introduction of foreign molecular patterns or dysregulation of self-expressed ligands would then trigger an innate immune response. As innate immune signaling cascades generally converge, particularly when type I interferon responses (referred to here as interferon responses) are elicited, it is often challenging to identify a single source or trigger for the response. Indeed, ligands from diverse sources can trigger the same sensor, and diverse sensors can elicit overlapping immune responses, such as transcription of interferon-stimulated genes (ISGs). Cases where specific ERE-expressed ligands uniquely initiate a signaling cascade, leading to a specific outcome, are the focus of intense investigation and are beginning to emerge (30, 31). However, much of the existing evidence points to a contribution of EREs to innate immune responses that may also be triggered by alternative self or foreign ligands, with the relative contribution of each source depending on the context (28, 29).

Innate Immune Stimulation by Long Terminal Repeat Elements

Similarly to exogenous retroviruses, individual steps in the typical replication cycle of ERVs create nucleic acid replication intermediates with the potential to engage multiple innate sensors (Figure 2). Single-stranded RNA (ssRNA) produced by ERV transcription is not inherently more immunogenic than other cytosolic RNA transcribed from cellular genes, but it may gain better access to the endosomal ssRNA sensors Toll-like receptor 7 (TLR7) and TLR8. The formation of transducing particles by complementation of defective ERVs provides the opportunity for such endosomal sensors to detect incoming ERV particles the same way incoming infectious viruses are detected. Moreover, for ERVs that have lost *env* and have adopted an intracellular life style, virus budding into the endoplasmic reticulum (ER) or endosomal vesicles may provide more direct access to endosomal sensors. More

broadly, TLR7 is also accessible to cytoplasmic RNA through autophagy (32). Studies of genetic deficiencies in TLR7 or associated molecules in mice uncovered their critical role in the control of infectious recombinants derived from endogenous MLV (17, 33). However, ssRNA recognition likely targeted rescued recombinants more than defective endogenous MLV precursors in this context. Extracellular HERV-K(HML2) RNA has also been proposed as an endogenous ligand for TLR7 and TLR8 (34).

Reverse transcription of ERV genomes may also trigger innate sensors. Cytosolic DNA sensors lack sequence specificity and can be activated by multiple sources of DNA (26). Evidence supporting a role for reverse-transcribed DNA comes from studies of RT inhibitors in humans and animal models (30, 35–37), although the use of such inhibitors does not discriminate between ERV and LINE-1 RT, and it is more likely that these effects are mediated by LINE-1 RT, which, in principle, can reverse transcribe any polyadenylated RNA. Nevertheless, ERV RT has been specifically implicated in innate immune activation in some settings (38).

Unrelated to the typical retrovirus replication cycle, bidirectional expression of ERVs can generate complementary RNA that forms long double-stranded RNA (dsRNA), particularly in the context of natural or epigenetic drug-induced hypomethylation of cancer cells, which has been the subject of excellent reviews (31). Bidirectional ERV transcription may result from the bidirectional promoter activity of the proviral LTRs or from alternative adjacent promoters. Of note, dsRNA can form from certain IFN- γ -inducible loci carrying antisense ERV integrations in the 3' untranslated region (UTR), pairing sense transcripts initiated by the STAT1-activated gene promoter and antisense transcripts initiated by the ERV LTR (39). ERV-derived dsRNA has been reported to activate TLR3 or MDA5 and downstream MAVS (40, 41). ERV-derived dsRNA produced by cancer cells has also been suggested to activate TLR3 in endothelial cells in mouse cancer models (42).

In addition to immunogenic nucleic acid replication intermediates, expression of canonical ERV proteins has also been associated with triggering of immune signaling or inflammation. Such activity may arise from the biological function of a particular ERV protein. For example, the Rec accessory protein of HERV-K(HML2) is a functional homolog of the Rev and Rex proteins of HIV-1 and human T cell leukemia virus I (HTLV-I), respectively, necessary for nuclear export of unspliced or partially spliced virus RNA (43). Rec-mediated export of potentially immunogenic RNA has been implicated in the transient innate immune activation during human embryo development, where ERVs are expressed as part of global epigenetic reprogramming (44).

The envelope glycoproteins of certain ERVs, including *Env2* in mice and HERV-K(HML2) in humans, have recently been suggested to initiate signaling cascades in immune and cancer cells, when ligated with anti-envelope antibodies (45–48). Further dissection of the signaling pathways and their potential role in virus replication will help determine whether such activity is part of the physiological function of envelope glycoproteins. Similarly, overexpression of Syncytin-1, an exapted envelope glycoprotein (49), triggers inflammatory responses in human astrocytes in vitro and in transgenic mice in vivo (50), and overexpression of HERV-K(HML2) envelope glycoprotein triggers death in human neurons

in vitro and in transgenic mice in vivo (51). An inflammatory response has also been reported in the brains of transgenic mice with genetic deletion of the epigenetic repressor *Trim28* in neural progenitor cells, which was linked with upregulation and accumulation of ERV-derived proteins, such as aggregates of intracisternal A particle (IAP) Gag proteins (52). Cell death has also been attributed to ERV protein accumulation following deletion of the epigenetic repressor *Setdb1* in pro-B cells (53). It remains unclear whether the reported toxicity of ERV proteins derives from their biological activities or their accumulation in sensitive cell types. Lastly, immunosuppressive activity has been described for a shared domain of several ERV envelope glycoproteins (49), but the underlying mechanism by which this activity may be exerted remains incompletely understood.

Innate Immune Stimulation by Non-LTR Elements

Nucleic acid species and replication intermediates derived from non-LTR retrotransposons also possess immunogenic potential (Figure 2). Transfected *Alu* RNA has been reported to activate TLR7 and TLR8, but also to increase the immunogenicity of associated proteins in ribonucleoprotein complexes (54, 55). Similarly to ERVs, bidirectional transcription of LINE-1 elements following loss of epigenetic regulators such as SETDB1 or components of the HUSH (human silencing hub) complex generates potentially immunogenic dsRNA in human cells (56, 57).

A connection between ERE derepression and the interferon response that follows treatment of cells with the epigenetic drug 5-aza-2'-deoxycytidine (5-Aza) has long been established. Leonova et al. (58) implicated dsRNA formation by SINE repeats, independently transcribed by polymerase III, in the induction of a suicidal interferon response triggered by 5-Aza treatment of murine primary cells, as well as cancer cell lines lacking p53. These findings suggested the existence of an alarm system monitoring the transcriptional activity of normally repressed EREs that, when triggered, leads to the induction of potent interferon responses. Subsequent studies with 5-Aza treatment of cancer cell lines (40, 41) or loss of the epigenetic repressor LSD1 (59) focused on ERVs as the potential dsRNA triggers of the interferon response. However, a major role for SINEs in the induction of the interferon response was recently highlighted by studies of the dsRNA-specific adenosine deaminase ADAR1, loss of which causes embryonically lethal interferon induction, dependent on recognition of self-dsRNA through MDA5 and its downstream adaptor MAVS (60–62). ADAR1 catalyzes adenosine-to-inosine (A-to-I) editing in dsRNA, thereby disrupting adenosine:uracil (A:U) base pairing and complementarity (63, 64). Investigation of ADAR1 targets revealed the dominant contribution of dsRNA structures formed by the transcription of inverted *Alu* repeats (63, 64). The presence of two or more *Alu* copies in reverse orientation in a single RNA transcript provides the sequence complementarity to form long hairpin loops that activate dsRNA sensors. Indeed, the lethal interferon response that develops in the absence of ADAR1 activity is thought to be driven by unedited dsRNA from inverted murine SINE (62) and human *Alu* repeats (65, 66), typically in the 3' UTR of a limited number of genes. Moreover, thorough investigation of immunogenic dsRNA bound to MDA5 in 5-Aza-treated cancer cells identified inverted *Alu* repeats, and not ERVs as previously thought, as the main source (67). Under these conditions, intronic and intergenic

inverted *Alu* repeats that are normally prevented from reaching the cytoplasm predominated the MDA5-bound fraction (67).

In addition to activating the MDA5-MAVS signaling cascade, self-dsRNA has recently been suggested to trigger inflammatory cell death through activation of ZBP1, a sensor of left-handed double-helical (Z-form) DNA and RNA (68–75). Where examined, the origin of Z-form dsRNA activating ZBP1 could be traced back to EREs. Indeed, activation of ZBP1 in mice with deficiency in necroptosis regulator RIPK1 or FADD (68) and in humans and mice with deficiency in SETDB1 (69) has implicated dsRNA from SINEs and ERVs, respectively. Moreover, several studies have now described an essential role for ZBP1 in the inflammatory cell death and fatal disease that develops when ADAR1 activity is reduced or when the Z α domain of its p150 isoform, responsible for Z-form dsRNA binding, is mutated (71–75). Mutation of the ADAR1 p150 Z α domain and, therefore, loss of binding to Z-form dsRNA do not appear to compromise the overall editing activity of ADAR1 (72, 73, 76). Although a small proportion of Z-form dsRNA, mostly derived from inverted SINE repeats, may be edited in a Z α domain–dependent manner (72, 76), differential interferon inducibility of ADAR1 isoforms may also produce such shifts in editing targets. These observations suggest that ADAR1 p150 may prevent ZBP1 activation by sequestering, rather than editing, Z-form dsRNA (72, 73, 76). The intersection of ADAR1 and ZBP1 pathways remains to be investigated, but it may involve physical interaction through the Z α domains binding to the same Z-form dsRNA molecule, or competition for immunogenic SINE-derived Z-form dsRNA (71, 75). Of note, while ZBP1 activation has been primarily connected with the induction of various forms of cell death, including necroptosis, recent studies also revealed an indispensable contribution to the MAVS-dependent interferon response triggered when ADAR1 Z-form dsRNA binding is lost (73, 75).

Thus, Z-form dsRNA, or incompletely edited A-form dsRNA derived from SINEs, can trigger multiple signaling cascades converging on the interferon response. A central role for SINE-derived dsRNA in triggering an interferon response in multiple settings is further supported by the observation that expression of SINEs that can generate Z-form dsRNA or are targeted for A-to-I editing by ADAR1 is also induced by interferon (68, 72, 77). Moreover, their sensors, including ZBP1, as well as the Z α domain–bearing p150 isoform of ADAR1, are also interferon inducible in mice (71) and in human cancer (78), pointing to a finely regulated feed-forward loop of interferon induction.

Potentially immunogenic nucleic acids are additionally produced by the reverse transcription of LINE-1 and SINE RNA, a process that relies on LINE-1 ORF2p (79). A role for ORF2p-mediated reverse transcription was supported by findings in rare syndromes caused by dysregulation of nucleic acid metabolism or sensing, such as Aicardi-Goutières syndrome (AGS) (29). LINE-1 derepression and elevated ORF2p activity have been proposed to trigger the interferon response and inflammation that accompanies cellular senescence (80, 81). Furthermore, ORF2p activity is considered responsible for the interferon response that increases tumor immunogenicity in mice and humans, an enhanced form of which has been described for the cancer-resistant blind mole rat (82).

To prevent recognition of nuclear DNA, the activity of DNA sensors, such as cGAS and its downstream adaptor STING, is typically restricted to the cytoplasm (26). Although target-primed reverse transcription mediated by ORF2p is restricted to the nucleus, recent evidence suggests that LINE-1 and SINE complementary DNA (cDNA) can also be made in the cytoplasm. Indeed, LINE-1 cDNA and self-primed *Alu* cDNA have been detected in the cytoplasmic fraction of murine or human cells (80, 83), providing a possible mechanism for triggering the interferon response.

Lastly, the fully functional ORF2p, as well as numerous ORF2p copies that retain only the endonuclease activity, can introduce DNA double-strand breaks, triggering a DNA damage response (84). In turn, genomic DNA damage can initiate the cGAS-STING signaling cascade, leading to an interferon response (85), thus providing an alternative means of interferon activation by ORF2p.

Endogenous Retroelements and Adaptive Immunity

The combinatorial process of T cell receptor (TCR) and B cell receptor (BCR) gene segment rearrangement produces random specificities, some of which may be directed against self-antigens. Consequently, powerful mechanisms including negative selection of potentially autoreactive antigen receptors; naturally suppressive cells, such as regulatory T cells (Tregs); and inhibitory ligand/receptor axes, such as PD-L1/PD-1, have evolved to ensure immunological tolerance to self-antigens. Nevertheless, autoreactive TCRs and BCRs do develop and do occasionally break through regulatory and inhibitory controls to cause autoimmunity.

As part of self, proteins encoded by EREs should also be immunologically tolerated by the adaptive immune repertoire. However, specific features of ERE proteins may substantially modify their immunogenic or tolerogenic activity.

Firstly, certain endogenous retroviral envelope glycoproteins exhibit long-recognized superantigen activity for reactive TCR V β families in mice and humans (86). These include several endogenous MMTV proviruses in mice (87) and a HERV-K18 provirus in the first intron of the *CD48* gene in humans (88). Superantigen reactivity may be particularly enriched in Tregs, which typically escape negative selection by self-antigens, including retroviral superantigens (89). Indeed, in the commonly used C57BL/6 strain of laboratory mice, 10% of all Tregs are reactive with an endogenous MMTV superantigen (90), and their function can be directly modulated by changes in endogenous MMTV superantigen expression during inflammation or unrelated infection (91, 92).

Secondly, not all ERE-encoded proteins are or need to be expressed under physiological conditions. In contrast to all other host proteins, which are evolutionarily selected for a given function, most ERE protein products are unlikely to have retained their original function or assumed new function. Their expression is not, therefore, subject to positive selection, and it may even be actively suppressed as part of the epigenetic control that the host imposes on the transcription of many EREs. Insufficient expression of such EREs under physiological conditions also leads to incomplete immunological tolerance of their products and instead

may create immunological ignorance. However, failure of these control mechanisms, as it occurs in cancer or in autoimmunity, may also lead to immune reactivity against these otherwise ignored antigens.

A third feature of certain ERE-encoded proteins that may enhance their immunogenicity is their assembly into inherently immunogenic structures, such as virus particles, capsids, and ribonucleoprotein complexes. Formation of transducing particles through complementation between defective proviruses, such as endogenous MLVs, and even restoration of infectivity through recombination, giving rise to fully infectious retroviruses (12), is greatly facilitated by better overall conservation of ERVs in laboratory mice than in humans. Nonetheless, formation of retrovirus particles by human ERVs has also long been suspected or directly observed. Indeed, retrovirus particles have been detected in normal human placentas (93–95), attributed to HERV-K proviruses (96). Similarly, retrovirus particles produced by human teratocarcinoma cell lines were traced to HERV-K proviruses (97). HERV-K(HML-2)-derived particles are present in the plasma of lymphoma patients and can mobilize HERV-K-related proviruses (98, 99). Although it remains unclear whether other human ERVs can produce particles, the formation of particles by at least some HERV-K proviruses would enhance the immunogenicity of HERV-K proteins, as well as proteins from other ERVs that may be incorporated into HERV-K particles.

Lastly, comparison of the HLA class I-presented epitopes derived from ERVs with those from other human proteins or exogenous viruses infecting humans revealed unique features of ERV-derived epitopes that could potentially enhance their immunogenicity (100). ERV epitopes showed higher sequence identity with virus than human protein epitopes and enrichment for amino acid residues in specific epitope positions (100).

Adaptive Immune Responses to Endogenous Retrovirus Proteins in Humans

While the contribution to immunogenicity of unique features of ERE-encoded proteins remains to be quantified, spontaneous adaptive immune responses against such proteins have been consistently reported both in mice and in humans. The most frequent targets of adaptive immune responses appear to be canonical ERV proteins, likely owing to better or earlier annotation of these proteins, examples of which are listed below.

ERV3–1 envelope—One of the earliest demonstrations of adaptive immune reactivity to an ERV-encoded protein in humans is that against the envelope glycoprotein of the *ERV3–1* provirus. A member of the HERV3 family (also known as HERV-R), *ERV3–1* is a single-copy provirus on chromosome 7q11.21 and one of the best-studied ERVs (101, 102). It is present only in old-world primates, except gorillas, and the human copy has sustained mutations in all ORFs except the *env* ORF (101, 102). The *ERV3–1* envelope glycoprotein is expressed in normal syncytiotrophoblasts and had been speculated to function similarly to syncytins, endogenous retroviral envelope glycoproteins that are essential for syncytiotrophoblast formation during placentation (49). However, despite its preservation, the *ERV3–1 env* ORF carries a polymorphic premature stop codon in ~1% of White people that does not compromise successful pregnancy (103, 104).

In early studies, sera from pregnant women, as well as those from patients with Sjögren syndrome or systemic lupus erythematosus (SLE), exhibited increased antibody reactivity against a peptide corresponding to a predicted epitope from the *ERV3-1* envelope glycoprotein and to recombinant protein (105). The highest antibody levels were found in the sera from mothers of infants with congenital heart block, a rare pregnancy-associated autoimmune disorder (105). However, in a subsequent study of 12 mothers of infants with congenital heart block, none was homozygous for the *ERV3-1 env* allele with the premature stop codon (103), arguing against a mechanism whereby lack of immunological tolerance due to lack of *ERV3-1* envelope expression in mothers permits the induction of a pathogenic response against this glycoprotein expressed in the fetus as a paternal alloantigen.

More recently, IgG antibodies reacting with the full-length *ERV3-1* envelope glycoprotein expressed on the target cell surface have been detected in sera from a fraction of juvenile-onset SLE patients, as well as patients with multisystem inflammatory syndrome in children (MIS-C) following SARS-CoV-2 infection (106). Collectively, these studies highlight a possible association of *ERV3-1* envelope antibodies with autoimmune manifestations.

HERV-K(HML-2)—Arguably, the most frequently targeted canonical ERV-encoded proteins in humans belong to the HERV-K(HML-2) family, which includes the most recent and most complete proviruses (19). Consequently, numerous HERV-K(HML-2) proviruses in the human genome carry intact and highly similar *gag*, *pol*, and *env* ORFs (19). In turn, it is often difficult to establish which particular HERV-K(HML-2) copy might have induced a T cell or B cell response that cross-reacts with the products of other similar copies. For example, of the 14 HERV-K(HML-2) proviruses with a fully or partially intact *env* ORF in the human genome, at least 8 encode products with 95–98% amino acid identity between them and 98–99% amino acid identity with the HERV-K(HML-2) envelope glycoprotein consensus.

T cell responses to HERV-K(HML-2) proteins have been detected primarily in the context of cancer, targeting epitopes encoded by *gag* or *pol* ORFs (107, 108). Antibody responses to HERV-K(HML-2) envelope glycoprotein and Gag precursor protein have also been frequently observed in several cancer indications (109–117). Notably, antibodies reactive with HERV-K particle–producing teratocarcinoma cells were also detected, albeit at low titers, in the sera of a small proportion (~4%) of healthy individuals and of pregnant women (117), suggesting that spontaneous autoreactivity against HERV-K(HML-2) envelope glycoproteins can arise in individuals without overt pathological manifestations. Antibodies, primarily of the IgM class, to recombinant HERV-K(HML-2) envelope glycoprotein and Gag precursor protein were also detected by ELISA in an independent study of healthy individuals and were reported to be reduced in psoriasis patients (118).

Antibody responses to recombinant HERV-K(HML-2) envelope glycoprotein have also been reported in a recent study of healthy individuals and SLE patients (119). These were primarily of IgG subclasses and were found at comparable titers between healthy donors and SLE patients. In this study, Tokuyama et al. (119) were able to pinpoint a particular HERV-K(HML-2) provirus on chromosome 1q22, *ERVK-7* (also known as *HERV-K102*), as the most likely source of the targeted envelope glycoprotein. This provirus exhibited

the highest expression among all HERV-K(HML-2) proviruses examined both in healthy donors and in SLE patients and was also upregulated in SLE patients. This study suggests that, similarly to *ERV3-1*, the product of a single HERV-K(HML-2) provirus alone may be immunogenic and responsible for the observed immune reactivity.

Titers of antibodies reacting with the consensus HERV-K(HML-2) envelope glycoprotein expressed on the cell surface were also comparably low in children and adolescents with autoimmune rheumatic diseases and age-matched controls, but substantially higher in MIS-C patients (106), indicating that such antibodies may be induced in the weeks following acute viral infection, such as with SARS-CoV-2, but not necessarily associated with autoimmune rheumatic diseases.

HERV-E—Members of the HERV-E family have been found to be highly immunogenic particularly in clear cell renal cell carcinoma (ccRCC). Graft-versus-tumor effect in a ccRCC patient who underwent hematopoietic stem cell transplantation led to the identification of a HERV-E provirus on chromosome 6q15 (also known as CT-RCC) as the source of the tumor-associated epitope targeted by donor CD8⁺ T cells (120). This provirus shows minimal expression in healthy tissues, including the kidney, but is strongly upregulated in ccRCC, likely through the combined effects of epigenetic derepression, inactivation of the *VHL* tumor suppressor gene, and overexpression of the HERV-E LTR transcription factor HIF-2 α (121). As a result, translation products from partial *gag-pro-pol* and *env* ORFs are produced specifically in ccRCC, triggering T cell responses (120, 122).

A more comprehensive search of potentially immunogenic HERVs in ccRCC additionally identified another HERV-E family member on chromosome 19q12, also expressed specifically in this indication, with products from partial *gag-pro-pol* ORFs targeted by a substantial proportion of tumor-reactive CD8⁺ T cells (123). The repeated identification of HERV-E family members as the triggers of adaptive immune responses in ccRCC underlines the commonalities in their regulation during development of this cancer type.

HERV-H—The HERV-H family is a large grouping of diverse proviruses with some unusual characteristics. Most proviruses were acquired before the split of old-world and new-world monkeys about 40 million years ago. Some members were acquired even earlier than that, whereas other members are specific to great apes (19). Despite their age, an atypically larger proportion of HERV-H proviruses have retained proviral sequences between the two LTRs, in contrast to most other families, whose members are predominantly in solo LTR form (124). The ORFs in the retained internal sequences are not necessarily intact, and indeed, the *env* genes of HERV-H family members show a decay rate that is typical of all ERVs and are highly mutated (124). Nevertheless, numerous HERV-H family members are expressed in healthy cells and may be linked with physiological processes, most notably transcriptional regulation of human stem cell pluripotency (125). Expression of HERV-H proviruses in healthy cells does not, however, negate the immunogenicity of all copies in the genome, and CD8⁺ T cell responses have been documented against products of one particular HERV-H provirus on chromosome Xp22.3 (126). This provirus is strongly upregulated in gastrointestinal cancers, and in addition to a partial *gag* ORF, it contains an atypical ORF upstream of the *pre-gag* ORF present in certain other retroviruses (127). The

product of this atypical retroviral ORF, rather than the reported *env* product (126), was the likely target recognized by CD8⁺ T cells in colorectal cancer cell lines.

HERV-W—In multiple sclerosis patients, elevated antibody reactivity has also been reported against peptides corresponding to the envelope glycoproteins of HERV-W proviruses (128, 129), including Syncytin-1, encoded by the *ERVW-1* provirus on chromosome 7q21.2 (49), and multiple sclerosis-associated retrovirus (MSRV), a speculated but so far elusive virus thought to be related to HERV-W proviruses (130). However, reactivity against insect cell-produced, full-length versions of the envelope glycoproteins revealed no reactivity against the putative MSRV envelope glycoprotein and very rare reactivity against Syncytin-1 (131).

Additional and unconventional ERE-encoded antigens—Recent technical advances in protein identification through mass spectrometry, as well as in detection of reactive T cells, have allowed more comprehensive discovery of putative antigens encoded by ERVs and other retroelements, particularly in the context of cancer.

Using a limited list of 66 annotated HERVs in conjunction with DNA barcode-labeled MHC-I multimers, Saini et al. (132) identified CD8⁺ T cells reactive with numerous peptides encoded by canonical retroviral ORFs in members of the HERV-K, HERV-H, HERV-W, and HERV-E families in patients with hematological cancers. CD8⁺ T cells reactive with some of these peptides were also found in healthy donors, albeit less frequently, suggesting clonal expansion in cancer patients (132). In an analysis of all annotated genomic EREs, two studies identified ORFs within LTR retroelements, as well as LINE-1, SINEs, and composite SVA elements, as the source of peptides presented by MHC-I in glioblastoma (133, 134). Antibodies reactive with LINE-1 ORF1p have been detected in most SLE patients, as well as in a smaller proportion of healthy individuals, and their titers were higher in active SLE (135).

The identification of peptides derived from the translation of ERE transcripts not previously thought to have coding potential is consistent with accumulating evidence for translation of unconventional ORFs, including 5' and 3' UTRs, pseudogenes, long noncoding RNAs, short ORFs, and alternative ORFs (136–142). However, in addition to unconventional ORFs present in a given RNA molecule, alternative splicing patterns and use of alternative start or polyadenylation sites can create diverse RNA isoforms transcribed from a given locus, carrying additional putative ORFs. Indeed, increased alternative splicing in most human cancer types may create novel peptides that include predicted MHC-I binders (143). Moreover, increased intron retention in ovarian cancer (144), acute myeloid leukemia (145), and SF3B1-mutated uveal melanoma (146) has been implicated in the creation of tumor-specific T cell epitopes.

Given their genomic structure and sheer copy numbers, EREs represent the largest source of alternative start, polyadenylation, and splicing sites in the genome (7, 8). Their transcriptional utilization, therefore, has the potential to increase RNA isoform diversity quite considerably (Figure 3). Transcripts originating from or including EREs are still incompletely annotated, but recent efforts using de novo transcriptome assemblies without bias against repetitive sequences have uncovered increased ERE utilization in

novel transcripts, often in a cancer type-specific manner (147, 149). Inclusion of EREs in chimeric transcripts also overlapping unique neighboring genomic regions creates the potential for translation of protein products (Figure 3), with far less homology with products from other parts of the genome than products translated from repetitive EREs alone and canonical retroviral ORFs that may also be expressed in healthy tissues (100, 147). Indeed, recent studies integrating improved annotation of ERE-overlapping transcripts with immunopeptidomic analyses highlighted the considerable contribution of such chimeric products to the antigenic identity of cancer cells (147–150, 151). Moreover, the overlap of antigenic transcripts identified in each study using different methodologies is partial, suggesting that the complete contribution of chimeric ERE-overlapping transcripts to cancer antigenicity has not been captured yet.

Adaptive Immune Responses to Endogenous Retrovirus Proteins in Animal Models

—Despite the continuously evolving relationship between ERVs and the immune system of the host, ERV products can be comparably immunogenic in humans and in animal models, but there are also notable differences. Spontaneous B cell responses to ERV products have long been detected in mice, particularly from autoimmune-prone strains, where they have been implicated in SLE development (152). These autoantibody responses were found to target predominantly the envelope glycoprotein of endogenous MLVs (152). Similarly, B and T cell responses are known to be induced against endogenous MLVs expressed in transplantable cell lines used in murine cancer models (153–156). In the commonly used C57BL/6 mice, these adaptive immune responses target most frequently the envelope glycoprotein of *Emv2*, a single-copy endogenous MLV that induces partial immunological tolerance (157). Although *Emv2* is a replication-defective provirus, recombination with other defective endogenous MLVs can restore its infectivity, resulting in fully infectious, now exogenous retrovirus. Restoration of *Emv2* infectivity was first described in B16 melanoma cells, where it gave rise to the melanoma-associated retrovirus (MelARV) that reinfected these cells multiple times (15, 158). Notably, it was the hunt for the immunodominant target of anti-B16 melanoma antibodies that ultimately lead to the discovery of MelARV (153, 154), suggesting that endogenous MLVs are major cancer antigens. Restoration of infectivity likely contributes significantly to the heightened immunogenicity of endogenous MLVs and has been reported in most transplantable cell lines (16, 18).

Restoration of *Emv2* infectivity, leading to vertical transmission of infectious MLVs, has also been documented in mouse strains with genetic deficiencies affecting antiretrovirus antibody production (17, 33), underscoring the notion that adaptive immune responses to endogenous MLVs are not only possible to induce but are also necessary to control the emergence and transmission of infectious recombinants.

In addition to MLVs, adaptive immune responses against the envelope glycoproteins of endogenous MMTVs have also been documented. For example, T cell responses to EL4 T cell lymphoma cells have long been shown to target epitopes from MMTV envelope glycoprotein (159). Similarly to MLVs, MMTVs can also exist as defective endogenous and infectious exogenous viruses in certain mouse strains, and they have been suggested to amplify their copies in transplantable murine lymphoma cell lines (160). More

recent studies, including those using high-throughput discovery of cancer antigens, have independently identified epitopes from canonical MLV and MMTV proteins that reinforce their immunogenicity. In a comprehensive study of tumor-specific antigens by Laumont et al. (150), the most immunogenic and protective epitope identified in EL4 cells was, in fact, derived from the MMTV envelope glycoprotein. The envelope glycoprotein of *Emv2* has also been validated as tumor antigen in the transplantable GL261 glioma model (161). Moreover, a genetic screen of chromatin regulators that modify immunogenicity of mouse cancer models identified derepressed endogenous retroviruses as the predominant targets in *Setdb1*-deficient cancer cells, with enrichment for peptides from *Emv2* envelope glycoprotein, as well as endogenous MLV Gag and polymerase (162). The recurrent identification of endogenous MLV and MMTV antigens in independent studies points to a major contribution to mouse cancer cell immunogenicity, likely accentuated by the increase in MLV and MMTV copy numbers in transplantable cell lines.

Consequences Of Host Immune Responses To Endogenous Retroelements

As EREs are closely intertwined with the host, immune reactivity against their products could be considered autoimmune in nature and, therefore, has the potential to cause pathology (Figure 4). A pathogenic contribution of ERE-triggered interferon responses is supported by findings in AGS and SLE, as well as in atrophic macular degeneration and age-related inflammation (29, 30, 37, 83). The nucleic acid sensors or metabolic enzymes implicated in the development of these conditions lack the specificity that could discriminate between ERE-derived and other self-nucleic acids (25, 26, 29). Nevertheless, a causative role for EREs is indicated by the use of RT inhibitors. These have shown efficacy in early human AGS trials (37), although results from mouse models of AGS have been conflicting (36, 163). A potential effect of RT inhibitors on SLE would also be important to establish. RT inhibitors also reduce inflammation in aged mice (80), and inspection of health insurance databases indicates that RT inhibitors given as part of pre-exposure prophylaxis to individuals not infected with HIV-1 may lower the risk of atrophic macular degeneration and type 2 diabetes (83, 164). However, proinflammatory effects of pre-exposure prophylaxis have also been reported in genital or gastrointestinal mucosae following topical application of RT inhibitors (165, 166), and in the gastrointestinal mucosa but not the blood following oral administration in individuals not infected with HIV-1 (167). Proinflammatory effects have also been reported in colorectal cancer patients treated with RT inhibitors and in similarly treated colorectal cancer cell lines, particularly those with p53 mutations (168), which are hypothesized to arise from residual ORF2p activity and lead to DNA damage or accumulation of DNA replication intermediates. Collectively, these findings underscore a nonredundant pathogenic potential of RT encoded by either LINE-1 or ERVs, which warrants further investigation into its source and regulation.

A pathogenic contribution of dysregulated innate immunity to EREs does not, however, preclude a beneficial role in physiological conditions (Figure 4). Indeed, innate immune activation by EREs has been suggested to augment the interferon response to exogenous, unrelated viruses, such as herpesviruses (169) or influenza A virus (170), although these viruses can also subvert or inhibit ERE-induced interferon responses. ERE-triggered proinflammatory signaling downstream of MDA5 and RIG-I is thought to be required for

hematopoietic stem and progenitor cell development (171). Activation of the cGAS-STING pathway by EREs has also been suggested to play a major role in promoting the homeostatic adaptive immune response to skin microbiota (38). Therefore, the induction of ERE transcription during infection or colonization with exogenous microbes may function as an intrinsic adjuvant necessary for tuning innate immune reactivity. A beneficial effect of EREs in immune defense from exogenous pathogens has also been described in a mouse model for infection with herpes simplex virus type 2, where the effect was independent of type I interferon signaling (172), suggesting additional mechanisms by which this interaction can manifest at epithelial barriers. Moreover, a large body of evidence suggests a host-protective role for EREs against cancer initiation and progression. Indeed, transcriptional upregulation of EREs in response to mutation of key epigenetic regulators, oncogene activation, cytokines, or natural and drug-induced genomic hypomethylation is associated with potent antitumor effects, attributed to innate immunity induced intrinsically in cancer cells, and increased tumor immunogenicity (39–41, 56, 59, 77, 78, 123, 162, 173–181), although ERE-induced genome instability has also been reported (182).

The link between ERE transcriptional activity and cell-intrinsic innate immune activation may provide tonic and amplifying signals for the innate immune system, as well as the means of effective immunosurveillance of cells damaged by stress, infection, mutation, or transformation (183, 184). Disruption of such physiological processes may, in turn, explain the pathogenic consequences of ERE activation. For example, excessive ERE activation promoted by a high-fat diet can turn homeostatic responses to the microbiota into inflammatory responses (38). Similarly, age-related inflammation, SLE, and AGS may represent the spectrum of ERE dysregulation. Lastly, ERE upregulation in cancer may also be detrimental, particularly when the interferon response pathways are disrupted (180, 185), which may also explain increased cancer susceptibility in old age, despite increased ERE activity (82).

A pathogenic potential of adaptive immune responses to EREs can also be envisioned. Although their titers were not elevated, antibodies reactive with HERV-K(HML-2) envelope glycoproteins were found to cause neutrophil activation in SLE patients (119). Antibodies to HERV and LINE-1 products have frequently been observed in autoimmune disorders, but also in otherwise healthy individuals, albeit often less frequently (106, 135, 172). Consequently, their contribution to disease initiation and progression remains unclear (186). Germline ERVs, as well as maternally transmitted infectious counterparts, induce a degree of T cell tolerance in mouse models (157, 187). However, B cells appear to escape deletional tolerance and ultimately mount an anti-ERV antibody response in the offspring that protects further vertical transmission of infectious viruses, without pathological signs (187). Moreover, induction of adaptive immune responses against LINE-1 ORF2p and HERV-K(HML-2) Gag and envelope proteins by vaccination was not associated with adverse or pathological findings (188). Adaptive immune responses against ERV-encoded antigens can afford protection against cancer, both in mouse models and in humans (189). However, whether such responses are additionally associated with paraneoplastic autoimmunity or adverse effects of checkpoint blockade remains an important question.

Concluding Remarks And Broader Perspective

While both innate and adaptive responses to diverse ERE products can be clearly elicited, the overall consequences for the host are context dependent. For a powerful defense weapon such as the adaptive immune system, its response to a given stimulus may range from host protection to immune pathology or autoimmunity (190), and responses to EREs seem to follow the same rules. Equally, innate immune activation by EREs may contribute to protection from cancer but also promote age-related inflammation (30). Such antagonistic pleiotropy has striking parallels with dichotomous regulation of cancer risk and lifespan by p53 (191), or cancer risk and regenerative ability (192).

Participation of EREs in multiple biological processes implies that evolutionary adaptation will be slower, but persistence in the host will be more likely. The adaptation of EREs and host immunity will naturally integrate stronger evolutionary pressure on the host immune system from a multitude of exogenous pathogens, as well as ERE effects on host biological processes other than the immune system. Fitness trade-offs associated with ERE activity in the host can manifest at many levels, but arguably, a major battleground is genome function (7, 8, 79, 183, 193). Indeed, genomic features of EREs, together with their abundance, provide the genetic diversity for evolution of new gene function, including of immune genes.

A growing list of examples comprises the interferon inducibility provided by *MER41* LTR elements to several genes, including the cytoplasmic DNA sensor *AIM2* (194), and the enhancer activity provided by an *ERV1* LTR promoting HLA-G expression in human extravillous trophoblasts at the fetal-maternal interface (195). ERE-driven innate immune activation of dendritic cells lacking TRIM28 is linked to transcriptional regulation of nearby immune genes, rather than ligation of innate sensors (196), and LINE-1-containing transcript isoforms of immune genes regulate naive T cell quiescence (197). Alternative splicing events exonizing EREs can create isoforms of immune gene transcripts with novel function, such as the soluble form of PD-L1 acting as a receptor antagonist (198), and polymorphic *Alu* integrations in intronic regions can cause skipping of nearby exons, affecting the function of immune genes such as *CD58*, which encodes lymphocyte function-associated antigen 3 (LFA-3) (199).

While major effects of EREs in immune gene function are likely selected over long evolutionary periods, the regulation of ERE activity, particularly in healthy cells, and the potential effects of EREs on gene function are largely unexplored. Great progress is being made in defining ERE-derived ligands and respective innate sensors and adaptive receptors, triggering of which leads to immune activation in health and disease. However, our understanding of the full interaction of EREs with the host immune system will only be completed with integration of their roles as immune gene regulators, as well as immune activators.

Acknowledgements

This work was supported by The Francis Crick Institute (FC001099), which receives its core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement number 101018670).

References

1. Brodsky FM. Stealth, sabotage and exploitation. *Immunol Rev.* 1999; 168: 5–11. [PubMed: 10399060]
2. Damian RT. Molecular mimicry: antigen sharing by parasite and host and its consequences. *Am Nat.* 1964; 98: 129–49.
3. Rojas M, Restrepo-Jiménez P, Monsalve DM, Pacheco Y, Acosta-Ampudia Y, et al. Molecular mimicry and autoimmunity. *J Autoimmun.* 2018; 95: 100–23. [PubMed: 30509385]
4. Alcami A. Viral mimicry of cytokines, chemokines and their receptors. *Nat Rev Immunol.* 2003; 3: 36–50. [PubMed: 12511874]
5. Hoffmann HH, Schneider WM, Rice CM. Interferons and viruses: an evolutionary arms race of molecular interactions. *Trends Immunol.* 2015; 36: 124–38. [PubMed: 25704559]
6. García-Sastre A. Ten strategies of interferon evasion by viruses. *Cell Host Microbe.* 2017; 22: 176–84. [PubMed: 28799903]
7. Rebollo R, Romanish MT, Mager DL. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet.* 2012; 46: 21–42. [PubMed: 22905872]
8. Goubert C, Zevallos NA, Feschotte C. Contribution of unfixed transposable element insertions to human regulatory variation. *Philos Trans R Soc Lond B.* 2020; 375 20190331 [PubMed: 32075552]
9. Wells JN, Feschotte C. A field guide to eukaryotic transposable elements. *Annu Rev Genet.* 2020; 54: 539–61. [PubMed: 32955944]
10. Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, et al. Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature.* 2010; 463: 84–87. [PubMed: 20054395]
11. Magiorkinis G, Gifford RJ, Katzourakis A, De Ranter Belshaw R. Env-less endogenous retroviruses are genomic superspreaders. *PNAS.* 2012; 109: 7385–90. [PubMed: 22529376]
12. Kassiotis G. Endogenous retroviruses and the development of cancer. *J Immunol.* 2014; 192: 1343–49. [PubMed: 24511094]
13. Stocking C, Kozak CA. Murine endogenous retroviruses. *Cell Mol Life Sci.* 2008; 65: 3383–98. [PubMed: 18818872]
14. Li M, Huang X, Zhu Z, Gorelik E. Sequence and insertion sites of murine melanoma-associated retrovirus. *J Virol.* 1999; 73: 9178–86. [PubMed: 10516025]
15. Pothlichet J, Mangeney M, Heidmann T. Mobility and integration sites of a murine C57BL/6 melanoma endogenous retrovirus involved in tumor progression in vivo. *Int J Cancer.* 2006; 119: 1869–77. [PubMed: 16708391]
16. Pothlichet J, Heidmann T, Mangeney M. A recombinant endogenous retrovirus amplified in a mouse neuroblastoma is involved in tumor growth in vivo. *Int J Cancer.* 2006; 119: 815–22. [PubMed: 16550601]
17. Young GR, Eksmond U, Salcedo R, Alexopoulou L, Stoye JP, Kassiotis G. Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature.* 2012; 491: 774–78. [PubMed: 23103862]
18. Ottina E, Levy P, Eksmond U, Merckenschlager J, Young GR, et al. Restoration of endogenous retrovirus infectivity impacts mouse cancer models. *Cancer Immunol Res.* 2018; 6: 1292–300. [PubMed: 30143537]
19. Bannert N, Kurth R. The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genom Hum Genet.* 2006; 7: 149–73.
20. Holloway JR, Williams ZH, Freeman MM, Bulow U, Coffin JM. Gorillas have been infected with the HERV-K (HML-2) endogenous retrovirus much more recently than humans and chimpanzees. *PNAS.* 2019; 116: 1337–46. [PubMed: 30610173]
21. Monde K, Satou Y, Goto M, Uchiyama Y, Ito J, et al. Movements of ancient human endogenous retroviruses detected in SOX2-expressing cells. *J Virol.* 2022; 96 (9) e0035622 [PubMed: 35420440]
22. Smit AF. Identification of a new, abundant superfamily of mammalian LTR-transposons. *Nucleic Acids Res.* 1993; 21: 1863–72. [PubMed: 8388099]

23. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBOJ*. 1990; 9: 3353–62.
24. Richardson SR, Doucet AJ, Kopera HC, Moldovan JB, Garcia-Perez JL, Moran JV. The influence of LINE-1 and SINE retrotransposons on mammalian genomes. *Microbiol Spectr*. 2015; 3: MDNA3-0061-2014
25. Medzhitov R. Toll-like receptors and innate immunity. *Nat Rev Immunol*. 2001; 1: 135–45. [PubMed: 11905821]
26. Wu J, Chen ZJ. Innate immune sensing and signaling of cytosolic nucleic acids. *Annu Rev Immunol*. 2014; 32: 461–88. [PubMed: 24655297]
27. Nozaki K, Li L, Miao EA. Innate sensors trigger regulated cell death to combat intracellular infection. *Annu Rev Immunol*. 2022; 40: 469–98. [PubMed: 35138947]
28. Miller KN, Victorelli SG, Salmonowicz H, Dasgupta N, Liu T, et al. Cytoplasmic DNA: sources, sensing, and role in aging and disease. *Cell*. 2021; 184: 5506–26. [PubMed: 34715021]
29. Crowl JT, Gray EE, Pestal K, Volkman HE, Stetson DB. Intracellular nucleic acid detection in autoimmunity. *Annu Rev Immunol*. 2017; 35: 313–36. [PubMed: 28142323]
30. Gorbunova V, Seluanov A, Mita P, McKerrow W, Fenyö D, et al. The role of retrotransposable elements in ageing and age-associated diseases. *Nature*. 2021; 596: 43–53. [PubMed: 34349292]
31. Ishak CA, De Carvalho DD. Reactivation of endogenous retroelements in cancer development and therapy. *Annu Rev Cancer Biol*. 2020; 4: 159–76.
32. Lee HK, Lund JM, Ramanathan B, Mizushima N, Iwasaki A. Autophagy-dependent viral recognition by plasmacytoid dendritic cells. *Science*. 2007; 315: 1398–401. [PubMed: 17272685]
33. Yu P, Lübben W, Slomka H, Gebler J, Konert M, et al. Nucleic acid-sensing Toll-like receptors are essential for the control of endogenous retrovirus viremia and ERV-induced tumors. *Immunity*. 2012; 37: 867–79. [PubMed: 23142781]
34. Dembny P, Newman AG, Singh M, Hinz M, Szczepek M, et al. Human endogenous retrovirus HERV-K(HML-2) RNA causes neurodegeneration through Toll-like receptors. *JCI Insight*. 2020; 5: e131093 [PubMed: 32271161]
35. Sinibaldi-Vallebona P, Lavia P, Garaci E, Spadafora C. A role for endogenous reverse transcriptase in tumorigenesis and as a target in differentiating cancer therapy. *Genes Chromosomes Cancer*. 2006; 45: 1–10. [PubMed: 16175572]
36. Beck-Engeser GB, Eilat D, Wabl M. An autoimmune disease prevented by anti-retroviral drugs. *Retrovirology*. 2011; 8: 91. [PubMed: 22067273]
37. Rice GI, Meyzer C, Bouazza N, Hully M, Boddart N, et al. Reverse-transcriptase inhibitors in the Aicardi-Goutières syndrome. *N Engl J Med*. 2018; 379: 2275–77. [PubMed: 30566312]
38. Lima-Junior DS, Krishnamurthy SR, Bouladoux N, Collins N, Han SJ, et al. Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. *Cell*. 2021; 184: 3794–811. e19 [PubMed: 34166614]
39. Canadas I, Thummalapalli R, Kim JW, Kitajima S, Jenkins RW, et al. Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. *Nat Med*. 2018; 24 (8) 1143–50. [PubMed: 30038220]
40. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell*. 2015; 162: 974–86. [PubMed: 26317466]
41. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, et al. DNA-Demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell*. 2015; 162: 961–73. [PubMed: 26317465]
42. Tavora B, Mederer T, Wessel KJ, Ruffing S, Sadjadi M, et al. Tumoural activation of TLR3-SLIT2 axis in endothelium drives metastasis. *Nature*. 2020; 586: 299–304. [PubMed: 32999457]
43. Magin C, Löwer R, Löwer J. cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. *J Virol*. 1999; 73: 9496–507. [PubMed: 10516058]
44. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*. 2015; 522: 221–25. [PubMed: 25896322]

45. Wang-Johanning F, Rycaj K, Plummer JB, Li M, Yin B, et al. Immunotherapeutic potential of anti-human endogenous retrovirus-K envelope protein antibodies in targeting breast tumors. *J Natl Cancer Inst.* 2012; 104: 189–210. [PubMed: 22247020]
46. LiM Radvanyi L, Yin B, Rycaj K, Li J, et al. Downregulation of human endogenous retrovirus type K (HERV-K) viral env RNA in pancreatic cancer cells decreases cell proliferation and tumor growth. *Clin Cancer Res.* 2017; 23: 5892–911. [PubMed: 28679769]
47. Lemaitre C, Tsang J, Bireau C, Heidmann T, Dewannieux M. A human endogenous retrovirus-derived gene that can contribute to oncogenesis by activating the ERK pathway and inducing migration and invasion. *PLOS Pathog.* 2017; 13 e1006451 [PubMed: 28651004]
48. Panova V, Attig J, Young GR, Stoye JP, Kassiotis G. Antibody-induced internalisation of retroviral envelope glycoproteins is a signal initiation event. *PLOS Pathog.* 2020; 16 e1008605 [PubMed: 32453763]
49. Lavialle C, Cornelis G, Dupressoir A, Esnault C, Heidmann O, et al. Paleovirology of ‘syncytins’, retroviral env genes exapted for a role in placentation. *Philos Trans R Soc Lond B.* 2013; 368 20120507 [PubMed: 23938756]
50. Antony JM, van Marle G, Opii W, Butterfield DA, Mallet F, et al. Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination. *Nat Neurosci.* 2004; 7: 1088–95. [PubMed: 15452578]
51. Li W, Lee MH, Henderson L, Tyagi R, Bachani M, et al. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci Transl Med.* 2015; 7 307ra153
52. Jönsson ME, Garza R, Sharma Y, Petri R, Södersten E, et al. Activation of endogenous retroviruses during brain development causes an inflammatory response. *EMBO J.* 2021; 40 e106423 [PubMed: 33644903]
53. Pasquarella A, Ebert A, Pereira de Almeida G, Hinterberger M, Kazerani M, et al. Retrotransposon derepression leads to activation of the unfolded protein response and apoptosis in pro-B cells. *Development.* 2016; 143: 1788–99. [PubMed: 27013243]
54. Green NM, Moody KS, Debatis M, Marshak-Rothstein A. Activation of autoreactive B cells by endogenous TLR7 and TLR3 RNA ligands. *J Biol Chem.* 2012; 287: 39789–99. [PubMed: 23019335]
55. Hung T, Pratt GA, Sundararaman B, Townsend MJ, Chaivorapol C, et al. The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. *Science.* 2015; 350: 455–59. [PubMed: 26382853]
56. Cuellar TL, Herzner AM, Zhang X, Goyal Y, Watanabe C, et al. Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. *J Cell Biol.* 2017; 216: 3535–49. [PubMed: 28887438]
57. Tunbak H, Enriquez-Gasca R, Tie CHC, Gould PA, Mlcochova P, et al. The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. *Nat Commun.* 2020; 11 5387 [PubMed: 33144593]
58. Leonova KI, Brodsky L, Lipchick B, Pal M, Novototskaya L, et al. p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing of repeats and noncoding RNAs. *PNAS.* 2013; 110: E89–98. [PubMed: 23236145]
59. Sheng W, LaFleur MW, Nguyen TH, Chen S, Chakravarthy A, et al. LSD1 ablation stimulates anti-tumor immunity and enables checkpoint blockade. *Cell.* 2018; 174 (3) 549–63. e19 [PubMed: 29937226]
60. Mannion NM, Greenwood SM, Young R, Cox S, Brindle J, et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep.* 2014; 9: 1482–94. [PubMed: 25456137]
61. Pestal K, Funk CC, Snyder JM, Price ND, Treuting PM, Stetson DB. Isoforms of RNA-editing enzyme ADAR1 independently control nucleic acid sensor MDA5-driven autoimmunity and multi-organ development. *Immunity.* 2015; 43: 933–44. [PubMed: 26588779]
62. Liddicoat BJ, Piskol R, Chalk AM, Ramaswami G, Higuchi M, et al. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science.* 2015; 349: 1115–20. [PubMed: 26275108]

63. Nishikura K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat Rev Mol Cell Biol.* 2016; 17: 83–96. [PubMed: 26648264]
64. Baker AR, Slack FJ. ADAR1 and its implications in cancer development and treatment. *Trends Genet.* 2022; 38 (8) 821–30. [PubMed: 35459560]
65. Ahmad S, Mu X, Yang F, Greenwald E, Park JW, et al. Breaching self-tolerance to Alu duplex RNA underlies MDA5-mediated inflammation. *Cell.* 2018; 172: 797–810. e13 [PubMed: 29395326]
66. Chung H, Calis JJA, Wu X, Sun T, Yu Y, et al. Human ADAR1 prevents endogenous RNA from triggering translational shutdown. *Cell.* 2018; 172: 811–24. e14 [PubMed: 29395325]
67. Mehdipour P, Marhon SA, Ettayebi I, Chakravarthy A, Hosseini A, et al. Epigenetic therapy induces transcription of inverted SINES and ADAR1 dependency. *Nature.* 2020; 588: 169–73. [PubMed: 33087935]
68. Jiao H, Wachsmuth L, Kumari S, Schwarzer R, Lin J, et al. Z-nucleic-acid sensing triggers ZBP1-dependent necroptosis and inflammation. *Nature.* 2020; 580: 391–95. [PubMed: 32296175]
69. Wang R, Li H, Wu J, Cai Z-Y, Li B, et al. Gut stem cell necroptosis by genome instability triggers bowel inflammation. *Nature.* 2020; 580: 386–90. [PubMed: 32296174]
70. Devos M, Tanghe G, Gilbert B, Dierick E, Verheirstraeten M, et al. Sensing of endogenous nucleic acids by ZBP1 induces keratinocyte necroptosis and skin inflammation. *J Exp Med.* 2020; 217 e20191913 [PubMed: 32315377]
71. Karki R, Sundaram B, Sharma BR, Lee S, Malireddi RKS, et al. ADAR1 restricts ZBP1-mediated immune response and PANoptosis to promote tumorigenesis. *Cell Rep.* 2021; 37 109858 [PubMed: 34686350]
72. Zhang T, Yin C, Fedorov A, Qiao L, Bao H, et al. ADAR1 masks the cancer immunotherapeutic promise of ZBP1-driven necroptosis. *Nature.* 2022; 606 (7914) 594–602. [PubMed: 35614224]
73. Jiao H, Wachsmuth L, Wolf S, Lohmann J, Nagata M, et al. ADAR1 averts fatal type I interferon induction by ZBP1. *Nature.* 2022; 607: 776–83. [PubMed: 35859176]
74. de Reuver R, Verdonck S, Dierick E, Nemegeer J, Hessmann E, et al. ADAR1 prevents autoinflammation by suppressing spontaneous ZBP1 activation. *Nature.* 2022; 607: 784–89. [PubMed: 35859175]
75. Hubbard NW, Ames JM, Maurano M, Chu LH, Somfleth KY, et al. ADAR1 mutation causes ZBP1-dependent immunopathology. *Nature.* 2022; 607: 769–75. [PubMed: 35859177]
76. Tang Q, Rigby RE, Young GR, Hvidt AK, Davis T, et al. Adenosine-to-inosine editing of endogenous Z-form RNA by the deaminase ADAR1 prevents spontaneous MAVS-dependent type I interferon responses. *Immunity.* 2021; 54: 1961–75. e5 [PubMed: 34525337]
77. Ishizuka JJ, Manguso RT, Cheruiyot CK, Bi K, Panda A, et al. Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature.* 2019; 565: 43–48. [PubMed: 30559380]
78. Liu H, Golji J, Brodeur LK, Chung FS, Chen JT, et al. Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nat Med.* 2019; 25: 95–102. [PubMed: 30559422]
79. Kazazian HH Jr, Moran JV. Mobile DNA in health and disease. *N Engl J Med.* 2017; 377: 361–70. [PubMed: 28745987]
80. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature.* 2019; 566: 73–78. [PubMed: 30728521]
81. Simon M, Van Meter M, Ablaeva J, Ke Z, Gonzalez RS, et al. LINE1 derepression in aged wild-type and SIRT6-deficient mice drives inflammation. *Cell Metab.* 2019; 29: 871–85. e5 [PubMed: 30853213]
82. Zhao Y, Oreskovic E, Zhang Q, Lu Q, Gilman A, et al. Transposon-triggered innate immune response confers cancer resistance to the blind mole rat. *Nat Immunol.* 2021; 22: 1219–30. [PubMed: 34556881]
83. Fukuda S, Varshney A, Fowler BJ, Wang SB, Narendran S, et al. Cytoplasmic synthesis of endogenous Alu complementary DNA via reverse transcription and implications in age-related macular degeneration. *PNAS.* 2021; 118 e2022751118 [PubMed: 33526699]

84. Kines KJ, Sokolowski M, deHaro DL, Christian CM, Belancio VP. Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci. *Nucleic Acids Res.* 2014; 42: 10488–502. [PubMed: 25143528]
85. Li T, Chen ZJ. The cGAS-cGAMP-STING pathway connects DNA damage to inflammation, senescence, and cancer. *J Exp Med.* 2018; 215: 1287–99. [PubMed: 29622565]
86. Simpson E, Takacs K, Altmann DM. Thymic repertoire selection by superantigens: presentation by human and mouse MHC molecules. *Thymus.* 1994; 23: 1–13. [PubMed: 7863543]
87. Acha-Orbea H, MacDonald HR. Superantigens of mouse mammary tumor virus. *Annu Rev Immunol.* 1995; 13: 459–86. [PubMed: 7612231]
88. Sutkowski N, Conrad B, Thorley-Lawson DA, Huber BT. Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. *Immunity.* 2001; 15: 579–89. [PubMed: 11672540]
89. Papiernik M. Natural CD4+ CD25+ regulatory T cells. Their role in the control of superantigen responses. *Immunol Rev.* 2001; 182: 180–89. [PubMed: 11722633]
90. Ribot J, Romagnoli P, van Meerwijk JP. Agonist ligands expressed by thymic epithelium enhance positive selection of regulatory T lymphocytes from precursors with a normally diverse TCR repertoire. *J Immunol.* 2006; 177: 1101–7. [PubMed: 16818767]
91. Myers L, Joedicke JJ, Carmody AB, Messer RJ, Kassiotis G, et al. IL-2-independent and TNF- α -dependent expansion of V β 5⁺ natural regulatory T cells during retrovirus infection. *J Immunol.* 2013; 190: 5485–95. [PubMed: 23645880]
92. Punkosdy GA, Blain M, Glass DD, Lozano MM, O'Mara L, et al. Regulatory T-cell expansion during chronic viral infection is dependent on endogenous retroviral superantigens. *PNAS.* 2011; 108: 3677–82. [PubMed: 21321220]
93. Kalter SS, Helmke RJ, Heberling RL, Panigel M, Fowler AK, et al. Brief communication: C-type particles in normal human placentas. *J Natl Cancer Inst.* 1973; 50: 1081–84. [PubMed: 4703773]
94. Dirksen ER, Levy JA. Virus-like particles in placentas from normal individuals and patients with systemic lupus erythematosus. *J Natl Cancer Inst.* 1977; 59: 1187–92. [PubMed: 903994]
95. Nelson J, Leong JA, Levy JA. Normal human placentas contain RNA-directed DNA polymerase activity like that in viruses. *PNAS.* 1978; 75: 6263–67. [PubMed: 83652]
96. Simpson GR, Patience C, Löwer R, Tönjes RR, Moore HD, et al. Endogenous D-type (HERV-K) related sequences are packaged into retroviral particles in the placenta and possess open reading frames for reverse transcriptase. *Virology.* 1996; 222: 451–56. [PubMed: 8806530]
97. Löwer R, Boller K, Hasenmaier B, Korbmacher C, Müller-Lantzsch N, et al. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. *PNAS.* 1993; 90: 4480–84. [PubMed: 8506289]
98. Contreras-Galindo R, Kaplan MH, Leissner P, Verjat T, Ferlenghi I, et al. Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J Virol.* 2008; 82: 9329–36. [PubMed: 18632860]
99. Contreras-Galindo R, Kaplan MH, Dube D, Gonzalez-Hernandez MJ, Chan S, et al. Human endogenous retrovirus type K (HERV-K) particles package and transmit HERV-K-related sequences. *J Virol.* 2015; 89: 7187–201. [PubMed: 25926654]
100. Larouche JD, Trofimov A, Hesnard L, Ehx G, Zhao Q, et al. Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. *Genome Med.* 2020; 12: 40. [PubMed: 32345368]
101. Andersson AC, Yun Z, Sperber GO, Larsson E, Blomberg J. ERV3 and related sequences in humans: structure and RNA expression. *J Virol.* 2005; 79: 9270–84. [PubMed: 15994821]
102. Bustamante Rivera YY, Brütting C, Schmidt C, Volkmer I, Staeger MS. Endogenous retrovirus 3—history, physiology, and pathology. *Front Microbiol.* 2017; 8: 2691 [PubMed: 29379485]
103. de Parseval N, Forrest G, Venables PJ, Heidmann T. ERV-3 envelope expression and congenital heart block: What does a physiological knockout teach us. *Autoimmunity.* 1999; 30: 81–83. [PubMed: 10435720]
104. de Parseval N, Heidmann T. Physiological knockout of the envelope gene of the single-copy ERV-3 human endogenous retrovirus in a fraction of the Caucasian population. *J Virol.* 1998; 72: 3442–45. [PubMed: 9525678]

105. Li JM, Fan WS, Horsfall AC, Anderson AC, Rigby S, et al. The expression of human endogenous retrovirus-3 in fetal cardiac tissue and antibodies in congenital heart block. *Clin Exp Immunol.* 1996; 104: 388–93. [PubMed: 9099920]
106. Deakin CT, Cornish GH, Ng KW, Faulkner N, Bolland W, et al. Favorable antibody responses to human coronaviruses in children and adolescents with autoimmune rheumatic diseases. *Med.* 2021; 2: 1093–109. e6 [PubMed: 34414384]
107. Rakoff-Nahoum S, Kuebler PJ, Heymann JJ, Sheehy ME, Ortiz GM, et al. Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. *AIDS Res Hum Retroviruses.* 2006; 22: 52–56. [PubMed: 16438646]
108. Bonaventura P, Alcazer V, Mutez V, Tonon L, Martin J, et al. Identification of shared tumor epitopes from endogenous retroviruses inducing high-avidity cytotoxic T cells for cancer immunotherapy. *Sci Adv.* 2022; 8 eabj3671 [PubMed: 35080970]
109. Wang-Johanning F, Radvanyi L, Rycaj K, Plummer JB, Yan P, et al. Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. *Cancer Res.* 2008; 68: 5869–77. [PubMed: 18632641]
110. Sauter M, Schommer S, Kremmer E, Remberger K, Dolken G, et al. Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas. *J Virol.* 1995; 69: 414–21. [PubMed: 7983737]
111. Sauter M, Roemer K, Best B, Afting M, Schommer S, et al. Specificity of antibodies directed against Env protein of human endogenous retroviruses in patients with germ cell tumors. *Cancer Res.* 1996; 56: 4362–65. [PubMed: 8813125]
112. Kleiman A, Senyuta N, Tryakin A, Sauter M, Karseladze A, et al. HERV-K(HML-2) GAG/ENV antibodies as indicator for therapy effect in patients with germ cell tumors. *Int J Cancer.* 2004; 110: 459–61. [PubMed: 15095315]
113. Ishida T, Obata Y, Ohara N, Matsushita H, Sato S, et al. Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. *Cancer Immun.* 2008; 8: 15. [PubMed: 19006261]
114. Humer J, Waltenberger A, Grassauer A, Kurz M, Valencak J, et al. Identification of a melanoma marker derived from melanoma-associated endogenous retroviruses. *Cancer Res.* 2006; 66: 1658–63. [PubMed: 16452225]
115. Hahn S, Ugurel S, Hanschmann KM, Strobel H, Tondera C, et al. Serological response to human endogenous retrovirus K in melanoma patients correlates with survival probability. *AIDS Res Hum Retroviruses.* 2008; 24: 717–23. [PubMed: 18462078]
116. Buscher K, Trefzer U, Hofmann M, Sterry W, Kurth R, Denner J. Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer Res.* 2005; 65: 4172–80. [PubMed: 15899808]
117. Boller K, Janssen O, Schuldes H, Tonjes RR, Kurth R. Characterization of the antibody response specific for the human endogenous retrovirus HTDV/HERV-K. *J Virol.* 1997; 71: 4581–88. [PubMed: 9151852]
118. Gupta R, Michaud HA, Zeng X, Debbaneh M, Arron ST, et al. Diminished humoral responses against and reduced gene expression levels of human endogenous retrovirus-K (HERV-K) in psoriasis. *J Transl Med.* 2014; 12: 256. [PubMed: 25224121]
119. Tokuyama M, Gunn BM, Venkataraman A, Kong Y, Kang I, et al. Antibodies against human endogenous retrovirus K102 envelope activate neutrophils in systemic lupus erythematosus. *J Exp Med.* 2021; 218 e20191766 [PubMed: 34019642]
120. Takahashi Y, Harashima N, Kajigaya S, Yokoyama H, Cherkasova E, et al. Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells. *J Clin Investig.* 2008; 118: 1099–109. [PubMed: 18292810]
121. Cherkasova E, Weisman Q, Childs RW. Endogenous retroviruses as targets for antitumor immunity in renal cell cancer and other tumors. *Front Oncol.* 2013; 3: 243. [PubMed: 24062992]
122. Cherkasova E, Scrivani C, Doh S, Weisman Q, Takahashi Y, et al. Detection of an immunogenic HERV-E envelope with selective expression in clear cell kidney cancer. *Cancer Res.* 2016; 76: 2177–85. [PubMed: 26862115]

123. Smith CC, Beckermann KE, Bortone DS, de Cubas AA, Bixby LM, et al. Endogenous retroviral signatures predict immunotherapy response in clear cell renal cell carcinoma. *J Clin Investig*. 2018; 128 (11) 4804–20. [PubMed: 30137025]
124. Gemmell P, Hein J, Katzourakis A. Phylogenetic analysis reveals that ERVs “die young” but HERV-H is unusually conserved. *PLOS Comput Biol*. 2016; 12 e1004964 [PubMed: 27295277]
125. Izsvák Z, Wang J, Singh M, Mager DL, Hurst LD. Pluripotency and the endogenous retrovirus HERVH: conflict or serendipity? *BioEssays*. 2016; 38: 109–17. [PubMed: 26735931]
126. Mullins CS, Linnebacher M. Endogenous retrovirus sequences as a novel class of tumor-specific antigens: an example of HERV-H env encoding strong CTL epitopes. *Cancer Immunol Immunother*. 2012; 61: 1093–100. [PubMed: 22187063]
127. Wentzensen N, Coy JF, Knaebel HP, Linnebacher M, Wilz B, et al. Expression of an endogenous retroviral sequence from the HERV-H group in gastrointestinal cancers. *Int J Cancer*. 2007; 121: 1417–23. [PubMed: 17546591]
128. Brudek T, Christensen T, Aagaard L, Petersen T, Hansen H, Moller-Larsen A. B cells and monocytes from patients with active multiple sclerosis exhibit increased surface expression of both HERV-H Env and HERV-W Env, accompanied by increased seroreactivity. *Retrovirology*. 2009; 6: 104. [PubMed: 19917105]
129. Mameli G, Cossu D, Cocco E, Frau J, Marrosu MG, et al. Epitopes of HERV-Wenv induce antigen-specific humoral immunity in multiple sclerosis patients. *J Neuroimmunol*. 2015; 280: 66–68. [PubMed: 25773158]
130. Hon GM, Erasmus RT, Matsha T. Multiple sclerosis-associated retrovirus and related human endogenous retrovirus-W in patients with multiple sclerosis: a literature review. *J Neuroimmunol*. 2013; 263: 8–12. [PubMed: 23993654]
131. Ruprecht K, Gronen F, Sauter M, Best B, Rieckmann P, Mueller-Lantzsch N. Lack of immune responses against multiple sclerosis-associated retrovirus/human endogenous retrovirus W in patients with multiple sclerosis. *J Neurovirol*. 2008; 14: 143–51. [PubMed: 18444086]
132. Saini SK, Ørskov AD, Bjerregaard AM, Unnikrishnan A, Holmberg-Thyden S, et al. Human endogenous retroviruses form a reservoir of T cell targets in hematological cancers. *Nat Commun*. 2020; 11 5660 [PubMed: 33168830]
133. Kong Y, Rose CM, Cass AA, Williams AG, Darwish M, et al. Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. *Nat Commun*. 2019; 10 5228 [PubMed: 31745090]
134. Bonté P-E, Arribas YA, Merlotti A, Carrascal M, Zhang JV, et al. Single-cell RNA-seq-based proteogenomics identifies glioblastoma-specific transposable elements encoding HLA-I-presented peptides. *Cell Rep*. 2022; 39 110916 [PubMed: 35675780]
135. Carter V, LaCava J, Taylor MS, Liang SY, Mustelin C, et al. High prevalence and disease correlation of autoantibodies against p40 encoded by long interspersed nuclear elements in systemic lupus erythematosus. *Arthritis Rheumatol*. 2020; 72: 89–99. [PubMed: 31342656]
136. Starck SR, Tsai JC, Chen K, Shodiya M, Wang L, et al. Translation from the 5' untranslated region shapes the integrated stress response. *Science*. 2016; 351 aad3867 [PubMed: 26823435]
137. Laumont CM, Daouda T, Laverdure JP, Bonnel E, Caron-Lizotte O, et al. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun*. 2016; 7 10238 [PubMed: 26728094]
138. Zhu Y, Orre LM, Johansson HJ, Huss M, Boekel J, et al. Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow. *Nat Commun*. 2018; 9: 903. [PubMed: 29500430]
139. Erhard F, Dölken L, Schilling B, Schlosser A. Identification of the cryptic HLA-I immunopeptidome. *Cancer Immunol Res*. 2020; 8: 1018–26. [PubMed: 32561536]
140. Calviello L, Hirsekorn A, Ohler U. Quantification of translation uncovers the functions of the alternative transcriptome. *Nat Struct Mol Biol*. 2020; 27: 717–25. [PubMed: 32601440]
141. Chen J, Brunner AD, Cogan JZ, Nunez JK, Fields AP, et al. Pervasive functional translation of noncanonical human open reading frames. *Science*. 2020; 367: 1140–46. [PubMed: 32139545]

142. Ouspenskaia T, Law T, Clauser KR, Klaeger S, Sarkizova S, et al. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat Biotechnol.* 2022; 40: 209–17. [PubMed: 34663921]
143. Kahles A, Lehmann KV, Toussaint NC, Huser M, Stark SG, et al. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell.* 2018; 34: 211–24. e6 [PubMed: 30078747]
144. Zhao Q, Laverdure JP, Lanoix J, Durette C, Côté C, et al. Proteogenomics uncovers a vast repertoire of shared tumor-specific antigens in ovarian cancer. *Cancer Immunol Res.* 2020; 8: 544–55. [PubMed: 32047025]
145. Ehx G, Larouche JD, Durette C, Laverdure JP, Hesnard L, et al. Atypical acute myeloid leukemiaspecific transcripts generate shared and immunogenic MHC class-I-associated epitopes. *Immunity.* 2021; 54: 737–52. e10 [PubMed: 33740418]
146. Bigot J, Lalanne AI, Lucibello F, Gueguen P, Houy A, et al. Splicing patterns in SF3B1-mutated uveal melanoma generate shared immunogenic tumor-specific neoepitopes. *Cancer Discov.* 2021; 11: 1938–51. [PubMed: 33811047]
147. Attig J, Young GR, Hosie L, Perkins D, Encheva-Yokoya V, et al. LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome Res.* 2019; 29: 1578–90. [PubMed: 31537638]
148. Jang HS, Shah NM, Du AY, Dailey ZZ, Pehrsson EC, et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat Genet.* 2019; 51: 611–17. [PubMed: 30926969]
149. Kazachenka A, Young GR, Attig J, Kordella C, Lamprianidou E, et al. Epigenetic therapy of myelodysplastic syndromes connects to cellular differentiation independently of endogenous retroelement derepression. *Genome Med.* 2019; 11: 86. [PubMed: 31870430]
150. Laumont CM, Vincent K, Hesnard L, Audemard É, Bonneil É, et al. Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med.* 2018; 10 eaau5516 [PubMed: 30518613]
151. Chong C, Müller M, Pak H, Harnett D, Huber F, et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun.* 2020; 11 1293 [PubMed: 32157095]
152. Baudino L, Yoshinobu K, Morito N, Santiago-Raber M-L, Izui S. Role of endogenous retroviruses in murine SLE. *Autoimmun Rev.* 2010; 10: 27–34. [PubMed: 20659589]
153. Rappaport I, Alterman AL, Braverman S, Stackpole CW. Syngeneic monoclonal antibodies to B16 melanoma viral antigens. *Cancer Res.* 1987; 47: 5391–96. [PubMed: 3652042]
154. Leong SP, Muller J, Yetter RA, Gorelik E, Takami T, Hearing VJ. Expression and modulation of a retrovirus-associated antigen by murine melanoma cells. *Cancer Res.* 1988; 48: 4954–58. [PubMed: 2842042]
155. Huang AY, Gulden PH, Woods AS, Thomas MC, Tong CD, et al. The immunodominant major histocompatibility complex class I-restricted antigen of a murine colon tumor derives from an endogenous retroviral gene product. *PNAS.* 1996; 93: 9730–35. [PubMed: 8790399]
156. Kershaw MH, Hsu C, Mondesire W, Parker LL, Wang G, et al. Immunization against endogenous retroviral tumor-associated antigens. *Cancer Res.* 2001; 61: 7920–24. [PubMed: 11691813]
157. Young GR, Ploquin MJ, Eksmond U, Wadwa M, Stoye JP, Kassiotis G. Negative selection by an endogenous retrovirus promotes a higher-avidity CD4⁺ T cell response to retroviral infection. *PLOS Pathog.* 2012; 8 e1002709 [PubMed: 22589728]
158. Li M, Xu F, Muller J, Hearing VJ, Gorelik E. Ecotropic C-type retrovirus of B16 melanoma and malignant transformation of normal melanocytes. *Int J Cancer.* 1998; 76: 430–36. [PubMed: 9579583]
159. Malarkannan S, Serwold T, Nguyen V, Sherman LA, Shastri N. The mouse mammary tumor virus env gene is the source of a CD8⁺ T-cell-stimulating peptide presented by a major histocompatibility complex class I molecule in a murine thymoma. *PNAS.* 1996; 93: 13991–96. [PubMed: 8943048]
160. Dudley J, Rissler R. Amplification and novel locations of endogenous mouse mammary tumor virus genomes in mouse T-cell lymphomas. *J Virol.* 1984; 49: 92–101. [PubMed: 6317898]

161. Song E, Mao T, Dong H, Boisserand LSB, Antila S, et al. VEGF-C-driven lymphatic drainage enables immunosurveillance of brain tumours. *Nature*. 2020; 577: 689–94. [PubMed: 31942068]
162. Griffin GK, Wu J, Iracheta-Vellve A, Patti JC, Hsu J, et al. Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. *Nature*. 2021; 595: 309–14. [PubMed: 33953401]
163. Achleitner M, Kleefisch M, Hennig A, Peschke K, Polikarpova A, et al. Lack of Treg1 causes systemic autoimmunity despite the presence of antiretroviral drugs. *J Immunol*. 2017; 199: 2261–69. [PubMed: 28835460]
164. Ambati J, Magagnoli J, Leung H, Wang SB, Andrews CA, et al. Repurposing anti-inflammasome NRTIs for improving insulin sensitivity and reducing type 2 diabetes development. *Nat Commun*. 2020; 11 4737 [PubMed: 32968070]
165. Hladik F, Burgener A, Ballweber L, Gottardo R, Vojtech L, et al. Mucosal effects of tenofovir 1% gel. *eLife*. 2015; 4 e04525 [PubMed: 25647729]
166. Keller MJ, Wood L, Billingsley JM, Ray LL, Goymer J, et al. Tenofovir disoproxil fumarate in-travaginal ring for HIV pre-exposure prophylaxis in sexually active women: a phase 1, single-blind, randomised, controlled trial. *Lancet HIV*. 2019; 6 e498-508 [PubMed: 31320290]
167. Hughes SM, Levy CN, Calienes FL, Stekler JD, Pandey U, et al. Treatment with commonly used antiretroviral drugs induces a type I/III interferon signature in the gut in the absence of HIV infection. *Cell RepMed*. 2020; 1 100096
168. Rajurkar M, Parikh AR, Solovyov A, You E, Kulkarni AS, et al. Reverse transcriptase inhibition disrupts repeat element life cycle in colorectal cancer. *Cancer Discov*. 2022; 12 (6) 1462–81. [PubMed: 35320348]
169. Karijolich J, Abernathy E, Glaunsinger BA. Infection-induced retrotransposon-derived noncoding RNAs enhance herpesviral gene expression via the NF- κ B pathway. *PLOS Pathog*. 2015; 11 e1005260 [PubMed: 26584434]
170. Schmidt N, Domingues P, Golebiowski F, Patzina C, Tatham MH, et al. An influenza virus-triggered SUMO switch orchestrates co-opted endogenous retroviruses to stimulate host antiviral immunity. *PNAS*. 2019; 116: 17399–408. [PubMed: 31391303]
171. Lefkopoulos S, Polyzou A, Derecka M, Bergo V, Clapes T, et al. Repetitive elements trigger RIG-I-like receptor signaling that regulates the emergence of hematopoietic stem and progenitor cells. *Immunity*. 2020; 53: 934–51. e9 [PubMed: 33159854]
172. Jayewickreme R, Mao T, Philbrick W, Kong Y, Treger RS, et al. Endogenous retroviruses provide protection against vaginal HSV-2 disease. *Front Immunol*. 2021; 12 758721 [PubMed: 35058919]
173. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015; 160: 48–61. [PubMed: 25594174]
174. Desai N, Sajed D, Arora KS, Solovyov A, Rajurkar M, et al. Diverse repetitive element RNA expression defines epigenetic and immunologic features of colon cancer. *JCI Insight*. 2017; 2 e91078 [PubMed: 28194445]
175. Goel S, DeCristo MJ, Watt AC, BrinJones H, Sceneay J, et al. CDK4/6 inhibition triggers anti-tumour immunity. *Nature*. 2017; 548: 471–75. [PubMed: 28813415]
176. Gannon HS, Zou T, Kiessling MK, Gao GF, Cai D, et al. Identification of ADAR1 adenosine deaminase dependency in a subset of cancer cells. *Nat Commun*. 2018; 9 5450 [PubMed: 30575730]
177. Leruste A, Tosello J, Ramos RN, Tauziède-Espariat A, Brohard S, et al. Clonally expanded T cells reveal immunogenicity of rhabdoid tumors. *Cancer Cell*. 2019; 36: 597–612. e8 [PubMed: 31708437]
178. Krug B, De Jay N, Harutyunyan AS, Deshmukh S, Marchione DM, et al. Pervasive H3K27 acetylation leads to ERV expression and a therapeutic vulnerability in H3K27M gliomas. *Cancer Cell*. 2019; 35: 782–97. e8 [PubMed: 31085178]
179. Shen JZ, Qiu Z, Wu Q, Finlay D, Garcia G, et al. FBXO44 promotes DNA replication-coupled repetitive element silencing in cancer cells. *Cell*. 2021; 184: 352–69. e23 [PubMed: 33357448]
180. Espinet E, Gu Z, Imbusch CD, Giese NA, Büscher M, et al. Aggressive PDACs show hypomethylation of repetitive elements and the execution of an intrinsic IFN program linked to a ductal cell of origin. *Cancer Discov*. 2021; 11: 638–59. [PubMed: 33060108]

181. Zhou X, Singh M, Sanz Santos G, Guerlavais V, Carvajal LA, et al. Pharmacologic activation of p53 triggers viral mimicry response thereby abolishing tumor immune evasion and promoting antitumor immunity. *Cancer Discov.* 2021; 11 (12) 3090–105. [PubMed: 34230007]
182. Gu Z, Liu Y, Zhang Y, Cao H, Lyu J, et al. Silencing of LINE-1 retrotransposons is a selective dependency of myeloid leukemia. *Nat Genet.* 2021; 53: 672–82. [PubMed: 33833453]
183. Kassiotis G, Stoye JP. Immune responses to endogenous retroelements: taking the bad with the good. *Nat Rev Immunol.* 2016; 16: 207–19. [PubMed: 27026073]
184. Ishak CA, Classon M, De Carvalho DD. Deregulation of retroelements as an emerging therapeutic opportunity in cancer. *Trends Cancer.* 2018; 4: 583–97. [PubMed: 30064665]
185. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun.* 2019; 10 4278 [PubMed: 31537801]
186. Young GR, Stoye JP, Kassiotis G. Are human endogenous retroviruses pathogenic? An approach to testing the hypothesis. *BioEssays.* 2013; 35: 794–803. [PubMed: 23864388]
187. Mavrommatis B, Baudino L, Levy P, Merkenschlager J, Eksmond U, et al. Dichotomy between T cell and B cell tolerance to neonatal retroviral infection permits T cell therapy. *J Immunol.* 2016; 197: 3628–38. [PubMed: 27647833]
188. Sacha JB, Kim IJ, Chen L, Ullah JH, Goodwin DA, et al. Vaccination with cancer- and HIV infection-associated endogenous retrotransposable elements is safe and immunogenic. *J Immunol.* 2012; 189: 1467–79. [PubMed: 22745376]
189. Kassiotis G, Stoye JP. Making a virtue of necessity: the pleiotropic role of human endogenous retroviruses in cancer. *Philos Trans R Soc Lond B.* 2017; 372 20160277 [PubMed: 28893944]
190. Hedrick SM. The acquired immune system: a vantage from beneath. *Immunity.* 2004; 21: 607–15. [PubMed: 15539148]
191. Bauer JH, Helfand SL. New tricks of an old molecule: lifespan regulation by p53. *Aging Cell.* 2006; 5: 437–40. [PubMed: 16968311]
192. Schäfer M, Werner S. Cancer as an overheating wound: an old hypothesis revisited. *Nat Rev Mol CellBiol.* 2008; 9: 628–38.
193. Burns KH. Our conflict with transposable elements and its implications for human disease. *Annu Rev Pathol.* 2020; 15: 51–70. [PubMed: 31977294]
194. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science.* 2016; 351: 1083–87. [PubMed: 26941318]
195. Ferreira LM, Meissner TB, Mikkelsen TS, Mallard W, O'Donnell CW, et al. A distant trophoblast-specific enhancer controls HLA-G expression at the maternal-fetal interface. *PNAS.* 2016; 113: 5364–69. [PubMed: 27078102]
196. Chikuma S, Yamanaka S, Nakagawa S, Ueda MT, Hayabuchi H, et al. TRIM28 expression on dendritic cells prevents excessive T cell priming by silencing endogenous retrovirus. *J Immunol.* 2021; 206: 1528–39. [PubMed: 33619215]
197. Marasca F, Sinha S, Vadalà R, Polimeni B, Ranzani V, et al. LINE1 are spliced in non-canonical transcript variants to regulate T cell quiescence and exhaustion. *Nat Genet.* 2022; 54 (2) 180–93. [PubMed: 35039641]
198. Ng KW, Attig J, Young GR, Ottina E, Papamichos SI, et al. Soluble PD-L1 generated by endogenous retroelement exaptation is a receptor antagonist. *eLife.* 2019; 8 e50256 [PubMed: 31729316]
199. Payer LM, Steranka JP, Ardeljan D, Walker J, Fitzgerald KC, et al. Alu insertion variants alter mRNA splicing. *Nucleic Acids Res.* 2019; 47: 421–31. [PubMed: 30418605]

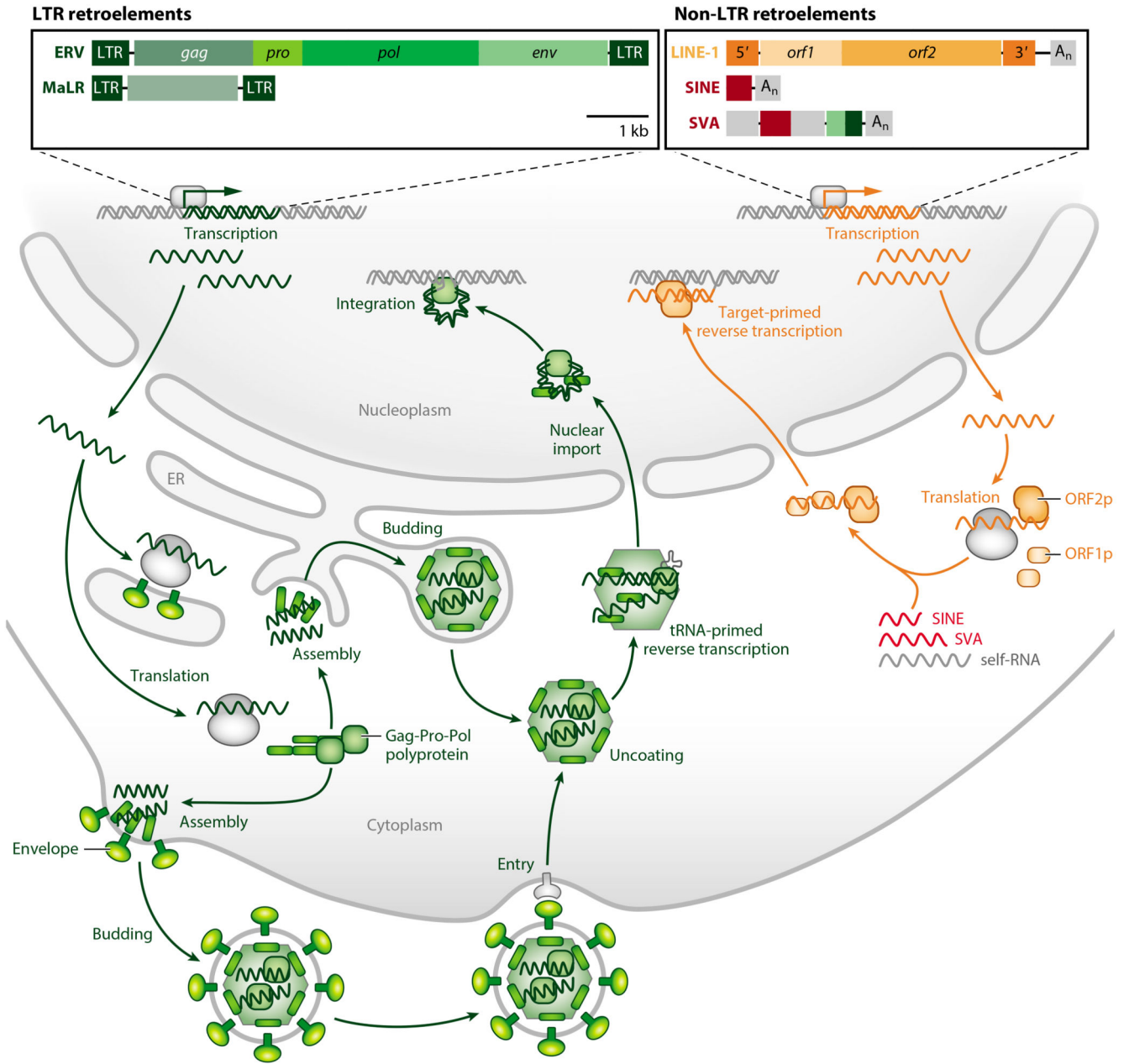


Figure 1. Genome structure and replication cycles of LTR (*left*) and non-LTR retroelements (*right*). Genomes of canonical LTR retroelements, including ERV and MaLR (*green*), and of non-LTR retroelements, including LINE-1 (*orange*), SINE, and SVA (*red*). SVAs are composite repetitive elements comprising a CCCTCT repeat section (*first gray rectangle*), two *Alu*-like sequences in reverse orientation, a VNTR section (*second gray rectangle*), and a region derived from ERV *env* and 5' LTR. Also depicted are the *env* (envelope), *gag* (group-specific antigen), and *pro-pol* (protease-polymerase) ORFs of the complete ERV and the *orf1* and *orf2* ORFs of the complete LINE-1. Also depicted are the poly-A regions (A_n) at the 3' end of non-LTR elements. Not depicted are accessory ORFs in more complex

ERVs [e.g., *rec* and *np9* in certain HERV-K(HML2) proviruses] or the *orf0* present in reverse orientation in the 5' UTR of LINE-1. Transcribed ERV RNA is exported to the cytoplasm, where it serves as mRNA for translation of functional proteins or as genomic RNA. Assembly of ERV particles typically takes place at the plasma membrane, with budding into the extracellular space, but ERVs lacking *env* and MaLR elements have adopted intracellular budding, such as into the ER or other subcellular locations. Reverse transcription of ERV genomic RNA is primed by host tRNA in the uncoated virus particle core and the pre-integration complex is then imported into the nucleus for integration of the virus DNA into the host DNA. Transcribed LINE-1 RNA is also exported to the cytoplasm acting as mRNA and genomic RNA, with the translated proteins, particularly ORF2p, which is made in limited amounts, exhibiting *cis* preference for the LINE-1 mRNA that has produced it. Nevertheless, SINE or SVA RNA, and, in principle, any polyadenylated host RNA (self-RNA), may also bind ORF2p. The complexes are imported into the nucleus, where ORF2p attacks host DNA, which is used to prime reverse transcription of associated RNA and integration of the complementary DNA (cDNA) copy. Abbreviations: ER, endoplasmic reticulum; ERV, endogenous retrovirus; LINE-1, long interspersed nuclear elements 1; LTR, long terminal repeat; MaLR, mammalian apparent LTR retrotransposon; ORF, open reading frame; SINE, short interspersed nuclear element; SVA, SINE-VNTR-*Alu*; UTR, untranslated region; VNTR, variable-number tandem repeat.

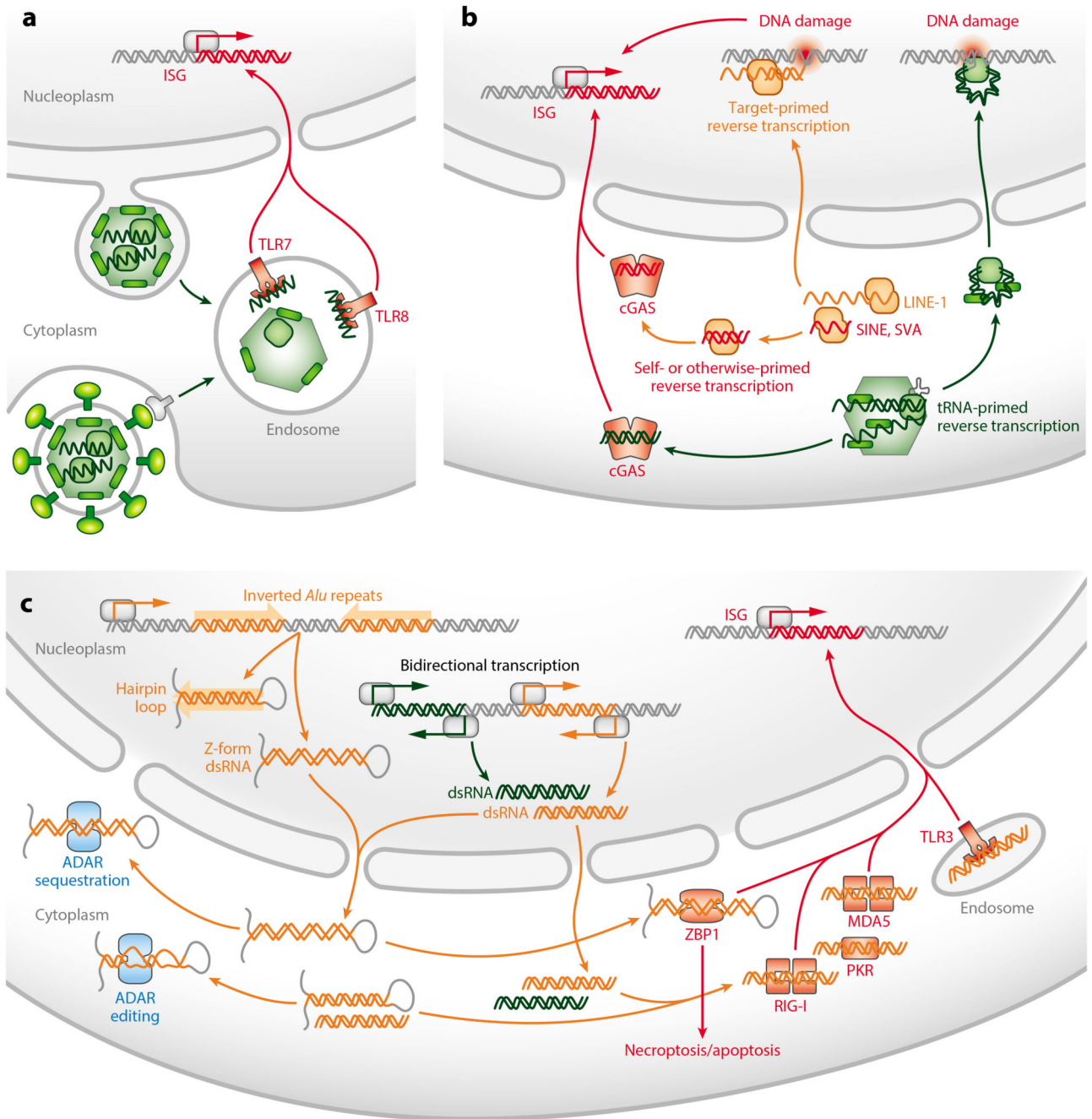


Figure 2. Immunogenicity of canonical and aberrant ERE nucleic acid replication intermediates. (a) Endosomal sensors of ssRNA, TLR7, and TLR8 may be triggered by ERV genomic RNA. This can derive from incoming extracellular virus particles accessing the endosome after cell entry or from intracellularly formed particles gaining access to the endosome through alternative routes, including autophagy (not depicted). Not depicted is the signaling cascade initiated by TLR7 and TLR8 ligation, which ultimately leads to the transcriptional induction of ISGs. (b) Cytoplasmic DNA sensors such as cGAS may be triggered by cDNA produced

by ERVs, as part of the typical ERV replication cycle, or by non-LTR retroelements through aberrant cytoplasmic reverse transcription. The latter can be self-primed (as in the case of *A/u*) or primed by an as yet unknown mechanism. Triggering of cGAS and its downstream adaptor STING (not depicted) then leads to an ISG response. Separately, the endonuclease activities of ERV polymerase and LINE-1 ORF2p catalyze DNA breaks during the replication cycle, and the ensuing DNA damage may indirectly trigger an ISG response. (c) Potentially immunogenic dsRNA is also aberrantly produced by LTR and non-LTR retroelements through distinct mechanisms. Bidirectional transcription of EREs generates intermolecular complementary RNA, forming dsRNA. Transcription of inverted SINE/*A/u* repeats generates regions of intramolecular complementarity leading to the formation of hairpin loops. More enigmatic is the formation of Z-form dsRNA, also enriched for SINE/*A/u* sequences. The immunogenicity of hairpin loop and Z-form dsRNA may be reduced by ADAR-mediated editing and sequestration, respectively, but increased dsRNA formation or insufficient ADAR activity permits the triggering of several dsRNA sensors, including MDA5, RIG-I, PKR, and TLR3, initiating signaling cascades that converge to an ISG response. The Z-form dsRNA-binding protein ZBP1, which typically induces necroptosis or apoptosis, also contributes to the ISG response triggered when ADAR activity is reduced. Abbreviations: cDNA, complementary DNA; dsRNA, double-stranded RNA; ERE, endogenous retroelement; ERV, endogenous retrovirus; ISG, interferon-stimulated gene; LINE-1, long interspersed nuclear elements 1; LTR, long terminal repeat; SINE, short interspersed nuclear element; ssRNA, single-stranded RNA; SVA, SINE-VNTR-*A/u*; TLR, Toll-like receptor; VNTR, variable-number tandem repeat.

Genomic locus



Transcript and protein isoforms

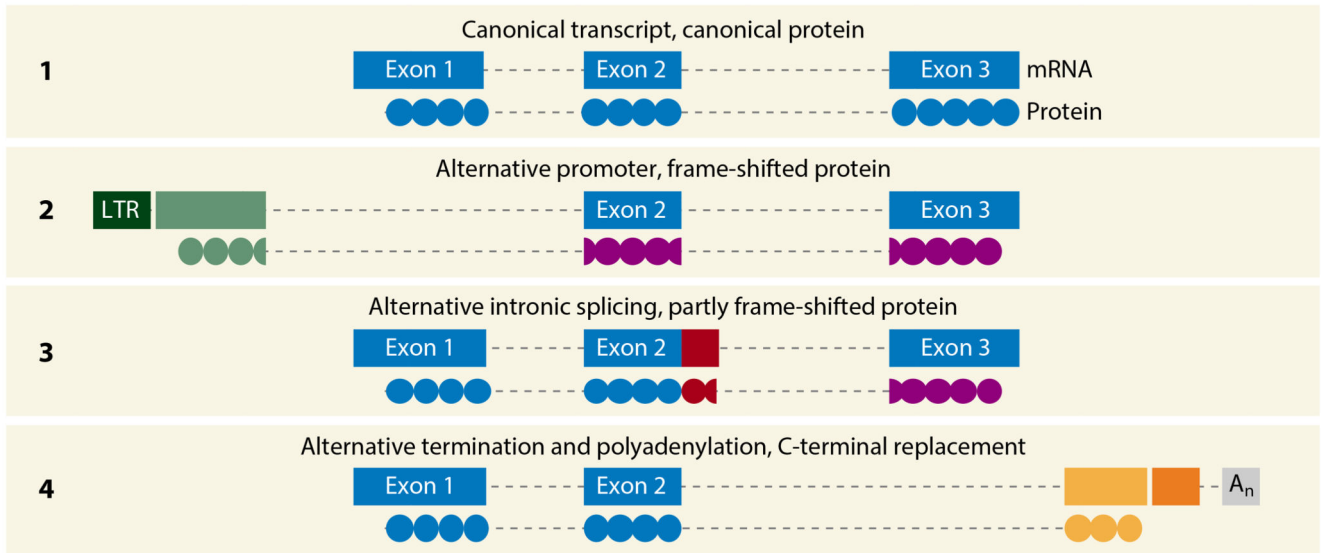


Figure 3.

Generation of aberrant protein products by transcriptional utilization of EREs. A hypothetical example of a protein-coding gene consisting of three exons with EREs integrated within and around the gene body (*top*) and potential alternative splicing isoforms of transcribed RNA (*bottom*). Isoform 1 represents the canonical mRNA and translated protein. Isoform 2 uses an LTR element as an alternative promoter, skipping the first canonical exon and shifting the translation frame. Isoform 3 uses an alternative splice donor site in an intronic SINE element instead of the canonical site at the end of exon 2, shifting the remaining translation. Isoform 4 uses a LINE-1 fragment as alternative terminal exon, replacing the C-terminal sequence of the translation product. Abbreviations: ERE, endogenous retroelement; LINE-1, long interspersed nuclear elements 1; LTR, long terminal repeat; SINE, short interspersed nuclear element.

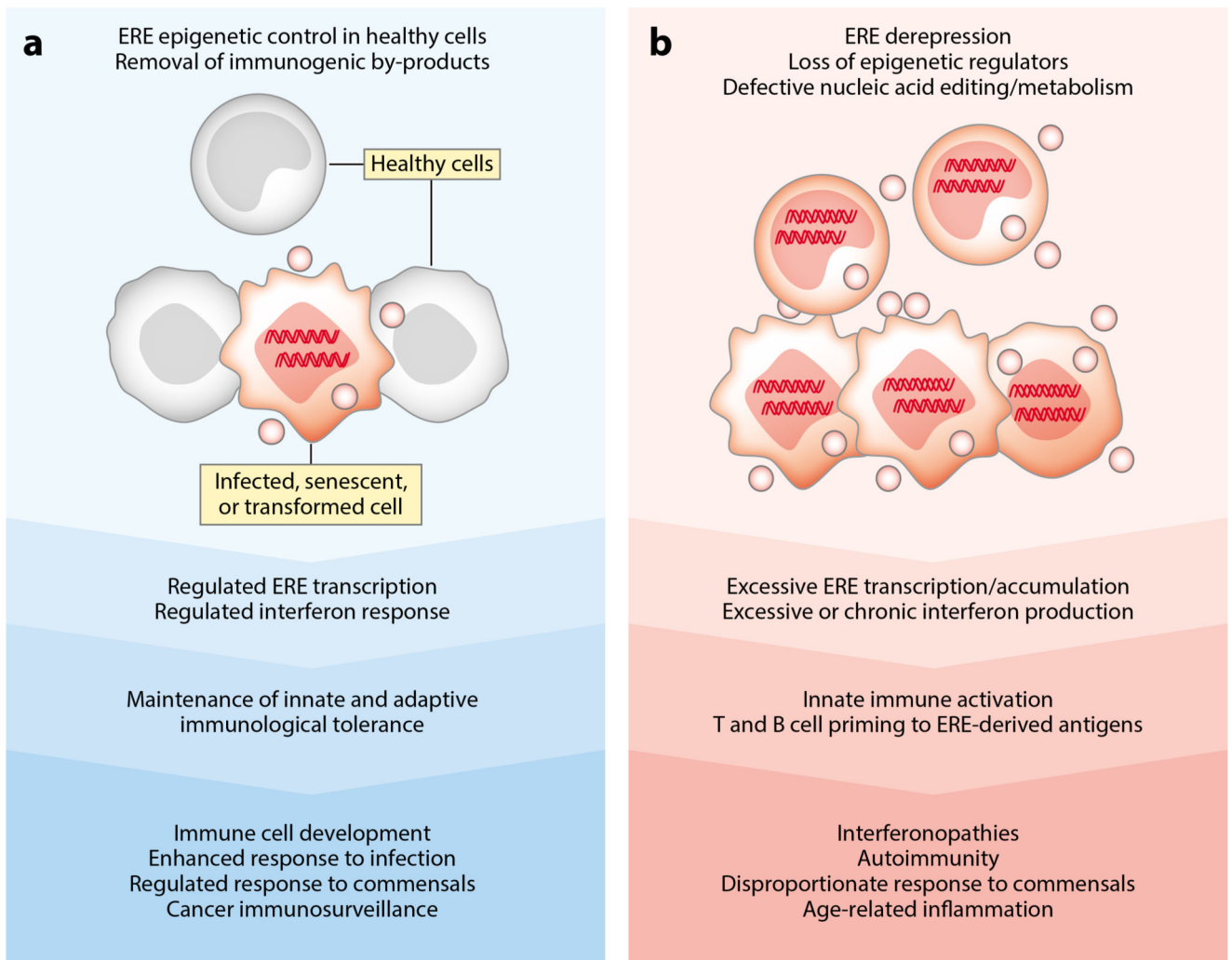


Figure 4.

Proposed consequences of ERE immunogenicity, depending on the degree of ERE activity. In physiological conditions (*a*), ERE transcriptional activity is epigenetically controlled and their products eliminated in most cells, with the exception of infected, transformed, senescent, or otherwise stressed cells. In turn, this prevents interferon responses and priming of adaptive immune cells against the low level of ERE products in healthy cells but permits immune reactions against elevated ERE products in stressed cells. Such regulated responses are thought to contribute to several physiological processes. In contrast, when ERE transcription is unleashed (*b*), through loss of epigenetic control, or their products accumulate, through loss of nucleic acid metabolism or editing machineries, in a sufficient number of otherwise healthy cells, the resulting excessive interferon and adaptive immune responses can trigger or contribute to a range of pathological conditions. Abbreviation: ERE, endogenous retroelement.