



How Retroviruses and Retrotransposons in Our Genome May Contribute to Autoimmunity in Rheumatological Conditions

Tomas Mustelin* and Kennedy C. Ukadike

Division of Rheumatology, Department of Medicine, University of Washington, Seattle, WA, United States

OPEN ACCESS

Edited by:

Gregg Joshua Silverman,
New York University, United States

Reviewed by:

Richard David Leslie,
Queen Mary University of London,
United Kingdom
Janine Adele Lamb,
The University of Manchester,
United Kingdom

*Correspondence:

Tomas Mustelin
tomas2@uw.edu

Specialty section:

This article was submitted to
Autoimmune and
Autoinflammatory Disorders,
a section of the journal
Frontiers in Immunology

Received: 11 August 2020

Accepted: 19 October 2020

Published: 13 November 2020

Citation:

Mustelin T and Ukadike KC
(2020) How Retroviruses and
Retrotransposons in Our Genome May
Contribute to Autoimmunity in
Rheumatological Conditions.
Front. Immunol. 11:593891.
doi: 10.3389/fimmu.2020.593891

More than 200 human disorders include various manifestations of autoimmunity. The molecular events that lead to these diseases are still incompletely understood and their causes remain largely unknown. Numerous potential triggers of autoimmunity have been proposed over the years, but very few of them have been conclusively confirmed or firmly refuted. Viruses have topped the lists of suspects for decades, and it seems that many viruses, including those of the Herpesviridae family, indeed can influence disease initiation and/or promote exacerbations by a number of mechanisms that include prolonged antiviral immunity, immune subverting factors, and mechanisms, and perhaps “molecular mimicry”. However, no specific virus has yet been established as being truly causative. Here, we discuss a different, but perhaps mechanistically related possibility, namely that retrotransposons or retroviruses that infected us in the past and left a lasting copy of themselves in our genome still can provoke an escalating immune response that leads to autoimmune disease. Many of these loci still encode for retroviral proteins that have retained some, or all, of their original functions. Importantly, these endogenous proviruses cannot be eliminated by the immune system the way it can eliminate exogenous viruses. Hence, if not properly controlled, they may drive a frustrated and escalating chronic, or episodic, immune response to the point of a frank autoimmune disorder. Here, we discuss the evidence and the proposed mechanisms, and assess the therapeutic options that emerge from the current understanding of this field.

Keywords: autoimmunity, retrotransposons, retroelements, nucleic acid sensors, reverse transcriptase, type I interferon, endogenous retroviruses

RELEVANT MOLECULAR CONCEPTS OF HUMAN AUTOIMMUNE DISEASES

An important feature of clinical autoimmunity is that patients tend to fall into a discrete number of reasonably well delineated and named disease entities (i.e. diagnoses), rather than spanning the full spectrum of autoimmunity against random antigens. Granted, there is variability and heterogeneity within most such disease entities; and some diagnoses may in fact represent more than one distinct

disease or a series of mechanistically different molecular “endotypes” of the disease. Nevertheless, patients that deviate radically from the typical disease profiles are rare. This pattern of many distinct diseases does not readily mesh with the commonly accepted notion that autoimmunity starts with a simple stochastic loss-of-tolerance blunder by a T cell. Rather, it seems that autoimmune diseases must be the result of distinct and unique pathophysiological processes that evolve over an extended period of time into a specific disease. The two concepts are of course not mutually exclusive, but they shape our thinking in different ways: while the former focuses autoimmunity research on T or B cell antigen receptor repertoire and mechanisms of central and peripheral tolerance, the notion that autoimmunity may arise from specific biological processes broadens the search for disease triggers and attempts to understand the escalation towards disease. The therapeutic ramifications of these two views are also distinct: a T or B cell-centric view calls for immunosuppressive or tolerizing approaches, while the concept of specific biological processes resulting in autoimmunity will look for specific modulation of such processes without the need for suppressing the normal function of the immune system. In this review, we follow the notion that autoimmune disease can have causes other than “stochastic mistakes of adaptive immunity”. We accept that T and B cells are critically important for autoimmunity, but we are not convinced that they initiate it.

Although individual autoimmune diseases can be clinically quite different from each other and often are associated with polymorphisms in different genes, and may respond to different targeted therapies, it is also clear that some diseases likely have overlapping pathogenic mechanisms; these mechanistically “related” diseases share cardinal features and symptoms and can co-occur in individual patients (e.g., SLE and secondary Sjögren’s syndrome). An example of a group of such “related” autoimmune disease are those characterized by elevated type I interferons (IFNs) (1–4), including large portions of systemic lupus erythematosus (SLE) (1, 5, 6), dermatomyositis (DM) (7), primary Sjögren’s syndrome (pSS) (8–10), and several others. Type I IFNs are a hallmark of anti-viral immunity, with which these diseases appear to share other features as well, including autoimmunity against a similar set of proteins involved in nucleic acid processing, as well as the nucleic acids themselves. As in autoimmune diseases, viral infections are often accompanied by fever, headache, loss of appetite, malaise, fatigue, arthralgias, and sometimes skin rash. During viral infections, these responses are transient, while in SLE they become chronic with an unpredictable and often episodic course.

We recently reviewed the currently known and proposed sources of pathogenic nucleic acids and how they can act to drive SLE-like autoimmunity (11). Briefly, the offending nucleic acid could be either DNA or RNA, or both. Pathogenic cytosolic DNA may leak out from the nucleus following extensive damage to chromosomal DNA or mitotic catastrophes (not very likely in autoimmunity). DNA can escape from defective mitochondria, or DNA can be synthesized by reverse-transcription of various species of RNA (particularly from retroelements). Extracellular

DNA may spill out from cells dying by a variety of programmed cell death mechanisms, or from commensal gut microbes, and then be internalized and sensed by immune cells. Pathogenic RNA can be (mis)generated and sensed intracellularly or end up in the extracellular space from which it can be internalized, for example as part of immune complexes, to be sensed by endosomal toll-like receptors (TLRs) in immune cells. In this review, we focus on RNA transcripts derived from endogenous retroviruses and retrotransposons and on the extrachromosomal DNA synthesized by reverse transcription of these RNA species. We also discuss the potential contributions of proteins generated by translation of these RNA transcripts, which may form more or less complete virions.

An important concept to keep in mind when contemplating how aberrant DNA or RNA drive autoimmunity is that a multitude of ancient and powerful mechanisms exist within our cells to effectively prevent the expression of potentially problematic sequences in our genome and to effectively degrade and remove aberrant DNA or RNA. These mechanisms are reviewed in section *Defense Mechanisms Against Retroviruses and Retrotransposons: Our Original Immunity*. Their importance to our health is perhaps best illustrated by the serious diseases that arise from mutations in the genes for several of these pathways, including Aicardi-Goutières Syndrome (AGS), which is characterized by constitutively elevated type I interferons and SLE-like autoimmunity. It presents at birth as a suspected neonatal viral infection, which is a medical emergency, but no exogenous virus can be found and the disease continues unabated. Over the years, AGS patients develop severe neurological deficits, perhaps due to direct neurotoxicity of type I IFNs. In regular polygenic SLE, however, it remains unclear if these defense mechanisms are weakened or simply overcome by an abundance of aberrant RNA or DNA. There are many potential variants of these scenarios. We have proposed that the clinical heterogeneity of SLE may be due, in part, to heterogeneity in which pathogenic nucleic acid molecules are present and which sensors and pathways they trigger in individual patients (11). Elucidation of these events may result in the recognition of distinct “endotypes” of SLE, each with its specific therapeutic opportunities.

ENDOGENOUS RETROVIRUSES AND RETROELEMENTS IN OUR GENOME

“By DNA sequence, we are more retroviral than human” is a provocative way of pointing out that a considerably larger portion of our human genome consists of sequences that once were RNA genomes of free and infectious retroviruses that were reverse transcribed into DNA and then pasted into our genome; they are more abundant (8%) than all the exons of our protein-coding “traditional” genes combined (about 1%) (12). Since most, if not all, genomes of eukaryotic and prokaryotic organisms on our planet share this feature of abundant inserted retroviral sequences, it is very likely that the 8% of our genome that is readily recognizable today as retroviral in origin

is, in fact, only the tip of the iceberg. Most such sequences are not positively selected for (but likely the opposite) and over evolutionary time lose their integrity by random mutations, deletions, recombinations, and other mechanisms. For these reasons, sequences older than 100 million years become increasingly difficult to recognize with acceptable confidence.

Since reverse-transcribed retroviral sequences are present in all kingdoms of life, it appears that this influx of genomic material started at the very dawn of cellular life (13–16). In fact, it is quite likely that it was instrumental to the evolution of larger and more diverse genomes: each newly incorporated reverse-transcribed retroviral genome adds ~ 9,500 base pairs to the genome (**Figure 1**), including three major protein-coding genes, *gag*, *pol*, and *env*, plus mRNA splicing sites, to generate transcripts that are translated into at least five proteins, each of which can be proteolytically processed into additional functional units. In addition, each retroviral integration brings two identical long terminal repeats (LTRs), one on each end of the insert, which contain clusters of highly efficient transcription factor

binding sites to control transcription of the insert, as well as adjacent regions. In fact, it has been estimated that more than 300,000 regulatory regions (including promoters and enhancers) in our genome are, or contain remnants of, ancient LTRs from otherwise long-lost retroviral inserts. It is also clear that many “traditional” genes are descendants of ancient retroviral *gag*, *pol*, or *env* genes that were co-opted for new uses (17). For example, the RNaseH and integrase domains of the retroviral *pol* gene served as starting material for fundamental building blocks of our immune system (18). It has also been suggested that mRNA splicing was originally a retroviral invention.

The 8% of our human genome that consists of recognizable proviruses (12) (*i.e.* the cDNA of retrovirus RNA genomes, or parts of them), are collectively termed the human endogenous retroviruses (HERVs). Strictly speaking, this term is not entirely accurate since the majority of these sequences were incorporated long before our hominin ancestors became *Homo sapiens*. Hence, the term HERVs should be viewed as the complement of retroviral sequences in their current state, which for essentially

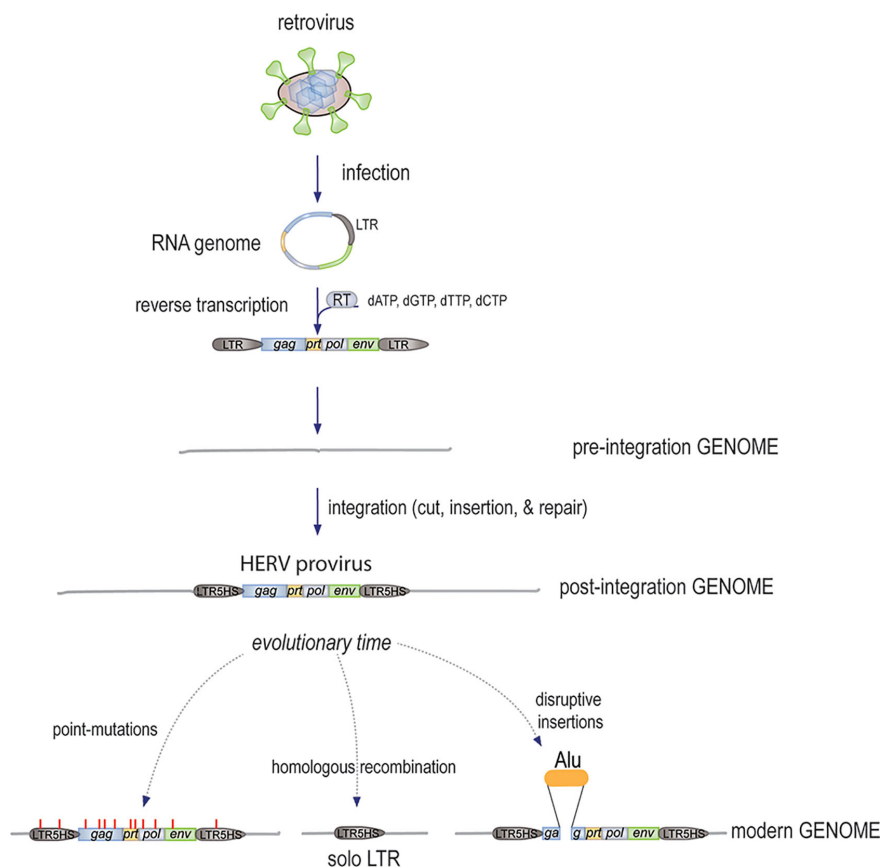


FIGURE 1 | The mechanism by which new HERVs were generated. An infectious free retrovirus infects a host germline cell, releases its circular RNA genome, which remains protected by the associated *gag*-derived nucleocapsid and other core proteins, while the *pol*-encoded RT synthesizes the first strand of linear cDNA starting with the LTR and ending after the second copy of the LTR, followed by second strand synthesis. The resulting dsDNA of approximately 9,500 bp is then inserted into the genome by the endonuclease activity of the Pol protein. The ends are then finalized by the DNA repair machinery with a few added nucleotides. Over evolutionary times, further changes to the HERVs included the accumulation of point-mutations (some introducing stop-codons), deletions by homologous recombination, and disruption by additional insertions.

all of them is different from what they looked like when they were free and contagious retroviruses that infected our ancestors and in the process reverse-transcribed their RNA genome and inserted it into the germline of their host. Although they were subsequently inherited in a Mendelian fashion by all descendants of the original host, there was little evolutionary pressure to maintain them in their intact form; more likely the opposite. Chimpanzees and gorillas have remarkably similar sets of retroviral loci (19–21), except for different mutations and the dozen or so new integrations that occurred in each species since our last shared ancestor lived approximately 6 million years ago.

In addition to the *bona fide* HERVs, an even larger portion of our genome, over 30%, consists of copies of non-LTR retrotransposons (12). Collectively, endogenous retroviruses and retrotransposons are referred to as retroelements. The non-LTR retrotransposons are classified as either “autonomous” or “non-autonomous” depending on whether they contain all the required components necessary for retrotransposition within themselves, or not. The most abundant class of autonomous retrotransposons are the long interspersed nuclear elements-1 (LINE-1 or L1 for short) (22). The biology and potential relevance of L1 in autoimmunity is discussed in section *Non-LTR Retrotransposons—LINE-1 and Alu Elements*.

The non-autonomous retrotransposons include the short interspersed nuclear elements (SINE), such as the Alu (23) elements and SINE-R, VTR, Alu (SVA) elements (24), which all depend on the L1 reverse transcriptase (RT) for their retrotransposition cycle. The Alu sequence itself appears to be a contracted form of the 7SL RNA (25), which is a component of the signal recognition particle. Alu elements have been extraordinarily successful in replicating within our genome: there are approximately 1 million of them in our genome, many of them within introns, where they may modulate gene transcription and mRNA processing.

The Many Families of Retroviral Sequences in Our Genome

The tens of thousands of retroviral sequences that exist in our genome belong to more than a dozen distinct families, which originally were distinct free retroviruses that infected our ancestors during different, but often overlapping, epochs of prehistorical times. Most of these sequences have accumulated numerous mutations and deletions, and some have been disrupted by insertions of other retrotransposons, e.g. Alu elements. The older sequences have more such alterations and have lost their ability to encode full-length retroviral proteins, but the more recently incorporated ones are more complete and still retain the capacity to encode fully functional proteins and to produce viral particles. However, it seems that none of the HERVs are fully infectious anymore.

A basic nomenclature divides Retroviridae into four classes: gammaretroviruses (class I), betaretroviruses (class II), spumaretroviruses (class III), and lentiviruses (class IV). The first three classes are represented in our genome. They are further divided based on the specific tRNA they use for priming of reverse transcription. In essence, the retroviral RT that generates

a DNA copy of the circular viral RNA genome uses a cellular tRNA complementary to a short motif in the viral LTR for priming of the reaction. The youngest of the Class II HERVs, for example, used lysine-tRNA for priming and are therefore classified as group K (for lysine), hence named HERV-K. The shortcomings of this classification, e.g. its lack of further taxonomic considerations, prompted other classification principles to be proposed. Unfortunately, these efforts to bring order only resulted in several parallel nomenclatures and, as a result, many loci have non-conforming and confusing names, as well as several synonymic designations. A more precise way to add specificity is to mention which chromosome the locus is on and exactly which nucleotide positions it occupies in the human reference genome, e.g., HERV-K119 occupies nucleotides 58,721,242–58,730,698 of chromosome 12.

In this review, we focus only on those families that have been proposed to be of potential relevance in human autoimmunity: primarily the “Human MMTV-like 2” (HML-2) subgroup of the Class II HERV-K (26) and the Class I provirus HERV-E (27). We would postulate, however, that individual HERVs that may be detrimental to our health could belong to any family. At the same time, we find it more likely that the most recently incorporated HERVs, which have retained much of their original features and still can produce virions (albeit all with reduced infectivity), are more likely to cause immune disorders resembling chronic viral infections than the older HERVs, which often are incomplete, and have been “domesticated” by frame-shifts, point-mutations, and stop codons. We accept, of course, that older HERVs may have acquired new properties by stochastic mutations and thereby gained the ability to drive unique pathologies unrelated to the mechanisms of typical antiviral immunity.

HERV-K (HML-2)—The Youngest and Most Intact HERV Family

Although the now (presumably) extinct free retrovirus that gave rise to the HERV-K (HML-2) provirus family (26) repeatedly infected our ancestors for tens of millions of years, we will probably never know what kind of disease it caused at that time. What we can conclude using computational tools from the ~120 genomic loci still present today is that HERV-K(HML-2) first entered our ancestral early hominin germline genome over 30 million years ago (28) and then continued to insert again and again into our germline genome until very recently in evolutionary time (29). Many other retroviruses stopped incorporating into our germline much earlier. Obviously, the potentially vast numbers of infections that did not result in germline insertions are invisible to us today, even if they likely were important for the life-cycle and spread of the virus. Hence, the HERV-K provirus loci in our genome represent a vast underestimate of the number of times the free virus infected and perhaps profoundly affected our ancestors. The insertional polymorphisms (*i.e.* only some people have some of them) (30–32) and the polymorphic deletions (33) observed in human populations today reveal that infections probably continued until times when modern humans were more numerous and

had spread out over larger geographical areas in the last 50–70,000 years. Exogenous HML-2 appears to have infected gorillas relatively recently as well (34).

The age of a HERV-K locus (*i.e.* the time since germline integration) can be estimated from the fact that the single LTR in the circular retroviral RNA genome is reverse-transcribed twice, resulting in two identical LTR copies, one in each end of the resulting genomic provirus (**Figure 1**). Since there is no evolutionary pressure to maintain these sequences, they are assumed to be subject to stochastic mutations at the standard background rate of approximately 0.5×10^{-9} per base-pair per year. This “molecular clock” obviously can only be applied to HERV-K loci that have retained both LTRs. Based on this logic, an alignment of the seven youngest HERV-K proviruses was used to deduce *in silico* what the sequence of the original infectious retroviruses most likely was. Albeit not necessarily 100% correct, the resulting sequence gave rise to a fully infectious retrovirus, termed HERV-K Phoenix (35), which has been studied for its tropism, cellular receptors, maturation, ultrastructure by electron microscopy, and ability to reverse transcribe and insert its cDNA into the genome of host cells (35).

The most recent human insertions of the HERV-K (HML-2) subfamily, *e.g.*, HERV-K113 at chromosome 19p12 (36), are also intact enough to produce virions (37), albeit with poor infectivity. Other full-length HERV-Ks with intact open reading-frames are HERV-K108a (at 7p22.1), HERV-K115 (8p23.1), HERV-K118 (at 11q22.1), and HERV-K119 (at 12q14.1). Another seemingly intact HERV-K provirus is located at Xp21.33 in approximately 2% of people, most of whom are of African ancestry (31). HERV-K113 and HERV-K115 are also insertionally polymorphic and exist in 15–30% of modern humans.

To the extent that we know, these youngest loci are transcriptionally silenced in healthy individuals by extensive DNA methylation and other epigenetic mechanisms, as one would assume for potentially dangerous loci. A consequence of this is that they probably remain silent during the development of T and B cell antigen receptor repertoires in early life, resulting in weak immunological tolerance against the proteins that they can encode. If this indeed is true, aberrant expression of these proteins would likely provoke both cellular and humoral immunity (38). There is supporting evidence for this assumption: increased transcription in malignancies of the breast (39) and prostate (40), and in HIV infected individuals (38, 41–45), leads to both (auto)antibodies against HERV-K proteins and HERV-K-specific T cells. Increased levels of HERV-K transcripts have also been detected in rheumatoid arthritis (RA) blood and synovial tissue (46, 47). The resulting immune response is discussed in more detail in section *Autoantibodies Against Retroviral Proteins in Autoimmunity*.

From the perspective of autoimmune diseases like RA and SLE, which are strongly female-biased, it is interesting to note that the 5' LTR of intact HERV-K loci contain many binding motifs for estrogen- and progesterone-regulated transcription factors. Indeed, these hormones can upregulate transcription many-fold (48). We have replicated this finding (unpublished).

HERV-K transcription is also increased by cigarette smoking (49, 50), another risk factor for RA (51, 52).

HERV-E and Other HERVs of Potential Significance

A body of literature describes findings of increased expression of HERV-E in autoimmune disease, particularly SLE (53), as well as the presence of autoantibodies against HERV-E proteins, such as p30 encoded by its *gag* gene (54, 55). HERV-E derived Env protein can be detected in psoriatic skin (56). Compared to healthy individuals, HERV-E mRNA is reportedly increased in T cells from SLE patients, its LTR is hypomethylated, and further expression can be induced by demethylating agents and UV irradiation (57, 58). It has been proposed that autoantibodies against p30 cross-react with class I HLA (55).

A detailed survey of transcripts derived from over 8,000 retroviral sequences in our genome by Akiko Iwasaki and her team (59) found that a large portion of HERV loci are transcribed in transformed cell lines and many also in cells from patients with SLE. Compared to healthy controls, a number of transcripts were more, and a few less, abundant in SLE patients. The overexpressed loci included several HERV-K, HERV-E, HERV-W, and ERV3 loci. The use of computational tools, such as ERVmap designed by these authors (59), or others (60), have begun to uncover the full complexity of this topic. Case in point: the number of spliced and processed mRNAs derived from all the HERVs theoretically rival those of the traditional genes in numbers. If one also includes all retrotransposon transcripts, which sometimes are derived from intronic or 5' and 3' UTRs of protein-coding genes, the overlap and complexity becomes truly challenging. An important question is which of all these retroviral and retrotransposon transcripts are, in fact, translated into polypeptides that may have consequences for health and disease? Do retroelement transcripts matter if they are not translated? What are the consequences of the production of various retroelement-encoded polypeptides? How might the numerous single nucleotide polymorphisms and other types of polymorphisms within HERVs and retrotransposons affect human health?

Non-LTR Retrotransposons—LINE-1 and Alu Elements

The L1 element (61) appears to represent a remnant of an ancient retrovirus that retained, or later acquired, a degree of autonomy through the conservation of a primordial RT, which endows it with the ability to transpose without having to leave the host cell. This mechanism has been extraordinarily successful over evolutionary time and L1 sequences now occupy 17% of our genome (12, 62–64). While most of the ~500,000 L1 copies are mutated and truncated, some ~180 copies are seemingly intact and a handful of them remain fully active today (65), *i.e.*, they continue to retrotranspose by the L1 “copy-and-paste” mechanism (**Figure 2**), occasionally disrupting genes or regulatory regions by novel insertions (66).

Full-length L1 is a 6-kb sequence with two open reading frames (ORF1 and ORF2) that encode for two proteins: the 40-kDa RNA binding ORF1p protein, and the 149-kDa ORF2p, which has both reverse transcriptase (RT) and endonuclease

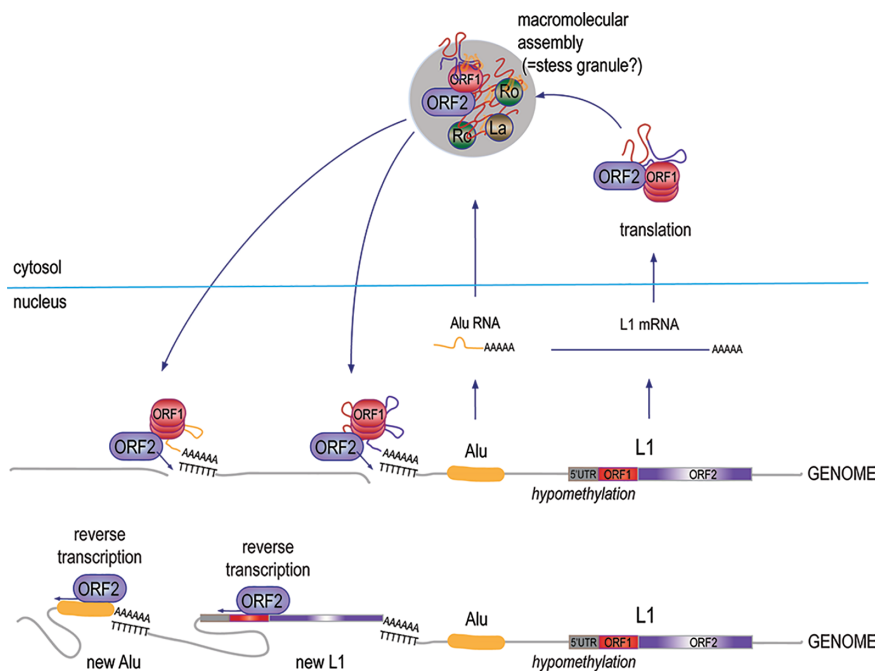


FIGURE 2 | The biology and replication cycle of the L1 element. Approximately, 500,000 copies of the L1 element, most of them truncated and mutated, exist in the human genome across all chromosomes, both within introns of protein-coding genes and in intergenic regions. Transcripts of the few 6-kb full-length L1 loci that retain intact open-reading frames are translated into the p40/ORF1p and p149/ORF2p proteins, which assemble in approximately a 20:1 stoichiometry into complexes with high affinity for RNA, particularly L1 mRNA, as well as Alu RNA. The complexes also include at least 10 other proteins, including Ro60 and La. Under permissive conditions, the RT activity of ORF2p makes a DNA copy of the associated RNA and inserts it into the genome, resulting in a new L1 element, a new Alu element, or a new pseudogene of an accidentally captured other mRNA.

activities. Retrotransposition occurs through a “copy-and-past” mechanism, where the primary transcript is captured by ORF1p, and is then reverse-transcribed by the RT domain of ORF2p primed by a nick made by the endonuclease domain of ORF2p in a genomic poly-T sequence, where the 3’ poly-A of the L1 RNA can anneal. The DNA repair machinery then patches up the 5’ end of the new insert. From time to time, ORF1p and ORF2p grab the wrong RNA, resulting in the reverse transcription and genomic insertion of an Alu element (67) or a spliced mRNA giving rise to an intron-less pseudogene (63). Even if most pseudogenes are inactive, this mechanism may have created genomic diversity and new material for natural selection to work with. Similarly, ORF2p-mediated reverse-transcription and insertion of a primary (un-spliced) mRNA would result in gene duplication; this may be how gene families were created over evolutionary time.

The RT of L1 ORF2p shows clear, but relatively distant, homology with the RT of the HERVs, suggesting that they all originate from a common ancestral RT. The L1 RT is closer in homology to the RT of hepatitis B virus, a “para-retrovirus”, that uses reverse transcription to generate a circular DNA that is not integrated into the host genome. Insects, like *Drosophila*, have retrotransposons (68). Yeast also have retrotransposons similar to L1, called Ty3 (69), as do baculoviruses (70), and prokaryotes

in the form of reverse-transcribing bacteriophage (71). Retrotransposons in plants (72) have been called “engines of evolution” (73). All of these examples attest to the truly old roots of these sequence elements.

There is evidence that L1 loci are transcriptionally active in SLE patients (74–76). This appears to correlate with a global decrease in DNA methylation, which is well documented in SLE (77, 78) and likely relates to the decreased expression of DNA methylases DNMT1 and DNMT3a (79, 80). Demethylating agents like 5-aza-2’deoxycytidine (81) also cause a dramatic upregulation of L1 and Alu element transcription in lymphocytes (82). In addition, transfer of 5-aza-2’deoxycytidine-treated T cells into healthy mice results in an SLE-like disease (83). The drugs that can induce “drug-induced lupus”, notably hydralazine and procainamide, are also demethylating agents (84). Other known triggers of lupus flares, like UV light, oxidative stress, inflammation and exogenous viruses also induce genomic hypomethylation (85, 86). L1 transcription can also be stimulated by female hormones, which further supports the notable female predilection of SLE and RA.

Among the many reasons to suspect that L1 plays a role in autoimmune disease is the observation that ORF1p resides mostly in macromolecular assemblies (Figure 2) that have been proposed to be stress granules (87) together with L1 RNA

and other RNA-binding proteins (88–91), such as Ro60, La (92), and U1 small nuclear ribonucleoprotein of 70 kDa (88), all well-established autoantigens in SLE and related diseases. The protein complexes also contain ORF2p and perhaps newly synthesized DNA made by its RT activity. If released from broken or dying cells, such protein complexes containing two apparently immunogenic proteins as well as both RNA and DNA would likely be of great interest to the immune system.

DEFENSE MECHANISMS AGAINST RETROVIRUSES AND RETROTRANSPOSONS: OUR ORIGINAL IMMUNITY

In support of the notion that incoming integrating retroviral genomes and autonomously retrotransposing sequences have posed serious threats to the integrity of host genomes since the early days of cellular life (despite also contributing positively to evolution), an elaborate set of defense mechanisms against retroviruses and retroelements are present in all cells (93–98). Many of these mechanisms were discovered in the course of HIV research as “restriction factors”. It also appears that many of these mechanisms continue to be critically important for human health. Conversely, many prevalent diseases, including cancer and autoimmunity, may be related to incomplete function of these mechanisms (61, 99).

All Roads Lead to Type I IFNs

Type I IFNs are a central hub of antiviral immunity (100). Therefore, it is not surprising that many of the defense mechanisms that cells use against retroviruses and retrotransposons also center on the induction of type I IFNs. Nevertheless, one should keep in mind that many defense mechanisms also have direct anti-viral functions and that many induce other pathways too. Interferons often play an amplifying role and increase the expression of these defenses in a positive feedback loop.

The main threat of a virus is its RNA or DNA genome, which will hijack the cellular biosynthetic machinery for its own replication and virion production, with detrimental and often lethal consequences for the host cell. Even more alarming, retroviruses will reverse transcribe their RNA genome and insert the resulting DNA into the host genome as a permanent provirus. To counteract these ancient foes, evolution has produced several cellular mechanisms for the detection of non-self RNA and DNA (11). Five principal pathways operate in the cytosol and on the cytosolic surface of intracellular organelles: the DNA-sensor “cyclic GMP-AMP synthase” (cGAS) (101), the RNA sensors “retinoic acid-inducible gene I” (RIG-I) (102) and “melanoma differentiation-associated gene 5” (MDA5) (102–104), and the two kinases “protein kinase RNA-activated” (PKR) (105, 106) and DNA-activated protein kinase (DNA-PK). A sixth pathway responds to extracellular DNA or RNA brought into the cell by receptor-mediated endocytosis and is initiated by Toll-like receptors (TLRs) 3, 7, 8, and 9 in the

endosome. A mechanism for cross-talk of the extracellular and intracellular sensing pathways consists of the transporter protein SIDT2 (107), which channels dsRNA through the endosome membrane into the cytosol, where it can trigger MDA5. There are several additional recently discovered DNA and RNA sensors, such as DDX1, 21, 36 and 41, IFI16, and Aim2 (108). All these pathways primarily promote type I IFN production through activation of IRF3 and related transcription factors. Some also activate signaling pathways that lead to the production of other cytokines. The resulting type I IFNs are secreted, bind to the type I IFN receptor, and signal through the JAK/STAT pathways to upregulate the expression of proteins with direct anti-viral activity, including nucleases, helicases, chaperones, and many of the sensors and their adapters and signaling proteins (100). Another important effect is the upregulation of MHC molecules to facilitate the recognition of the virally infected cell by cytotoxic T cells.

Most patients with SLE (or related diseases) have elevated levels of type I IFNs (3, 4, 109), which is best detected as the high expression of IFN-stimulated genes (ISGs), now referred to as the “IFN signature” and seen in 70–90% of SLE patient populations world-wide (5, 110–113), as well as in patients with pSS (10, 113), systemic sclerosis (114, 115), polymyositis (PM) and DM (7, 116), and in a small subset of rheumatoid arthritis (RA) (117, 118). The elevated IFNs include not only IFN α and IFN β , but also the less known IFN ϵ , IFN κ , and IFN ω , as well as type II IFN (IFN γ) (10), and type III IFNs (IFN γ 1, IFN γ 2, and IFN γ 3) (119), which collectively appear to play an important role in pathogenesis (2, 120, 121). Type I and III IFNs are functionally overlapping (all genes induced by type III IFNs are also induced by type I IFNs), but IFN γ is instrumental in a distinct aspect of the immune system, namely the activation of T helper 1 cells, cytotoxic T cells, natural killer (NK) cells, and other elements of a general immune response. Nevertheless, some 900 of the 1,300 ISGs induced by IFN γ are also induced by type I IFNs, which induces a total of over 1,500 ISGs, suggesting much overlap in downstream consequences.

Type I IFNs have a spectrum of effects on the immune system and beyond, particularly upregulating numerous aspects of anti-viral defense. They stimulate emergency myelopoiesis (122), monocyte differentiation into myeloid dendritic cells (123, 124), antigen presentation, cytotoxic T cell differentiation (125), and B cell differentiation into plasma cells (126). These hallmarks of anti-viral immunity also characterize SLE and other autoimmune conditions.

DNases, RNases, and Aicardi-Goutières Syndrome

To neutralize dangerous DNA or RNA, cells express a number of DNases and RNases, the function of which also prevent untimely triggering of DNA and RNA sensors. Remarkable insights into the dynamic biology behind these processes was gleaned from studies of the monogenic disease known as Aicardi-Goutières syndrome (AGS) (127–132), in which loss-of-function mutations in any one of a number of enzymes lead to constitutively high production of type I interferons (IFNs), neurological deficits due

to IFN toxicity, and autoimmunity that resembles SLE very much. Loss of the cytosolic DNase *TREX1* (99, 129, 133, 134) causes the accumulation of non-chromosomal DNA made by L1 ORF2p (135, 136), while mutations in any of the three subunits of *RNASEH2* (129, 132) cause the accumulation of DNA:RNA heteroduplexes made by ongoing reverse transcription (132). Another AGS gene, *SAMHD1* (137, 138), directly counteracts reverse transcription by dephosphorylating the required deoxy-nucleotide triphosphates. Together, these defects show that IFN-driving aberrant DNA apparently results from reverse transcription of cellular RNAs at a surprisingly high spontaneous rate. The only cellular enzyme capable of this reverse transcription is the ORF2p of L1, which is a highly efficient RT (61, 139, 140). IFN production (141) triggered by L1 can use many cellular RNA templates, including its own mRNA (63, 64) or Alu transcripts, to generate DNA species that drive the interferon production pathway. This mechanism also operates in cellular senescence (142).

In a mouse model of AGS, the *Trex1*^{-/-} mouse (99), the animals develop a systemic inflammation with immune cell infiltrates in many organs and they die early from a severe carditis. These animals can be rescued from death by treatment with the RT inhibitors tenofovir plus nevirapine (143), indicating that reverse transcription is a key step in the pathogenesis of systemic inflammation in this model. However, there is also a published paper refuting these data (144). More importantly, the IFN signature can be reduced substantially in AGS patients by RT inhibitors used for the treatment of HIV (145).

Factors That Reduce Retrovirus Infectivity and Retrotransposition

The fact that the vast majority of HERVs and retrotransposons have been rendered largely inactive and/or harmless (to the best of our knowledge) attests to the power of the spectrum of defensive mechanisms employed by cells both acutely and over evolutionary time. The default acute mechanism employed by cells to silence unwanted or dangerous genes is the modification of deoxycytosine in DNA by methylation. This modification also facilitates the addition of suppressive histone H3 K4-dimethyl marks to keep these loci transcriptionally silent. In this context, it is interesting to note that many of the drugs notorious for causing “drug-induced lupus”, such as hydralazine and procainamide, are demethylating agents. Experimentally, 5-aza-deoxycytosine can also be used to reduce the methylation of the genome. This also causes an increase in the expression of L1 and many HERVs. Ultraviolet light (UVB) also reduces genomic methylation, likely with the same de-repression of retroelement transcription. UVB is also a well-recognized trigger of lupus flares.

A good example of retrotransposon control is seen with the large number of interrupted retrotranspositions of L1 in which the reverse transcription was terminated before it reached the 5' end. As a result of this, many L1 copies lack portions of ORF1 or the 5'UTR regulatory region and cannot retrotranspose. Many of the mutations seen in HERVs and L1 may have been deliberately introduced by the APOBEC family (146) of IFN-inducible cytidine deaminases, which recognize viral or retroviral

sequences and rapidly introduce mutations into them (147). This mechanism has been shown to be effective at reducing the virulence of new retroviruses (148, 149) and the ability of retrotransposons to replicate (150).

Another example is Moloney leukemia virus 10 (MOV10), an ATP-dependent helicase that unwinds L1 RNA during reverse transcription to reduce the retrotransposition (151–153) process in a somewhat unclear manner (95). It also participates in the defense against retroviruses (154). MOV10 is located in the macromolecular complex of RNA-binding proteins that also includes L1 ORF1p and ORF2p, as well as SLE autoantigens Ro60 and La (89–91).

There are numerous additional cellular mechanisms to counteract each step of the retrovirus life cycle and the retrotransposition of repetitive elements (155). Many of these mechanisms also serve as defenses against other types of viruses and many of them were uncovered in the course of HIV research (98). It is presently not known if any of these mechanisms are compromised in patients with SLE or other autoimmune conditions.

RNA Interference, Argonaute, Piwi, and Other Nucleic Acid-Based Defenses

Many prokaryotes employ an interesting defense mechanism in which short pieces of the genetic material of past pathogens are kept in a region of the genome to serve as recognition modules for the defense against reinfection. This mechanism (known as CRISPR/Cas9) has an RNA-based counterpart in eukaryotes, including humans, in the form of RNA interference mechanisms that utilize retrotransposon-derived miRNAs, piRNAs, and potentially antisense transcripts from HERVs. Particularly, the piRNA/Piwi pathway appears to be important for protecting the integrity of the germline genome (156). One of the effectors of these still incompletely understood pathways is Z-DNA binding protein-1 (ZBP-1) (157), also known as DAI, which binds dsRNA or DNA that adopt the Z-conformation. ZBP-1 then activates both the IRF3 pathway for type I IFN production and the RIP3 kinase pathway that triggers cell death by necroptosis (158). The relevance of ZBP-1 in Crohn's disease (159) and other inflammatory conditions (158) was recently demonstrated.

MOLECULAR MECHANISMS BY WHICH ENDOGENOUS RETROVIRUSES AND RETROTRANSPOSONS COULD CAUSE IMMUNE PATHOLOGY LEADING TO AUTOIMMUNITY

A number of mechanisms have been proposed over the years for how HERVs or L1 could cause diseases like cancer and autoimmunity. In cancer, it is thought that some combination of active retrotransposition catalyzed by L1 ORF2p (160), genomic recombinations caused by highly similar repetitive sequences (e.g., LTRs), or promoter/enhancer effects of HERV LTRs or L1 5' UTRs can create gross chromosomal abnormalities, tumor suppressor disruptions, and loss of

normal transcriptional control. In autoimmunity, other aspects of HERV and L1 biology are probably more relevant.

It should be stated upfront that despite the many genetic, experimental, and supportive correlative findings, it is still possible that none of the myriad of HERV and L1 sequences that constitute at least 25% of our genome (close to 40% if one includes Alu elements and other SINEs) have any role at all in autoimmunity because they have been sufficiently “domesticated” and have lost their immunogenicity and potential to raise an anti-viral response or to skew biological processes in any meaningful way. That said, there are many plausible aspects to the general hypothesis that HERVs and/or L1 can promote or even trigger autoimmunity (17, 161, 162). We believe that the vast majority of retroelements are harmless, some even beneficial (17), but that a select few are dangerous and participate in the pathogenesis of common autoimmune diseases like SLE or RA by mechanisms that are discussed below.

Autoantibodies Against Retroviral Proteins in Autoimmunity

In the 1990s, a number of researchers made the surprising discovery that serum immunoglobulins from patients with RA, SLE, or other autoimmune diseases, reacted with Human Immunodeficiency Virus (HIV) proteins, *e.g.*, p24 of the HIV capsid (163–166), even if these patients had never encountered the virus. Such HIV-reactive antibodies were found in exceedingly few healthy subjects, but reportedly in up to 60% of RA patients. A likely answer to this conundrum was provided by the subsequent discovery (167) that endogenous retroviruses in the human genome, particularly HERV-K (168), are transcriptionally activated in some RA patients (46, 169). This raised the possibility that HIV-reactive antibodies in patients are, in fact, antibodies against HERV proteins that have a sufficient degree of sequence homology with HIV proteins. Indeed, two papers (170, 171) reported that 16% of RA patients have antibodies against an epitope in the HERV-K envelope protein (amino acids 19–37). It should be noted that the percentage of positive patients in the earlier papers was higher, presumably because the tested antigens were full-length proteins in their native state, while later papers mainly used selected peptides and therefore may have missed many autoantibodies.

We have replicated the detection of elevated IgG autoantibodies against HERV-K Env proteins (not peptides) in RA patients (submitted for publication). These antibodies were also present in pediatric patients with juvenile idiopathic arthritis (JIA), and they were higher in smokers than in non-smokers. Anti-Env antibodies were also detected in some control (*i.e.* non-RA) individuals, in SLE patients (172), and in patients with breast cancer (173, 174) and other hormonally driven cancers. Several HERV-K loci are reportedly transcriptionally active in these cancers (175–177), perhaps through the action of sex hormones on the HERV-K LTRs (48). Notably, RA is a female-biased disease with a 4:1 female-to-male ratio. Autoantibodies have also been reported against proteins of HERV-E, particularly the Gag protein of HERV-E clone 4-1 (178). We would not be surprised if it was found that patients have autoantibodies against additional HERV proteins.

Could Autoantibodies Against Retroviral Proteins Be Directly Pathogenic?

The presence of anti-HERV autoantibodies in patients with autoimmune diseases prompted many researchers to wonder if HERVs play a role in autoimmunity. One proposed mechanism to connect these autoantibodies to autoimmune disease was the hypothesis of “molecular mimicry”, postulating that the relevant epitopes for these anti-HERV autoantibodies may be sufficiently similar to amino acid motifs in self-proteins to cross-react with such *bona fide* autoantigens and result in autoimmunity. In our view, this hypothesis (which has also been proposed for exogenous viruses) seems rather unlikely and it is not supported by patient-based data. While a few instances of three to four identical amino acid residues can be found in HERV proteins and proteins like collagen or IgG, these were not shown to be epitopes for autoantibodies. This hypothesis also assumes that retroelement proteins are not “self”, in contrast to proteins encoded by traditional genes, and that humoral or cellular immunity only against the latter could be pathogenic. In addition, central tolerance against the relevant epitopes in traditional self-proteins should automatically also prevent the same sequence from being immunogenic when present in a different class of self-proteins, *e.g.*, HERVs.

Another simple hypothesis focuses on the plasma membrane location of Env proteins. When expressed, these transmembrane glycoproteins cluster into microdomains together with intracellular Gag to form virions that eventually bud off to leave the host cell. During this time, autoantibodies against Env would be predicted to bind the exposed Env with potential consequences like complement fixation or antibody-dependent cellular cytotoxicity (**Figure 3**). In both cases, the cells expressing Env can be killed under circumstances that would be pro-inflammatory. In this scenario, inflammation would follow the same pattern as in antiviral immunity, except that the offending virus is a HERV and the response would be that of autoimmunity. Since the HERV in question is irrevocably fixed in the host genome, it cannot be eradicated by the immune response and any cell that subsequently expresses Env would be treated as a virally re-infected cell by a recall immune response. This could well result in a pattern akin to what we see in clinically relevant autoimmune disorders.

Could Protein Citrullination be Linked to Retroviral Proteins?

Among the autoimmune diseases, RA is unique in that post-translational deimination of arginine residues, also known as citrullination, plays an important role in creating autoantigens. While a low amount of citrullination is a part of normal physiology, much elevated levels of citrullination of proteins that perhaps never are citrullinated at all in healthy individuals can be induced by a process termed “lethal hypercitrullination” or “leukotoxic hypercitrullination” (179, 180). This reaction is induced by any agent that creates pores in the plasma membrane that are large enough to allow Ca^{2+} to rush into cells, such as the membrane-attack complex of activated complement, polymerized perforin from cytotoxic T cells or NK cells, or

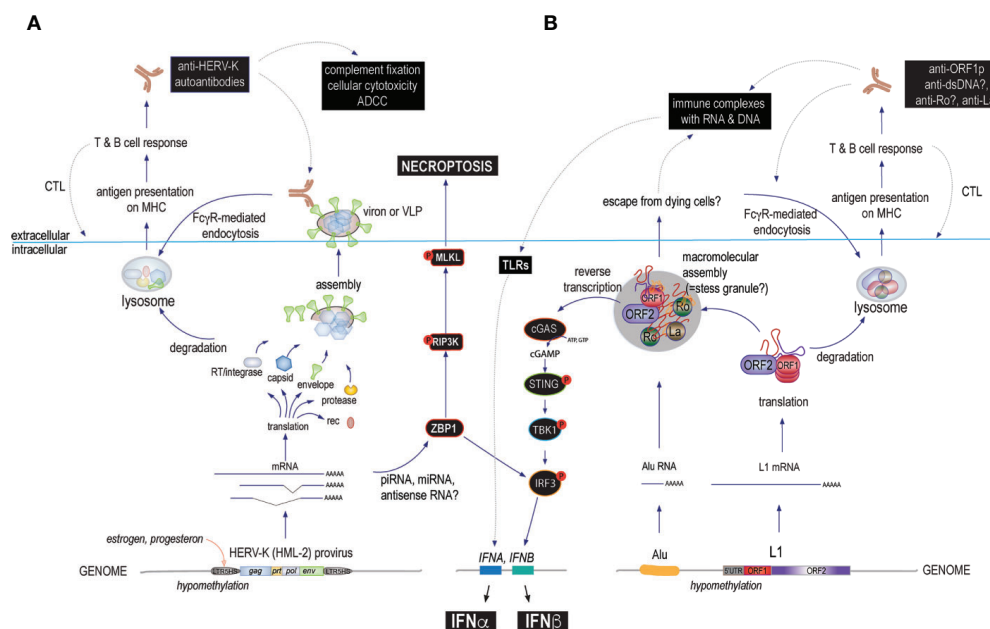


FIGURE 3 | The main proposed mechanisms by which HERVs and L1s may cause immune responses that could escalate to autoimmune disease if they become chronic or recurring. **(A)** starting at the bottom (intracellularly): environmental or internal factors first reduce the suppressive DNA methylation of 5' LTRs allowing transcription factors such as those regulated by female hormones to actively transcribe one or several HERV loci. The resulting transcripts are spliced and processed and some of them are translated into proteins, while others can associate with antisense RNA, small interfering RNAs of the miRNA or piRNA classes, or form internal loop structures that are recognized by ZBP-1 (or other sensors or RNase-based enzymes), which can signal through IRF3 to IFN production or, alternatively, via the RIP3 kinase pathway to cell death by necroptosis. HERV proteins can assemble into more or less complete virions that may remain exposed on the cell surface or even bud off as mature virions. These proteins can be degraded and processed for antigen-presentation on class I or II MHC molecules to activate T cells and B cells to generate both cellular and humoral immunity. Both arms target cells expressing the relevant HERV for immune attack. **(B)** L1 may drive a similar immune response, except that the ORF2p RT can generate intracellular DNA to directly trigger cGAS or other DNA sensors resulting in type I IFN production (primarily IFN β). Another unique feature of L1 biology is the assembly of the two L1 proteins with other RNA binding proteins, many of which are well-known autoantigens, into aggregates that also contain L1 RNA and other RNA species, such as Alu element RNA or processed mRNAs. These bodies may also contain DNA newly synthesized by the ORF2p RT and hence will even more resemble virions if they are released from dying cells. This would result in an immune response and autoantibodies against several of their components, such as Ro60, in addition to the L1 proteins themselves. Lastly, such autoantibodies would further promote the uptake of L1 protein/RNA/DNA complexes by Fc γ R-expressing immune cells, including plasmacytoid dendritic cells and further stimulate type I IFN production (primarily IFN α) through TLR pathways. In all these scenarios, the repeated boosting of anti-HERV or anti-L1 immune responses would lead to increasingly powerful immune-mediated destruction of the cells that express them. Disease-relevant pathways are indicated with black boxes.

certain bacterial toxins like α -toxin from *Staphylococcus aureus*. Among immune cells, neutrophils are the most prone to undergo a strong hypercitrullination reaction (180). Hence, if the neutrophils express Env in an RA patient with anti-Env autoantibodies, these cells could well be killed by complement or by Env-specific cytotoxic T cells. This hypothetical model of neutrophil killing and hypercitrullination linked to anti-HERV immunity is under investigation in our laboratory.

Autoantibodies Against Retrotransposon Proteins and Associated Proteins

We recently reported that SLE patients also frequently have autoantibodies against the L1-encoded ORF1p protein (181), which is physically associated with Ro, La, snRNP70, and other well-known SLE autoantigens (87, 88, 91, 182) together with RNA in macromolecular assemblies (which may be stress-granules). Indeed, SLE autoantibodies recognized several other proteins in purified ORF1p-containing complexes (181). Furthermore, anti-ORF1p titers correlated with SLE disease

activity, lupus nephritis, anti-dsDNA levels, and complement consumption (181). They were also present in pediatric lupus patients with newly diagnosed disease (our unpublished observation). In contrast, RA patients were negative for anti-ORF1p antibodies with only sporadic exceptions.

At this time, we do not know whether anti-ORF1p autoantibodies are pathogenic *per se*. Because ORF1p is intracellular, and hence out of reach for extracellular antibodies, we are inclined to believe that they do not contribute directly to pathology in SLE and merely reflect the expression of immunogenic ORF1p. However, it is likely that they form immune complexes with extracellular ORF1p, which could escape from dying or broken cells (Figure 3). If so, intact ORF1p likely would have bound RNA and exist in a complex with Ro60, La, and other proteins that it associates with in the cells. Such immune complexes could induce type I IFNs upon internalization by plasmacytoid dendritic cells and could promote antibody responses against all the proteins and nucleic acids in these immune complexes. In support of this

possibility, isolated anti-Ro60 immune complexes from SLE patients were shown to contain RNA from both Alu elements and L1 (182).

Direct Pathogenic Effects of Proteins Encoded by HERVs and Retrotransposons

After the initial discovery of autoantibodies reactive against HERV-encoded polypeptides in the 1990s, researchers began to search for the corresponding mRNAs and proteins in a variety of cells and tissues, such as RA synovium (183, 184), initially using degenerate primers and little insight into the sheer multitude of retroviral and retrotransposon sequences in our genome. With the sequencing of the human genome (12), the complexity of the issue became more obvious. To this day, there are relatively few commercially available reagents to detect HERV or L1 proteins and only a limited number of computational tools to analyze retroelement transcription profiles in RNA-Seq data sets. Furthermore, it is now clear that large numbers of retroelement loci are transcribed even in healthy individuals (59), including many that probably do not encode for proteins or only for short peptides if translated at all.

The presence of autoantibodies against HERV and L1 proteins suggests that they are at least moderately immunogenic. An important unanswered question is to what extent retroelement-encoded polypeptides are expressed in the thymus and bone marrow during T and B cell antigen receptor repertoire selection and the formation of central tolerance. From the acute use of extensive DNA methylation to silence the transcription of unwanted and potentially dangerous sequences, we project that the youngest HERVs are least likely there and that tolerance against many of them is weak or absent.

In agreement with this notion, a recent paper (185) demonstrated that pancreatic islets in non-obese diabetic (NOD) mice (but not control mice) release microvesicles, which contain endogenous retroviral Gag and Env proteins, probably in the form of complete or near-complete virions. The NOD mice developed antibodies against these proteins, as well as specific T cells, which caused diabetes when adoptively transferred. Elimination of Gag prevented diseases. These data show that abnormal activation and expression of endogenous retroviruses can trigger an anti-retroviral immune response and autoimmunity (185). This scenario is also depicted on the left side of **Figure 3**.

Besides acting as antigens for the host immune response, the proteins and peptides encoded by HERVs and L1 have other properties that could be relevant (17). For example, retroviral Env and Gag proteins can mediate cell fusion events, while mature and processed Env of HERV-K, which form transmembrane trimers, binds to heparan sulfate-containing surface proteins (186). Hence, aberrant expression of these proteins may cause pathological fusion and adhesion events that could prove problematic. In addition, certain Env portions may have immunomodulatory effects.

The “superantigen” encoded by the *env* gene of HERV-K18, which was incorporated into the human genome 7.8–14.4 million years ago and has accrued a number of amino acid

substitutions in its *env* gene, made a big splash in the field when its T cell activating properties and presence in type 1 diabetes patients was first reported in *Cell* in 1997 (187). The proposition was that this protein causes a polyclonal T cell activation, which then leads to autoimmunity. Subsequent papers were not in full agreement (188) and despite subsequent papers finding this “superantigen” expressed in JIA patients (189) and that it is inducible by Herpes viruses (190), the interest in HERV-K18 slowly waned. Nevertheless, the lesson from this specific case is that random mutations may not only reduce the ability of a retroviral component to cause pathology, but may also, by chance, give them new and dangerous properties. Certain sequences in the Env protein also appear to have immunosuppressive properties. How such motifs might act and whether they were important for the life-cycle of free infectious HERV-Ks are not known.

The L1-encoded ORF1p and ORF2p proteins have been detected in samples from patients with SLE or pSS (74) by immunoblotting and immunohistochemistry. ORF1p was present in kidney biopsies from lupus nephritis patients and in salivary gland from pSS patients. This staining coincided with IFN β in glandular cells and with IFN α in the infiltrating immune cells. As activation of L1 elements in autoimmune patients (74–76) appears to involve demethylation of the 5' UTR (74, 76, 78), these authors also analyzed the methylation of CpG sites in this regulatory region and found it to be reduced in patients with elevated L1 expression. L1 ORF2p is also present in the ductal cells of salivary gland biopsies from patients with pSS (191).

We have detected ORF1p protein in neutrophils of juvenile and adult SLE patients, including in low-density granulocytes from these patients (submitted for publication). These findings are compatible with the emerging role of the neutrophil as a cell type of interest in SLE pathogenesis.

Could Transcripts From Proviruses and Retroelements Be Pathogenic?

While primary transcripts from endogenous retroviral or retrotransposon loci are synthesized by the same machinery that transcribes and processes other genes and therefore should be indistinguishable from any other cellular transcripts, two recent papers published in *Nature* (158, 159) showed that they have a propensity to form double-stranded structures, perhaps through transcription of the complementary strand as well. They found that such dsRNAs are recognized by the host defense protein ZBP-1, which binds both DNA and RNA in their Z-conformation. This binding activates ZBP-1 to trigger type I interferon production through IRF3, as well as activation of the RIP3 kinase pathway leading to cell death by necroptosis. Necroptosis, in turn, is a very immunogenic process and leads to autoimmune disease (158, 159).

Another possibility is that some species of retroelement RNA may have features or motifs that somehow resemble viral RNA and therefore are recognized by the RNA-sensors RIG-I or MDA5 (192), a challenging task given the abundance of cellular RNA species. Antisense sequences (transcripts from the other DNA strand) could provide the answer. The delicate

balance between the recognition of self- versus foreign RNA is well illustrated by the *IFIH1* A946T allele, which encodes a variant of MDA5 that enhances anti-viral immunity, but increases the risk of autoimmunity (193, 194).

It is curious that some retroelement RNAs like Alu transcripts need to be edited by adenine deamination catalyzed by ADAR1 to remain harmless (195). In the absence of this editing (e.g., loss of ADAR1), these RNAs form double-stranded hairpin loops that trigger the RNA sensors leading to AGS (196). There is also evidence that some 30% of SLE patients have constitutively activated RNA sensors, detectable as an aggregation of the downstream mitochondrial antiviral signaling (MAVS) adaptor protein (197). It is not known what RNA species was responsible for this activation in the patients.

Reverse Transcriptases, DNA Sensors, and Type I Interferons

A mechanism with clear potential for pathogenicity is the conversion of retroelement RNA into extrachromosomal DNA by reverse transcription. If not rapidly degraded by the DNase TREX1 (or other DNases), such aberrant DNA will trigger DNA sensors like cGAS, which in turn drive the production of type I IFNs. This is apparently what happens constitutively in patients with loss-of-function mutations in *TREX1*. It also appears to occur in at least a subset of SLE patients: the second messenger 2'3'-cyclic-guanosine-adenosine-monophosphate (cGAMP), which is synthesized exclusively by cGAS upon DNA binding, was detected by mass spectrometry in 7 of 30 SLE patients (198). While it may seem that this represents a small portion of SLE patients, it is important to recognize that the data represent a single snap-shot in time for each patient and that cGAMP is a short-lived second messenger present in minute quantities. Thus, it may well be that cGAMP is periodically elevated in many more SLE patients than reported.

While the DNA species that triggers cGAS in SLE patients remains unknown, there are only two likely possibilities, as we have discussed before (11): mitochondrial DNA or DNA synthesized by a cellular RT. There are only three types of RTs in our genome: telomerase (*TERT*), the RTs encoded by the *pol* genes of HERVs, and ORF2p encoded by L1. Telomerase only synthesizes TTAGGG repeats in the ends of our diploid chromosomes using the *TERC* RNA template (199, 200), while retroviral RTs supposedly only convert the RNA genome of an incoming retrovirus to a DNA provirus that is inserted into the genome during acute infection. Hence, L1 ORF2p is the most likely to produce DNA that can trigger type I IFN production through cGAS activation in SLE patients. ORF2p has robust RT activity (139, 140, 201), which is key for retrotransposition (202), and is sensitive to some clinically used RT inhibitors (203, 204).

Possible Roles of Genomic Alterations Resulting From Retrotransposition in Autoimmunity

Lastly, a unique potential mechanism by which retrotransposons, and perhaps also HERVs, could impact human health is by

retrotransposition, *i.e.* by inserting a brand-new reverse-transcribed copy into a new genomic location. This can occur early in embryogenesis (205) when the genome is broadly hypomethylated and extensively transcribed, including retroelements of all classes. During this time, RT activity is high, extrachromosomal DNA is readily detectable, and L1 elements are capable of active retrotransposition (206). In fact, more than a hundred novel genetic diseases have been found to be caused by L1 retrotranspositions into vital genes (207), disrupting their regulation or function. It is also clear that active L1 retrotransposition occurs in certain neurons during development of the central nervous system (208) and that this creates somatic mosaicism of unclear neurological relevance (209). It is also possible that similar L1 retrotranspositions could occur in immune cells to generate T cells with abnormal behavior, leading to autoimmunity. One could imagine that the disruption of the gene for an important negative regulator of immunity, for example in a hematopoietic stem cell, could result in populations of overly reactive T cells. While this is an interesting possibility, there is no evidence of it at this time other than in cancer.

A different type of genomic alterations also occurs in malignant cells, namely the recombination of retroelement loci that have a high degree of sequence identity but located in different chromosomes. The large number of single LTRs ("solo-LTRs") throughout our genome appear to be the result of such recombinations. Again, there is no evidence for this type of genomic alterations in human autoimmunity, but it might be worthwhile to search for them.

Malignancy-Related Autoimmunity

An immune response against a tumor involves the recognition of tumor-specific epitopes, which consist of point-mutations in common self-proteins, aberrantly spliced or modified proteins, or proteins that normally are only expressed during early embryonic development (*i.e.* carcinoembryonic antigens). Particularly when the immune response against the tumor is strong and succeeds in eliminating the malignant cells, there is an obvious risk of epitope spreading and further escalation into autoimmunity. There are numerous examples of such "paraneoplastic" syndromes and collateral damage, for example vitiligo (*i.e.* the killing of normal melanocytes) in patients with melanoma after a successfully boosted anti-tumor immune response.

Carcinogenesis typically involves a broad genomic demethylation and de-repression of many genes, including carcinoembryonic antigens, as well as numerous HERVs and L1s. It therefore seems very likely that proteins encoded by these retroelements serve as tumor-specific antigens. This notion is supported by the presence of anti-HERV-K autoantibodies and HERV-K-specific T cells in breast cancer (173, 174). If this response is successful and eradicates the (pre)malignant cells, the immunological memory of these antigens would readily serve to re-engage the immune system if the same proteins were expressed again in a non-malignant cell type. There is currently no clear evidence that patients with autoimmune

disease would have fought off a malignancy prior to developing their autoimmunity, but this possibility should be explored if possible.

TESTING THE HYPOTHESES

Resolving the question of whether HERVs and/or L1 retrotransposons contribute to human autoimmune diseases is a monumental challenge. Ultimately, only significant effects of therapeutic interventions in double-blinded, placebo-controlled human clinical trials will be able to conclusively prove causalities that have been inferred from molecular mechanistic experiments, correlative patient observations, and, perhaps, animal models. Disproving a proposed mechanism is even harder.

Regarding animal models, due to the approximately 80–100 million years of evolutionary distance between mice and humans, there are fundamental differences between our repertoires of retroviruses and retrotransposons. Most of the human HERVs are more recent, and do not exist in the mouse genome. Instead, mice still have many fully competent retroviruses, such as mouse mammary tumor virus (MMTV), which is distantly related to the HERV-K family. Disease models based on MMTV would be difficult to interpret. It is very interesting to note, however, that the autoimmunity-prone strains of mice (e.g. MRL, or NZB) have larger sets of active endogenous retroviruses than other strains. Mice also have numerous L1 insertions, albeit not quite as many as humans.

It will be important to construct a set of plausible molecular hypotheses of pathogenesis and to test them for supportive patient-based evidence, or, conversely, lack thereof. Since many diseases like SLE are clinically heterogeneous, one cannot assume that the molecular mechanisms that underpin them will be the same in every patient. Instead, it might be more productive to start with the assumption that each disease contains two or more “endotypes”, *i.e.* patients with a clinically similar disease, but with distinct molecular underpinnings. This concept is well accepted in oncology and is making inroads in respiratory diseases, where asthma is now understood to contain “eosinophilic asthma”, “Th2-high asthma”, “epithelial asthma”, and “allergen/IgE asthma”, each of which respond well to different targeted therapies. These endotypes of asthma are not possible to distinguish clinically, but require laboratory measurements of eosinophils, cytokines, or IgE to classify and treat.

To advance, and eventually elucidate, the role of HERVs and retrotransposons in human autoimmunity, some major tasks that should be undertaken include:

- A comprehensive cataloging of the expression of HERVs and retrotransposon transcripts that are differentially expressed in cells and tissues from a range of autoimmune and other diseases, with an emphasis of disease-to-disease comparisons, and a filtering of the data sets by the capacity of transcripts to be translated.
- A broader and more detailed characterization of which proteins encoded by HERVs and L1 that become targets of autoantibodies in patients with different autoimmune diseases.
- The use of high-sensitivity proteomics to determine when and where such proteins are present, including an analysis of whether they are expressed in the thymus.
- The identification of T cell epitopes on retroelement proteins and characterization of the relevant T cells.
- Identification of the exact nature and source of nucleic acid species that drive type I IFNs in patients with type I IFN gene signature.
- An evaluation of the roster of anti-retroelement defense mechanisms to identify disease-related deficiencies or abnormalities in patients.
- A thoughtful selection of testable drug targets based on plausible mechanisms, followed by the development of therapeutic molecules that will enable clinical trials with relevant pharmacodynamic measures (biomarkers) to ensure that the drugs have the desired biological impact before asking if they have therapeutic effects.

THERAPEUTIC IMPLICATIONS

Currently used therapeutic regimens for autoimmune conditions consist of more or less broadly immunosuppressive drugs, which often provide unsatisfactory efficacy and/or unacceptable safety concerns. The development of more efficacious drugs with more precision and therefore, hopefully, improved safety profiles is very challenging as long as our understanding of the molecular mechanisms that initiate and propagate these diseases is so incomplete. We anticipate that scientific discoveries and breakthroughs in the coming years will open up new avenues for the development of better new drugs (210). If some of the hypotheses discussed in this review have any merit, what might new drug targets look like?

An already tested opportunity is the direct inhibition of type I IFNs, which appear to be instrumental in the pathogenesis of SLE and related diseases (111, 211). These drugs have been both encouraging and disappointing. Anifrolumab, an antibody that blocks the type I IFN receptor (IFNAR1) used by all type I IFNs, met with a statistically significant efficacy in phase 2 clinical trials in SLE (212) and met its primary endpoint in one of two phase 3 trials. In contrast, two different antibodies that block IFN α alone (213), sifalimumab and rontalizumab, were efficacious only in a small subset of SLE patients. Together, these trials reveal that type I IFNs beyond the 13 isotypes of IFN α are important, at least in some patients. Furthermore, although anifrolumab was clinically more efficacious, and neutralized the IFN signature by over 90% in the treated patient population, it did not benefit all the included patients.

Inhibitors of protein kinases that mediate the signaling from IFNAR, such as the JAK-family kinase TYK2, or signaling pathways that lead to type I IFN production, like TBK1, or the

necroptosis-inducing RIP3K, could also prove therapeutic even if these pathways are in broader use. Another possibility is to intervene in the biology that produces the nucleic acid species that initiate type I IFN production. Because anti-viral immune responses that include production of type I IFNs also include many additional pathways (although many of them amplified by IFNs), such therapeutics may be more efficacious than IFNAR blockade. Drugs that stimulate DNases or RNases, augment their function, or prevent their negative regulation, or that inhibit the enzymes that produce the offending DNA or RNA, e.g. reverse transcription, could be therapeutic. Indeed, RT inhibitors are efficacious in IFN-driven systemic inflammation observed in the *Trex1*^{-/-} (a DNase) mouse model of AGS (143) and reduce the IFN signature in patients with AGS (145). RT inhibitors also eliminated all symptoms of the pSS-like diffuse salivary and lacrimal gland inflammation in HIV patients in a small clinical trial (214). It will be interesting to see if RT inhibitors would also benefit SLE patients.

The DNA- or RNA-sensors that are activated by aberrant or excessive nucleic acid species could also be targeted by inhibitors. Indeed, c-GAS inhibitors are under development in a growing number of companies, which is also true for drugs targeting RIG-I or MDA5. An obvious risk with these drugs is that they can compromise anti-viral immunity. This is also true for antibodies that block type I IFNs yet appears to be manageable.

As we learn more about the immunogenicity of retroelement proteins and how various retroelement transcripts act to trigger ZBP-1 or RNA sensors, additional drug targets likely emerge. However, it currently seems that most of the transcriptional regulation and transcript processing of retroelements is mediated by the same cellular factors that regulate and process our traditional protein-coding genes. Retroviral LTRs, for example, use a host of transcription factors that regulate numerous other genes as well. Nevertheless, it is possible that RNA-based therapeutic molecules could be designed that more selectively interfere with pathogenic sequences. This space is still totally unexplored.

DIAGNOSTIC AND PROGNOSTIC IMPLICATIONS

The heterogeneity in clinical manifestations of autoimmune diseases (215) often make them challenging to accurately diagnose and properly treat, and to determine their prognosis with any precision. This challenge has prompted the development of various high-sensitivity and -specificity clinical

and laboratory classification criteria and disease activity indices to help in the management of patients with these diseases. However, while these methods are reasonably objective and helpful in clinical trials, there is wide variability in their applicability owing to the often-seen discordance between laboratory evidence of immunologic activity and actual physical evidence of clinical disease activity.

This discordance between laboratory and clinical disease activity leaves plenty of room for improved or adjunctive diagnostic methods that could help in closing that gap, and perhaps also help to improve the prognostication of the diseases relative to treatment. Given the significant correlations of anti-HERV-K Env antibody with RA disease activity (170, 171) (also our unpublished data) and anti-L1 ORF1p antibody with SLE disease activity (181), the potential for their use as diagnostic and prognostic markers is not too far-fetched. By the same token, with regard to potential surrogate markers for innate immune system activity in disease flares (i.e. type I IFN production in SLE), a similar approach could be undertaken whereby IFN signature gene expression is routinely monitored. However, we understand that these principles are still experimental and thus require more investigations for the validation of their actual clinical utility.

AUTHOR CONTRIBUTIONS

TM and KU contributed equally to the writing of this review and share accountability for its content. All authors contributed to the article and approved the submitted version.

FUNDING

Our work is supported by NIH grants T32 AR007108 (to KU), R21AR075134 (TM), R01 AR074939 (TM), and R21 AR077266 (TM).

ACKNOWLEDGMENTS

We apologize to all those colleagues whose many important papers we could not cite due to space constraints. We thank Christian Lood, John LaCava, Felipe Andrade, Martin Taylor, Galit Alter, Dan Stetson, Keith Elkon, and other colleagues for providing valuable feedback and suggestions.

REFERENCES

- Bengtsson AA, Sturfelt G, Truedsson L, Blomberg J, Alm G, Vallin H, et al. Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies. *Lupus* (2000) 9(9):664–71. doi: 10.1191/096120300674499064
- Ronnblom L, Alm GV. A pivotal role for the natural interferon alpha-producing cells (plasmacytoid dendritic cells) in the pathogenesis of lupus. *J Exp Med* (2001) 194(12):F59–63. doi: 10.1084/jem.194.12.f59
- Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U.S.A.* (2003) 100(5):2610–5. doi: 10.1073/pnas.0337679100
- Crow MK, Kirou KA, Wohlgemuth J. Microarray analysis of interferon-regulated genes in SLE. *Autoimmunity* (2003) 36(8):481–90. doi: 10.1080/08916930310001625952
- Ronnblom L, Eloranta ML, Alm GV. The type I interferon system in systemic lupus erythematosus. *Arthritis Rheum* (2006) 54(2):408–20. doi: 10.1002/art.21571

6. Ronnblom L, Alm GV, Eloranta ML. Type I interferon and lupus. *Curr Opin Rheumatol* (2009) 21(5):471–7. doi: 10.1097/BOR.0b013e32832e089e
7. Baechler EC, Bauer JW, Slattery CA, Ortmann WA, Espe KJ, Novitzke J, et al. An interferon signature in the peripheral blood of dermatomyositis patients is associated with disease activity. *Mol Med* (2007) 13(1-2):59–68. doi: 10.2119/2006-00085.Baechler
8. Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML, et al. Activation of the type I interferon system in primary Sjogren's syndrome: a possible etiopathogenic mechanism. *Arthritis Rheum* (2005) 52(4):1185–95. doi: 10.1002/art.20998
9. Peck AB, Nguyen CQ. Transcriptome analysis of the interferon-signature defining the autoimmune process of Sjogren's syndrome. *Scand J Immunol* (2012) 76(3):237–45. doi: 10.1111/j.1365-3083.2012.02749.x
10. Hall JC, Baer AN, Shah AA, Criswell LA, Shiboski CH, Rosen A, et al. Molecular Subsetting of Interferon Pathways in Sjogren's Syndrome. *Arthritis Rheumatol* (2015) 67(9):2437–46. doi: 10.1002/art.39204
11. Mustelin T, Lood C, Giltiay NV. Sources of Pathogenic Nucleic Acids in Systemic Lupus Erythematosus. *Front Immunol* (2019) 10:1028:1028. doi: 10.3389/fimmu.2019.01028
12. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* (2001) 409(6822):860–921. doi: 10.1038/35057062
13. Andrianov BV, Reznik NL, Gorelova TV, Zolotova LI. The retrotransposon Tv1 forms infectious virus-like particles in some lines of *Drosophila virilis*. *Dokl Biochem Biophys* (2005) 400:76–9. doi: 10.1007/s10628-005-0037-3
14. Bohmdorfer G, Luxa K, Frosch A, Garber K, Tramontano A, Jelenic S, et al. Virus-like particle formation and translational start site choice of the plant retrotransposon Tto1. *Virology* (2008) 373(2):437–46. doi: 10.1016/j.virol.2007.11.029
15. Beliakova-Bethell N, Beckham C, Giddings TH Jr., Winey M, Parker R, Sandmeyer S. Virus-like particles of the Ty3 retrotransposon assemble in association with P-body components. *RNA* (2006) 12(1):94–101. doi: 10.1261/rna.2264806
16. Volff JN, Lehrach H, Reinhardt R, Chourrout D. Retroelement dynamics and a novel type of chordate retrovirus-like element in the miniature genome of the tunicate *Oikopleura dioica*. *Mol Biol Evol* (2004) 21(11):2022–33. doi: 10.1093/molbev/msh207
17. Grandi N, Tramontano E. Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses. *Front Immunol* (2018) 9:2039:2039. doi: 10.3389/fimmu.2018.02039
18. Moelling K, Broecker F. The reverse transcriptase-RNase H: from viruses to antiviral defense. *Ann N Y Acad Sci* (2015) 1341:126–35. doi: 10.1111/nyas.12668
19. Grandi N, Cadeddu M, Blomberg J, Mayer J, Tramontano E. HERV-W group evolutionary history in non-human primates: characterization of ERV-W orthologs in Catarrhini and related ERV groups in Platyrrhini. *BMC Evol Biol* (2018) 18(1):6. doi: 10.1186/s12862-018-1125-1
20. Mayer J, Meese E, Mueller-Lantzsch N. Human endogenous retrovirus K homologous sequences and their coding capacity in Old World primates. *J Virol* (1998) 72(3):1870–5. doi: 10.1128/JVI.72.3.1870-1875.1998
21. Macfarlane CM, Badge RM. Genome-wide amplification of proviral sequences reveals new polymorphic HERV-K(HML-2) proviruses in humans and chimpanzees that are absent from genome assemblies. *Retrovirology* (2015) 12:35. doi: 10.1186/s12977-015-0162-8
22. Penzkofer T, Jager M, Figlerowicz M, Badge R, Mundlos S, Robinson PN, et al. L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes. *Nucleic Acids Res* (2017) 45(D1):D68–73. doi: 10.1093/nar/gkw925
23. Deininger P. Alu elements: know the SINEs. *Genome Biol* (2011) 12(12):236. doi: 10.1186/gb-2011-12-12-236
24. Ostertag EM, Goodier JL, Zhang Y, Kazazian HH Jr. SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am J Hum Genet* (2003) 73(6):1444–51. doi: 10.1086/380207
25. Ullu E, Tschudi C. Alu sequences are processed 7SL RNA genes. *Nature* (1984) 312(5990):171–2. doi: 10.1038/312171a0
26. Subramanian RP, Wildschutte JH, Russo C, Coffin JM. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology* (2011) 8:90. doi: 10.1186/1742-4690-8-90
27. Le Dantec C, Vallet S, Brooks WH, Renaudineau Y. Human endogenous retrovirus group E and its involvement in diseases. *Viruses* (2015) 7(3):1238–57. doi: 10.3390/v7031238
28. Hohn O, Hanke K, Bannert N. HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. *Front Oncol* (2013) 3:246:246. doi: 10.3389/fonc.2013.00246
29. Jha AR, Nixon DF, Rosenberg MG, Martin JN, Deeks SG, Hudson RR, et al. Human endogenous retrovirus K106 (HERV-K106) was infectious after the emergence of anatomically modern humans. *PLoS One* (2011) 6(5):e20234. doi: 10.1371/journal.pone.0020234
30. Kahyo T, Yamada H, Tao H, Kurabe N, Sugimura H. Insertionally polymorphic sites of human endogenous retrovirus-K (HML-2) with long target site duplications. *BMC Genomics* (2017) 18(1):487. doi: 10.1186/s12864-017-3872-6
31. Wildschutte JH, Williams ZH, Montesion M, Subramanian RP, Kidd JM, Coffin JM. Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc Natl Acad Sci U.S.A.* (2016) 113(16):E2326–34. doi: 10.1073/pnas.1602336113
32. Belshaw R, Dawson AL, Woolven-Allen J, Redding J, Burt A, Tristem M. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. *J Virol* (2005) 79(19):12507–14. doi: 10.1128/JVI.79.19.12507-12514.2005
33. Lenz J. HERV-K HML-2 diversity among humans. *Proc Natl Acad Sci U.S.A.* (2016) 113(16):4240–2. doi: 10.1073/pnas.1603569113
34. 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. *Proc Natl Acad Sci U.S.A.* (2019) 116(4):1337–46. doi: 10.1073/pnas.1814203116
35. Dewannieux M, Harper F, Richaud A, Letzelter C, Ribet D, Pierron G, et al. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res* (2006) 16(12):1548–56. doi: 10.1101/gr.5565706
36. Beimforde N, Hanke K, Ammar I, Kurth R, Bannert N. Molecular cloning and functional characterization of the human endogenous retrovirus K113. *Virology* (2008) 371(1):216–25. doi: 10.1016/j.virol.2007.09.036
37. Boller K, Schonfeld K, Lischer S, Fischer N, Hoffmann A, Kurth R, et al. Human endogenous retrovirus K113 is capable of producing intact viral particles. *J Gen Virol* (2008) 89(Pt 2):567–72. doi: 10.1099/vir.0.83534-0
38. Michaud HA, de Mulder M, SenGupta D, Deeks SG, Martin JN, Pilcher CD, et al. Trans-activation, post-transcriptional maturation, and induction of antibodies to HERV-K (HML-2) envelope transmembrane protein in HIV-1 infection. *Retrovirology* (2014) 11:10. doi: 10.1186/1742-4690-11-10
39. Montesion M, Bhardwaj N, Williams ZH, Kuperwasser C, Coffin JM. Mechanisms of HERV-K (HML-2) Transcription during Human Mammary Epithelial Cell Transformation. *J Virol* (2018) 92(1):e01258-17. doi: 10.1128/JVI.01258-17
40. Goering W, Ribarska T, Schulz WA. Selective changes of retroelement expression in human prostate cancer. *Carcinogenesis* (2011) 32(10):1484–92. doi: 10.1093/carcin/bgr181
41. Contreras-Galindo R, Kaplan MH, Markovitz DM, Lorenzo E, Yamamura Y. Detection of HERV-K(HML-2) viral RNA in plasma of HIV type 1-infected individuals. *AIDS Res Hum Retroviruses* (2006) 22(10):979–84. doi: 10.1089/aid.2006.22.979
42. Contreras-Galindo R, Gonzalez M, Almodovar-Camacho S, Gonzalez-Ramirez S, Lorenzo E, Yamamura Y. A new Real-Time-RT-PCR for quantitation of human endogenous retroviruses type K (HERV-K) RNA load in plasma samples: increased HERV-K RNA titers in HIV-1 patients with HAART non-suppressive regimens. *J Virol Methods* (2006) 136(1-2):51–7. doi: 10.1016/j.jviromet.2006.03.029
43. Gonzalez-Hernandez MJ, Swanson MD, Contreras-Galindo R, Cookinham S, King SR, Noel RJ Jr., et al. Expression of human endogenous retrovirus type K (HML-2) is activated by the Tat protein of HIV-1. *J Virol* (2012) 86(15):7790–805. doi: 10.1128/JVI.07215-11
44. Bhardwaj N, Maldarelli F, Mellors J, Coffin JM. HIV-1 infection leads to increased transcription of human endogenous retrovirus HERV-K (HML-2) proviruses in vivo but not to increased virion production. *J Virol* (2014) 88(19):11108–20. doi: 10.1128/JVI.01623-14

45. de Mulder M, SenGupta D, Deeks SG, Martin JN, Pilcher CD, Hecht FM, et al. Anti-HERV-K (HML-2) capsid antibody responses in HIV elite controllers. *Retrovirology* (2017) 14(1):41. doi: 10.1186/s12977-017-0365-2
46. Reynier F, Verjat T, Turrel F, Imbert PE, Marotte H, Mouglin B, et al. Increase in human endogenous retrovirus HERV-K (HML-2) viral load in active rheumatoid arthritis. *Scand J Immunol* (2009) 70(3):295–9. doi: 10.1111/j.1365-3083.2009.02271.x
47. Ehlhardt S, Seifert M, Schneider J, Ojak A, Zang KD, Mehraein Y. Human endogenous retrovirus HERV-K (HML-2) Rec expression and transcriptional activities in normal and rheumatoid arthritis synovia. *J Rheumatol* (2006) 33(1):16–23.
48. Ono M, Kawakami M, Ushikubo H. Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *J Virol* (1987) 61(6):2059–62. doi: 10.1128/JVI.61.6.2059-2062.1987
49. Gabriel U, Steidler A, Trojan L, Michel MS, Seifarth W, Fabarius A. Smoking increases transcription of human endogenous retroviruses in a newly established in vitro cell model and in normal urothelium. *AIDS Res Hum Retroviruses* (2010) 26(8):883–8. doi: 10.1089/aid.2010.0014
50. Bergallo M, Galliano I, Dapra V, Pirra A, Montanari P, Pavan M, et al. Transcriptional Activity of Human Endogenous Retroviruses in Response to Prenatal Exposure of Maternal Cigarette Smoking. *Am J Perinatol* (2019) 36(10):1060–5. doi: 10.1055/s-0038-1675768
51. Hedström AK, Klareskog L, Alfredsson L. Exposure to passive smoking and rheumatoid arthritis risk: results from the Swedish EIRA study. *Ann Rheum Dis* (2018) 77(7):970–2. doi: 10.1136/annrheumdis-2018-212973
52. Vittecoq O, Richard L, Banse C, Lequerre T. The impact of smoking on rheumatoid arthritis outcomes. *Joint Bone Spine* (2018) 85(2):135–8. doi: 10.1016/j.jbspin.2017.12.004
53. Talotta R, Atzeni F, Laska MJ. The contribution of HERV-E clone 4-1 and other HERV-E members to the pathogenesis of rheumatic autoimmune diseases. *APMIS* (2020) 128(5):367–77. doi: 10.1111/apm.13039
54. Ogasawara H, Hishikawa T, Sekigawa I, Hashimoto H, Yamamoto N, Maruyama N. Sequence analysis of human endogenous retrovirus clone 4-1 in systemic lupus erythematosus. *Autoimmunity* (2000) 33(1):15–21. doi: 10.3109/08916930108994105
55. Ogasawara H, Kaneko H, Hishikawa T, Sekigawa I, Takasaki Y, Hashimoto H, et al. Molecular mimicry between human endogenous retrovirus clone 4-1 and HLA class I antigen with reference to the pathogenesis of systemic lupus erythematosus. *Rheumatol (Oxford)* (1999) 38(11):1163–4. doi: 10.1093/rheumatology/38.11.1163
56. Bessis D, Moles JP, Basset-Seguin N, Tesniere A, Arpin C, Guillou JJ. Differential expression of a human endogenous retrovirus E transmembrane envelope glycoprotein in normal, psoriatic and atopic dermatitis human skin. *Br J Dermatol* (2004) 151(4):737–45. doi: 10.1111/j.1365-2133.2004.06116.x
57. Wu Z, Mei X, Zhao D, Sun Y, Song J, Pan W, et al. DNA methylation modulates HERV-E expression in CD4+ T cells from systemic lupus erythematosus patients. *J Dermatol Sci* (2015) 77(2):110–6. doi: 10.1016/j.jdermsci.2014.12.004
58. Wang X, Zhao C, Zhang C, Mei X, Song J, Sun Y, et al. Increased HERV-E clone 4-1 expression contributes to DNA hypomethylation and IL-17 release from CD4(+) T cells via miR-302d/MBD2 in systemic lupus erythematosus. *Cell Commun Signal* (2019) 17(1):94. doi: 10.1186/s12964-019-0416-5
59. Tokuyama M, Kong Y, Song E, Jayewickreme T, Kang I, Iwasaki A. ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci U.S.A.* (2018) 115(50):12565–72. doi: 10.1073/pnas.1814589115
60. Iniguez LP, de Mulder Rougvie M, Stearrett N, Jones RB, Ormsby CE, Reyes-Teran G, et al. Transcriptomic analysis of human endogenous retroviruses in systemic lupus erythematosus. *Proc Natl Acad Sci U.S.A.* (2019) 116(43):21350–1. doi: 10.1073/pnas.1907705116
61. Volkman HE, Stetson DB. The enemy within: endogenous retroelements and autoimmune disease. *Nat Immunol* (2014) 15(5):415–22. doi: 10.1038/ni.2872
62. Weiss RA, Stoye JP. Virology. Our viral inheritance. *Science* (2013) 340(6134):820–1. doi: 10.1126/science.1235148
63. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* (2000) 24(4):363–7. doi: 10.1038/74184
64. Ostertag EM, Kazazian HHJr. Biology of mammalian L1 retrotransposons. *Annu Rev Genet* (2001) 35:501–38. doi: 10.1146/annurev.genet.35.102401.091032
65. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, et al. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U.S.A.* (2003) 100(9):5280–5. doi: 10.1073/pnas.0831042100
66. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev* (2009) 23(11):1303–12. doi: 10.1101/gad.1803909
67. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* (2003) 35(1):41–8. doi: 10.1038/ng1223
68. Kanda T, Hattori K, Saigo K. Variation in number of long, open reading frames among the members of copia, a Drosophila retrotransposon, and cDNAs of copia-related RNA in virus-like particles. *Nucleic Acids Res* (1986) 14(10):4369. doi: 10.1093/nar/14.10.4369
69. Hansen LJ, Chalker DL, Sandmeyer SB. Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. *Mol Cell Biol* (1988) 8(12):5245–56. doi: 10.1128/MCB.8.12.5245
70. Ozers MS, Friesen PD. The Env-like open reading frame of the baculovirus-integrated retrotransposon TED encodes a retrovirus-like envelope protein. *Virology* (1996) 226(2):252–9. doi: 10.1006/viro.1996.0653
71. Inouye S, Sunshine MG, Six EW, Inouye M. Retronophage phi R73: an E. coli phage that contains a retroelement and integrates into a tRNA gene. *Science* (1991) 252(5008):969–71. doi: 10.1126/science.1709758
72. Sabot F, Schulman AH. Parasitism and the retrotransposon life cycle in plants: a hitchhiker's guide to the genome. *Hered (Edinb)* (2006) 97(6):381–8. doi: 10.1038/sj.hdy.6800903
73. Galindo-Gonzalez L, Mhiri C, Deyholos MK, Grandbastien MA. LTR-retrotransposons in plants: Engines of evolution. *Gene* (2017) 626:14–25. doi: 10.1016/j.gene.2017.04.051
74. Mavragani CP, Sagalovskiy I, Guo Q, Nezos A, Kapsogeorgou EK, Lu P, et al. Expression of Long Interspersed Nuclear Element 1 Retroelements and Induction of Type I Interferon in Patients With Systemic Autoimmune Disease. *Arthritis Rheumatol* (2016) 68(11):2686–96. doi: 10.1002/art.39795
75. Mavragani CP, Nezos A, Sagalovskiy I, Seshan S, Kirou KA, Crow MK. Defective regulation of L1 endogenous retroelements in primary Sjogren's syndrome and systemic lupus erythematosus: Role of methylating enzymes. *J Autoimmun* (2018) 88:75–82. doi: 10.1016/j.jaut.2017.10.004
76. Huang X, Su G, Wang Z, Shanguan S, Cui X, Zhu J, et al. Hypomethylation of long interspersed nucleotide element-1 in peripheral mononuclear cells of juvenile systemic lupus erythematosus patients in China. *Int J Rheum Dis* (2014) 17(3):280–90. doi: 10.1111/1756-185X.12239
77. Lim YW, Sanz LA, Xu X, Hartono SR, Chedin F. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutieres syndrome. *Elife* (2015) 4:e08007. doi: 10.7554/eLife.08007
78. Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JL, Rodriguez-Ubrea J, et al. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* (2010) 20(2):170–9. doi: 10.1101/gr.100289.109
79. Balada E, Ordi-Ros J, Serrano-Acedo S, Martinez-Lostao L, Rosa-Leyva M, Vilardell-Tarres M. Transcript levels of DNA methyltransferases DNMT1, DNMT3A and DNMT3B in CD4+ T cells from patients with systemic lupus erythematosus. *Immunology* (2008) 124(3):339–47. doi: 10.1111/j.1365-2567.2007.02771.x
80. Nawrocki MJ, Majewski D, Puszczewicz M, Jagodzinski PP. Decreased mRNA expression levels of DNA methyltransferases type 1 and 3A in systemic lupus erythematosus. *Rheumatol Int* (2017) 37(5):775–83. doi: 10.1007/s00296-017-3711-8
81. Momparler RL. Molecular, cellular and animal pharmacology of 5-aza-2'-deoxycytidine. *Pharmacol Ther* (1985) 30(3):287–99. doi: 10.1016/0169-4727(85)90001-0
82. Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene* (2008) 27(3):404–8. doi: 10.1038/sj.onc.1210631
83. Quddus J, Johnson KJ, Gavalchin J, Amento EP, Chrisp CE, Yung RL, et al. Treating activated CD4+ T cells with either of two distinct DNA methyltransferase inhibitors, 5-azacytidine or procainamide, is sufficient to cause a lupus-like disease in syngeneic mice. *J Clin Invest* (1993) 92(1):38–53. doi: 10.1172/JCI116576

84. Cornacchia E, Golbus J, Maybaum J, Strahler J, Hanash S, Richardson B. Hyalazine and procainamide inhibit T cell DNA methylation and induce autoreactivity. *J Immunol* (1988) 140(7):2197–200.
85. Lieberman MW, Beach LR, Palminter RD. Ultraviolet radiation-induced metallothionein-I gene activation is associated with extensive DNA demethylation. *Cell* (1983) 35(1):207–14. doi: 10.1016/0092-8674(83)90223-4
86. Li TH, Schmid CW. Differential stress induction of individual Alu loci: implications for transcription and retrotransposition. *Gene* (2001) 276(1–2):135–41. doi: 10.1016/S0378-1119(01)00637-0
87. Goodier JL, Zhang L, Vetter MR, Kazazian HHJr. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Mol Cell Biol* (2007) 27(18):6469–83. doi: 10.1128/MCB.00332-07
88. Goodier JL, Cheung LE, Kazazian HHJr. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res* (2013) 41(15):7401–19. doi: 10.1093/nar/gkt512
89. Taylor MS, Altukhov I, Molloy KR, Mita P, Jiang H, Adney EM, et al. Dissection of affinity captured LINE-1 macromolecular complexes. *Elife* (2018) 7:e30094. doi: 10.7554/eLife.30094
90. Taylor MS, LaCava J, Dai L, Mita P, Burns KH, Rout MP, et al. Characterization of L1-Ribonucleoprotein Particles. *Methods Mol Biol* (2016) 1400:311–38. doi: 10.1007/978-1-4939-3372-3_20
91. Taylor MS, LaCava J, Mita P, Molloy KR, Huang CR, Li D, et al. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell* (2013) 155(5):1034–48. doi: 10.1016/j.cell.2013.10.021
92. Maraiá RJ, Intine RV. Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function. *Mol Cell Biol* (2001) 21(2):367–79. doi: 10.1128/MCB.21.2.367-379.2001
93. Duggal NK, Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* (2012) 12(10):687–95. doi: 10.1038/nri3295
94. Urbano VD, De Crignis E, Re MC. Host Restriction Factors and Human Immunodeficiency Virus (HIV-1): A Dynamic Interplay Involving All Phases of the Viral Life Cycle. *Curr HIV Res* (2018) 16(3):184–207. doi: 10.2174/1570162X16666180817115830
95. Goodier JL, Cheung LE, Kazazian HHJr. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet* (2012) 8(10):e1002941. doi: 10.1371/journal.pgen.1002941
96. Merindol N, Berthoux L. Restriction Factors in HIV-1 Disease Progression. *Curr HIV Res* (2015) 13(6):448–61. doi: 10.2174/1570162X13666150608104412
97. Pyndiah N, Telenti A, Rausell A. Evolutionary genomics and HIV restriction factors. *Curr Opin HIV AIDS* (2015) 10(2):79–83. doi: 10.1097/COH.0000000000000138
98. Ghimire D, Rai M, Gaur R. Novel host restriction factors implicated in HIV-1 replication. *J Gen Virol* (2018) 99(4):435–46. doi: 10.1099/jgv.0.001026
99. Stetson DB, Ko JS, Heidmann T, Medzhitov R. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* (2008) 134(4):587–98. doi: 10.1016/j.cell.2008.06.032
100. Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, et al. A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* (2011) 472(7344):481–5. doi: 10.1038/nature09907
101. Gao D, Li T, Li XD, Chen X, Li QZ, Wight-Carter M, et al. Activation of cyclic GMP-AMP synthase by self-DNA causes autoimmune diseases. *Proc Natl Acad Sci U.S.A.* (2015) 112(42):E5699–705. doi: 10.1073/pnas.1516465112
102. Yoneyama M, Kikuchi M, Matsumoto K, Imaizumi T, Miyagishi M, Taira K, et al. Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J Immunol* (2005) 175(5):2851–8. doi: 10.4049/jimmunol.175.5.2851
103. Loo YM, Gale MJr. Immune signaling by RIG-I-like receptors. *Immunity* (2011) 34(5):680–92. doi: 10.1016/j.immuni.2011.05.003
104. Keating SE, Baran M, Bowie AG. Cytosolic DNA sensors regulating type I interferon induction. *Trends Immunol* (2011) 32(12):574–81. doi: 10.1016/j.it.2011.08.004
105. Radetskyy R, Daher A, Gatignol A. ADAR1 and PKR, interferon stimulated genes with clashing effects on HIV-1 replication. *Cytokine Growth Factor Rev* (2018) 40:48–58. doi: 10.1016/j.cytogfr.2018.03.007
106. Wang TT, Li ZG, Wang Q. The RNA-Specific Adenosine Deaminase ADAR1 Inhibits Human Protein Kinase R Activation. *Viral Immunol* (2018) 31(7):537–8. doi: 10.1089/vim.2018.0056
107. Nguyen TA, Smith BRC, Tate MD, Belz GT, Barrios MH, Elgass KD, et al. SIDT2 Transports Extracellular dsRNA into the Cytoplasm for Innate Immune Recognition. *Immunity* (2017) 47(3):498–509 e6. doi: 10.1016/j.immuni.2017.08.007
108. Vance RE. Cytosolic DNA Sensing: The Field Narrows. *Immunity* (2016) 45(2):227–8. doi: 10.1016/j.immuni.2016.08.006
109. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* (2003) 197(6):711–23. doi: 10.1084/jem.20021553
110. Ronnblom L, Eloranta ML. The interferon signature in autoimmune diseases. *Curr Opin Rheumatol* (2013) 25(2):248–53. doi: 10.1097/BOR.0b013e32835c7e32
111. Eloranta ML, Ronnblom L. Cause and consequences of the activated type I interferon system in SLE. *J Mol Med (Berl)* (2016) 94(10):1103–10. doi: 10.1007/s00109-016-1421-4
112. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol* (2014) 66(6):1583–95. doi: 10.1002/art.38628
113. Bodewes ILA, Al-Ali S, van Helden-Meeuwsen CG, Maria NI, Tarn J, Lendrem DW, et al. Systemic interferon type I and type II signatures in primary Sjogren's syndrome reveal differences in biological disease activity. *Rheumatol (Oxford)* (2018) 57(5):921–30. doi: 10.1093/rheumatology/kex490
114. Eloranta ML, Franck-Larsson K, Lovgren T, Kalamajski S, Ronnblom A, Rubin K, et al. Type I interferon system activation and association with disease manifestations in systemic sclerosis. *Ann Rheum Dis* (2010) 69(7):1396–402. doi: 10.1136/ard.2009.121400
115. Wuttge DM, Lood C, Tufvesson E, Scheja A, Truedsson L, Bengtsson AA, et al. Increased serum type I interferon activity in early systemic sclerosis patients is associated with antibodies against Sjogren's syndrome antigens and nuclear ribonucleoprotein antigens. *Scand J Rheumatol* (2013) 42(3):235–40. doi: 10.3109/03009742.2012.736532
116. Niewold TB, Kariuki SN, Morgan GA, Shrestha S, Pachman LM. Elevated serum interferon-alpha activity in juvenile dermatomyositis: associations with disease activity at diagnosis and after thirty-six months of therapy. *Arthritis Rheum* (2009) 60(6):1815–24. doi: 10.1002/art.24555
117. Rodriguez-Carrio J, Alperi-Lopez M, Lopez P, Ballina-Garcia FJ, Suarez A. Heterogeneity of the Type I Interferon Signature in Rheumatoid Arthritis: A Potential Limitation for Its Use As a Clinical Biomarker. *Front Immunol* (2017) 8:2007:2007. doi: 10.3389/fimmu.2017.02007
118. El-Sherbiny YM, Psarras A, Md Yusof MY, Hensor EMA, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. *Sci Rep* (2018) 8(1):5793. doi: 10.1038/s41598-018-24198-1
119. Zahn S, Rehkemper C, Kummerer BM, Ferring-Schmidt S, Bieber T, Tuting T, et al. Evidence for a pathophysiological role of keratinocyte-derived type III interferon (IFNλ) in cutaneous lupus erythematosus. *J Invest Dermatol* (2011) 131(1):133–40. doi: 10.1038/jid.2010.244
120. Blomberg S, Eloranta ML, Cederblad B, Nordlin K, Alm GV, Ronnblom L. Presence of cutaneous interferon-alpha producing cells in patients with systemic lupus erythematosus. *Lupus* (2001) 10(7):484–90. doi: 10.1191/096120301678416042
121. Ronnblom LE, Alm GV, Oberg KE. Possible induction of systemic lupus erythematosus by interferon-alpha treatment in a patient with a malignant carcinoid tumour. *J Intern Med* (1990) 227(3):207–10. doi: 10.1097/BOR.0b013e32835c7e32
122. Buechler MB, Teal TH, Elkon KB, Hamerman JA. Cutting edge: Type I IFN drives emergency myelopoiesis and peripheral myeloid expansion during chronic TLR7 signaling. *J Immunol* (2013) 190(3):886–91. doi: 10.4049/jimmunol.1202739
123. Lee PY, Li Y, Kumagai Y, Xu Y, Weinstein JS, Kellner ES, et al. Type I interferon modulates monocyte recruitment and maturation in chronic inflammation. *Am J Pathol* (2009) 175(5):2023–33. doi: 10.2353/ajpath.2009.090328 doi: S0002-9440(10)60712-2 [pii].

124. Blanco P, Palucka AK, Gill M, Pascual V, Banchereau J. Induction of dendritic cell differentiation by IFN- α in systemic lupus erythematosus. *Science* (2001) 294(5546):1540–3. doi: 10.1126/science.1064890
125. Kolumam GA, Thomas S, Thompson LJ, Sprent J, Murali-Krishna K. Type I interferons act directly on CD8 T cells to allow clonal expansion and memory formation in response to viral infection. *J Exp Med* (2005) 202(5):637–50. doi: 10.1084/jem.20050821
126. Joo H, Coquery C, Xue Y, Gayet I, Dillon SR, Punaro M, et al. Serum from patients with SLE instructs monocytes to promote IgG and IgA plasmablast differentiation. *J Exp Med* (2012) 209(7):1335–48. doi: 10.1084/jem.20111644
127. Aicardi J, Goutieres F. A progressive familial encephalopathy in infancy with calcifications of the basal ganglia and chronic cerebrospinal fluid lymphocytosis. *Ann Neurol* (1984) 15(1):49–54. doi: 10.1002/ana.410150109
128. Crow YJ, Manel N. Aicardi-Goutieres syndrome and the type I interferonopathies. *Nat Rev Immunol* (2015) 15(7):429–40. doi: 10.1038/nri3850
129. Crow YJ, Chase DS, Lowenstein Schmidt J, Szykiewicz M, Forte GM, Gornall HL, et al. Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1. *Am J Med Genet A* (2015) 167A(2):296–312. doi: 10.1002/ajmg.a.36887
130. Crow YJ, Vanderver A, Orcesi S, Kuijpers TW, Rice GI. Therapies in Aicardi-Goutieres syndrome. *Clin Exp Immunol* (2014) 175(1):1–8. doi: 10.1111/cei.12115
131. Ramantani G, Kohlhase J, Hertzberg C, Innes AM, Engel K, Hunger S, et al. Expanding the phenotypic spectrum of lupus erythematosus in Aicardi-Goutieres syndrome. *Arthritis Rheum* (2010) 62(5):1469–77. doi: 10.1002/art.27367
132. Pokatayev V, Hasin N, Chon H, Cerritelli SM, Sakhuja K, Ward JM, et al. RNase H2 catalytic core Aicardi-Goutieres syndrome-related mutant invokes cGAS-STING innate immune-sensing pathway in mice. *J Exp Med* (2016) 213(3):329–36. doi: 10.1084/jem.20151464
133. Ablasser A, Hemmerling I, Schmid-Burgk JL, Behrendt R, Roers A, Hornung V. TREX1 Deficiency Triggers Cell-Autonomous Immunity in a cGAS-Dependent Manner. *J Immunol* (2014) 192(12):5993–7. doi: 10.4049/jimmunol.1400737
134. Yang YG, Lindahl T, Barnes DE. Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease. *Cell* (2007) 131(5):873–86. doi: 10.1016/j.cell.2007.10.017
135. Li P, Du J, Goodier JL, Hou J, Kang J, Kazazian HH Jr., et al. Aicardi-Goutieres syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion. *Nucleic Acids Res* (2017) 45(8):4619–31. doi: 10.1093/nar/gkx178
136. Thomas CA, Tejwani L, Trujillo CA, Negraes PD, Herai RH, Mesci P, et al. Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. *Cell Stem Cell* (2017) 21(3):319–31 e8. doi: 10.1016/j.stem.2017.07.009
137. Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, et al. Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet* (2009) 41(7):829–32. doi: 10.1038/ng.373
138. Zhao K, Du J, Han X, Goodier JL, Li P, Zhou X, et al. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1. *Cell Rep* (2013) 4(6):1108–15. doi: 10.1016/j.celrep.2013.08.019
139. Mathias SL, Scott AF, Kazazian HH Jr., Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science* (1991) 254(5039):1808–10. doi: 10.1126/science.1722352
140. Clements AP, Singer MF. The human LINE-1 reverse transcriptase: effect of deletions outside the common reverse transcriptase domain. *Nucleic Acids Res* (1998) 26(15):3528–35. doi: 10.1093/nar/26.15.3528
141. Stetson DB, Medzhitov R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* (2006) 24(1):93–103. doi: 10.1016/j.immuni.2005.12.003
142. De Cecco M, Ito T, Petraschen AP, Elias AE, Skvir NJ, Criscione SW, et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* (2019) 566(7742):73–8. doi: 10.1038/s41586-018-0784-9
143. Beck-Engeser G, Eilat D, Wabl M. An autoimmune disease prevented by anti-retroviral drugs. *Retrovirology* (2011) 8(1):91. doi: 10.1186/1742-4690-8-91
144. Achleitner M, Kleefisch M, Hennig A, Peschke K, Polikarpova A, Oertel R, et al. Lack of Trex1 Causes Systemic Autoimmunity despite the Presence of Antiretroviral Drugs. *J Immunol* (2017) 199(7):2261–9. doi: 10.4049/jimmunol.1700714
145. Rice GI, Meyzer C, Bouazza N, Hully M, Bodaert N, Semeraro M, et al. Reverse-Transcriptase Inhibitors in the Aicardi-Goutieres Syndrome. *N Engl J Med* (2018) 379(23):2275–7. doi: 10.1056/NEJMc1810983
146. Refsland EW, Harris RS. The APOBEC3 family of retroelement restriction factors. *Curr Top Microbiol Immunol* (2013) 371:1–27. doi: 10.1007/978-3-642-37765-5_1
147. Pang W, Song JH, Lu Y, Zhang XL, Zheng HY, Jiang J, et al. Host Restriction Factors APOBEC3G/3F and Other Interferon-Related Gene Expressions Affect Early HIV-1 Infection in Northern Pig-Tailed Macaque (Macaca leonina). *Front Immunol* (2018) 9:1965:1965. doi: 10.3389/fimmu.2018.01965
148. Lindic N, Budic M, Petan T, Knisbacher BA, Levanon EY, Lovsin N. Differential inhibition of LINE1 and LINE2 retrotransposition by vertebrate AID/APOBEC proteins. *Retrovirology* (2013) 10:156. doi: 10.1186/1742-4690-10-156
149. Bogerd HP, Wiegand HL, Doehle BP, Lueders KK, Cullen BR. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res* (2006) 34(1):89–95. doi: 10.1093/nar/gkj416
150. Mita P, Wudzinska A, Sun X, Andrade J, Nayak S, Kahler DJ, et al. LINE-1 protein localization and functional dynamics during the cell cycle. *Elife* (2018) 7:e30058. doi: 10.7554/eLife.30058
151. Li X, Zhang J, Jia R, Cheng V, Xu X, Qiao W, et al. The MOV10 helicase inhibits LINE-1 mobility. *J Biol Chem* (2013) 288(29):21148–60. doi: 10.1074/jbc.M113.465856
152. Arjan-Odedra S, Swanson CM, Sherer NM, Wolinsky SM, Malim MH. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology* (2012) 9:53. doi: 10.1186/1742-4690-9-53
153. Choi J, Hwang SY, Ahn K. Interplay between RNASEH2 and MOV10 controls LINE-1 retrotransposition. *Nucleic Acids Res* (2018) 46(4):1912–26. doi: 10.1093/nar/gkx1312
154. Wang X, Han Y, Dang Y, Fu W, Zhou T, Ptak RG, et al. Moloney leukemia virus 10 (MOV10) protein inhibits retrovirus replication. *J Biol Chem* (2010) 285(19):14346–55. doi: 10.1074/jbc.M110.109314
155. Goodier JL. Restricting retrotransposons: a review. *Mob DNA* (2016) 7:16. doi: 10.1186/s13100-016-0070-z
156. Toth KF, Pezic D, Stuwe E, Webster A. The piRNA Pathway Guards the Germline Genome Against Transposable Elements. *Adv Exp Med Biol* (2016) 886:51–77. doi: 10.1007/978-94-017-7417-8_4
157. Kuriakose T, Kanneganti TD. ZBP1: Innate Sensor Regulating Cell Death and Inflammation. *Trends Immunol* (2018) 39(2):123–34. doi: 10.1016/j.it.2017.11.002
158. Jiao H, Wachsmuth L, Kumari S, Schwarzer R, Lin J, Eren RO, et al. Z-nucleic-acid sensing triggers ZBP1-dependent necroptosis and inflammation. *Nature* (2020) 580(7803):391–5. doi: 10.1038/s41586-020-2129-8
159. Wang R, Li H, Wu J, Cai ZY, Li B, Ni H, et al. Gut stem cell necroptosis by genome instability triggers bowel inflammation. *Nature* (2020) 580(7803):386–90. doi: 10.1038/s41586-020-2127-x
160. Sciamanna I, Sinibaldi-Vallebona P, Serafino A, Spadafora C. LINE-1-encoded reverse Transcriptase as a target in cancer therapy. *Front Biosci (Landmark Ed)* (2018) 23:1360–9. doi: 10.2741/4648
161. Greenig M. HERVs, immunity, and autoimmunity: understanding the connection. *PeerJ* (2019) 7:e6711. doi: 10.7717/peerj.6711
162. Tovo PA, Rabbone I, Tinti D, Galliano I, Trada M, Dapra V, et al. Enhanced expression of human endogenous retroviruses in new-onset type 1 diabetes: Potential pathogenetic and therapeutic implications. *Autoimmunity* (2020) 53(5):283–8. doi: 10.1080/08916934.2020.1777281
163. Talal N, Dauphinee MJ, Dang H, Alexander SS, Hart DJ, Garry RF. Detection of serum antibodies to retroviral proteins in patients with primary Sjogren's syndrome (autoimmune exocrinopathy). *Arthritis Rheum* (1990) 33(6):774–81. doi: 10.1002/art.1780330603
164. Fraziano M, Montesano C, Lombardi VR, Sammarco I, De Pisa F, Mattei M, et al. Epitope specificity of anti-HIV antibodies in human and murine

- autoimmune diseases. *AIDS Res Hum Retroviruses* (1996) 12(6):491–6. doi: 10.1089/aid.1996.12.491
165. Dang H, Dauphinee MJ, Talal N, Garry RF, Seibold JR, Medsger TAJr, et al. Serum antibody to retroviral gag proteins in systemic sclerosis. *Arthritis Rheum* (1991) 34(10):1336–7. doi: 10.1002/art.1780341022
 166. Talal N, Garry RF, Schur PH, Alexander S, Dauphinee MJ, Livas IH, et al. A conserved idiotype and antibodies to retroviral proteins in systemic lupus erythematosus. *J Clin Invest* (1990) 85(6):1866–71. doi: 10.1172/JCI114647
 167. Nelson PN, Lever AM, Smith S, Pitman R, Murray P, Perera SA, et al. Molecular investigations implicate human endogenous retroviruses as mediators of anti-retroviral antibodies in autoimmune rheumatic disease. *Immunol Invest* (1999) 28(4):277–89. doi: 10.3109/08820139909060862
 168. Garcia-Montojo M, Doucet-O'Hare T, Henderson L, Nath A. Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit Rev Microbiol* (2018) 44(6):715–38. doi: 10.1080/1040841X.2018.1501345
 169. Freimanis G, Hooley P, Ejtehadi HD, Ali HA, Veitch A, Rylance PB, et al. A role for human endogenous retrovirus-K (HML-2) in rheumatoid arthritis: investigating mechanisms of pathogenesis. *Clin Exp Immunol* (2010) 160(3):340–7. doi: 10.1111/j.1365-2249.2010.04110.x
 170. Marnett G, Erre GL, Caggia E, Mura S, Cossu D, Bo M, et al. Identification of a HERV-K env surface peptide highly recognized in Rheumatoid Arthritis (RA) patients: a cross-sectional case-control study. *Clin Exp Immunol* (2017) 189(1):127–31. doi: 10.1111/cei.12964
 171. Nelson PN, Roden D, Nevill A, Freimanis GL, Trella M, Ejtehadi HD, et al. Rheumatoid arthritis is associated with IgG antibodies to human endogenous retrovirus gag matrix: a potential pathogenic mechanism of disease? *J Rheumatol* (2014) 41(10):1952–60. doi: 10.3899/jrheum.130502
 172. Bengtsson A, Blomberg J, Nived O, Pipkorn R, Toth L, Sturfelt G. Selective antibody reactivity with peptides from human endogenous retroviruses and nonviral poly(amino acids) in patients with systemic lupus erythematosus. *Arthritis Rheum* (1996) 39(10):1654–63. doi: 10.1002/art.1780391007
 173. Wang-Johanning F, Radvanyi L, Rycaj K, Plummer JB, Yan P, Sastry KJ, et al. Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. *Cancer Res* (2008) 68(14):5869–77. doi: 10.1158/0008-5472.CAN-07-6838
 174. Wang-Johanning F, Rycaj K, Plummer JB, Li M, Yin B, Frerich K, et al. Immunotherapeutic potential of anti-human endogenous retrovirus-K envelope protein antibodies in targeting breast tumors. *J Natl Cancer Inst* (2012) 104(3):189–210. doi: 10.1093/jnci/djr540
 175. Zhou F, Li M, Wei Y, Lin K, Lu Y, Shen J, et al. Activation of HERV-K Env protein is essential for tumorigenesis and metastasis of breast cancer cells. *Oncotarget* (2016) 7(51):84093–117. doi: 10.18632/oncotarget.11455
 176. Goering W, Schmitt K, Dostert M, Schaal H, Deenen R, Mayer J, et al. Human endogenous retrovirus HERV-K(HML-2) activity in prostate cancer is dominated by a few loci. *Prostate* (2015) 75(16):1958–71. doi: 10.1002/pros.23095
 177. Johanning GL, Malouf GG, Zheng X, Esteva FJ, Weinstein JN, Wang-Johanning F, et al. Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype. *Sci Rep* (2017) 7:41960. doi: 10.1038/srep41960
 178. Hishikawa T, Ogasawara H, Kaneko H, Shirasawa T, Matsuura Y, Sekigawa I, et al. Detection of antibodies to a recombinant gag protein derived from human endogenous retrovirus clone 4-1 in autoimmune diseases. *Viral Immunol* (1997) 10(3):137–47. doi: 10.1089/vim.1997.10.137
 179. Romero V, Fert-Bober J, Nigrovic PA, Darrah E, Haque UJ, Lee DM, et al. Immune-mediated pore-forming pathways induce cellular hypercitrullination and generate citrullinated autoantigens in rheumatoid arthritis. *Sci Transl Med* (2013) 5(209):209ra150. doi: 10.1126/scitranslmed.3006869
 180. Zhou Y, Chen B, Mittereder N, Chaerkady R, Strain M, An LL, et al. Spontaneous Secretion of the Citrullination Enzyme PAD2 and Cell Surface Exposure of PAD4 by Neutrophils. *Front Immunol* (2017) 8:1200:1200. doi: 10.3389/fimmu.2017.01200
 181. Carter V, LaCava J, Taylor MS, Liang SY, Mustelin C, Ukadike KC, 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(1):89–99. doi: 10.1002/art.41054
 182. Hung T, Pratt GA, Sundararaman B, Townsend MJ, Chaivorapol C, Bhargale T, et al. The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. *Science* (2015) 350(6259):455–9. doi: 10.1126/science.aac7442
 183. Nakagawa K, Brusica V, McColl G, Harrison LC. Direct evidence for the expression of multiple endogenous retroviruses in the synovial compartment in rheumatoid arthritis. *Arthritis Rheum* (1997) 40(4):627–38. doi: 10.1002/art.1780400407
 184. Gaudin P, Ijaz S, Tuke PW, Marcel F, Paraz A, Seigneurin JM, et al. Infrequency of detection of particle-associated MSRV/HERV-W RNA in the synovial fluid of patients with rheumatoid arthritis. *Rheumatol (Oxford)* (2000) 39(9):950–4. doi: 10.1093/rheumatology/39.9.950
 185. Bhashratty R, Regn D, Rahman MJ, Marquardt K, Fink E, Hu WY, et al. Type 1 diabetes pathogenesis is modulated by spontaneous autoimmune responses to endogenous retrovirus antigens in NOD mice. *Eur J Immunol* (2017) 47(3):575–84. doi: 10.1002/eji.201646755
 186. Robinson-McCarthy LR, McCarthy KR, Raaben M, Piccinotti S, Nieuwenhuis J, Stubbs SH, et al. Reconstruction of the cell entry pathway of an extinct virus. *PLoS Pathog* (2018) 14(8):e1007123. doi: 10.1371/journal.ppat.1007123
 187. Conrad B, Weissmahr RN, Boni J, Arcari R, Schupbach J, Mach B. A human endogenous retroviral superantigen as candidate autoimmune gene in type 1 diabetes. *Cell* (1997) 90(2):303–13. doi: 10.1016/s0092-8674(00)80338-4
 188. Lapatschek M, Durr S, Lower R, Magin C, Wagner H, Miethke T. Functional analysis of the env open reading frame in human endogenous retrovirus IDDMK(1,2)22 encoding superantigen activity. *J Virol* (2000) 74(14):6386–93. doi: 10.1128/JVI.74.14.6386-6393.2000
 189. Sicat J, Sutkowski N, Huber BT. Expression of human endogenous retrovirus HERV-K18 superantigen is elevated in juvenile rheumatoid arthritis. *J Rheumatol* (2005) 32(9):1821–31.
 190. Oakes B, Hoagland-Henefield M, Komaroff AL, Erickson JL, Huber BT. Human endogenous retrovirus-K18 superantigen expression and human herpesvirus-6 and human herpesvirus-7 viral loads in chronic fatigue patients. *Clin Infect Dis* (2013) 56(10):1394–400. doi: 10.1093/cid/cit086
 191. Kalogirou EM, Piperi EP, Tosios KI, Tsiambas E, Fanourakis G, Sklavounou A. Ductal cells of minor salivary glands in Sjogren's syndrome express LINE-1 ORF2p and APOBEC3B. *J Oral Pathol Med* (2018) 47(2):179–85. doi: 10.1111/jop.12656
 192. Kell AM, Gale MJr. RIG-I in RNA virus recognition. *Virology* (2015) 479–480:110–21. doi: 10.1016/j.virol.2015.02.017
 193. Robinson T, Kariuki SN, Franek BS, Kumabe M, Kumar AA, Badaracco M, et al. Autoimmune disease risk variant of IFIH1 is associated with increased sensitivity to IFN-alpha and serologic autoimmunity in lupus patients. *J Immunol* (2011) 187(3):1298–303. doi: 10.4049/jimmunol.1100857
 194. Gorman JA, Hundhausen C, Errett JS, Stone AE, Allenspach EJ, Ge Y, et al. The A946T variant of the RNA sensor IFIH1 mediates an interferon program that limits viral infection but increases the risk for autoimmunity. *Nat Immunol* (2017) 18(7):744–52. doi: 10.1038/ni.3766
 195. Chung H, Calis JJA, Wu X, Sun T, Yu Y, Sarbanes SL, et al. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* (2018) 172(4):811–24 e14. doi: 10.1016/j.cell.2017.12.038
 196. Rice GI, Kasher PR, Forte GM, Mannion NM, Greenwood SM, Szykiewicz M, et al. Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. *Nat Genet* (2012) 44(11):1243–8. doi: 10.1038/ng.2414
 197. Shao WH, Shu DH, Zhen Y, Hilliard B, Priest SO, Cesaroni M, et al. Prion-like Aggregation of Mitochondrial Antiviral Signaling Protein in Lupus Patients Is Associated With Increased Levels of Type I Interferon. *Arthritis Rheumatol* (2016) 68(11):2697–707. doi: 10.1002/art.39733
 198. An J, Durcan L, Karr RM, Briggs TA, Rice GI, Teal TH, et al. Expression of Cyclic GMP-AMP Synthase in Patients With Systemic Lupus Erythematosus. *Arthritis Rheum* (2017) 69(4):800–7. doi: 10.1002/art.40002
 199. Chan SR, Blackburn EH. Telomeres and telomerase. *Philos Trans R Soc Lond B Biol Sci* (2004) 359(1441):109–21. doi: 10.1098/rstb.2003.1370
 200. Autexier C, Lue NF. The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem* (2006) 75:493–517. doi: 10.1146/annurev.biochem.75.103004.142412

201. Piskareva O, Denmukhametova S, Schmatchenko V. Functional reverse transcriptase encoded by the human LINE-1 from baculovirus-infected insect cells. *Protein Expr Purif* (2003) 28(1):125–30. doi: 10.1016/S1046-5928(02)00655-1
202. Dombroski BA, Feng Q, Mathias SL, Sassaman DM, Scott AF, Kazazian HHJr., et al. An in vivo assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* (1994) 14(7):4485–92. doi: 10.1128/MCB.14.7.4485
203. Dai L, Huang Q, Boeke JD. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. *BMC Biochem* (2011) 12:18. doi: 10.1186/1471-2091-12-18
204. Jones RB, Garrison KE, Wong JC, Duan EH, Nixon DF, Ostrowski MA. Nucleoside analogue reverse transcriptase inhibitors differentially inhibit human LINE-1 retrotransposition. *PLoS One* (2008) 3(2):e1547. doi: 10.1371/journal.pone.0001547
205. Yin Y, Zhou L, Yuan S. Enigma of Retrotransposon Biology in Mammalian Early Embryos and Embryonic Stem Cells. *Stem Cells Int* (2018) 2018:6239245. doi: 10.1155/2018/6239245
206. Russell SJ, Stalker L, LaMarre J. PIWI, piRNAs and Retrotransposons: Complex battles during reprogramming in gametes and early embryos. *Reprod Domest Anim* (2017) 52 Suppl 4:28–38. doi: 10.1111/rda.13053
207. Hancks DC, Kazazian HHJr. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* (2012) 22(3):191–203. doi: 10.1016/j.gde.2012.02.006
208. Mukherjee S, Sharma D, Upadhyaya KC. L1 Retrotransposons Are Transcriptionally Active in Hippocampus of Rat Brain. *Prague Med Rep* (2016) 117(1):42–53. doi: 10.14712/23362936.2016.4
209. Suarez NA, Macia A, Muotri AR. LINE-1 retrotransposons in healthy and diseased human brain. *Dev Neurobiol* (2018) 78(5):434–55. doi: 10.1002/dneu.22567
210. Talotta R, Atzeni F, Laska MJ. Retroviruses in the pathogenesis of systemic lupus erythematosus: Are they potential therapeutic targets? *Autoimmunity* (2020) 53(4):177–91. doi: 10.1080/08916934.2020.1755962
211. Ronnblom L. The importance of the type I interferon system in autoimmunity. *Clin Exp Rheumatol* (2016) 34(4 Suppl 98):21–4.
212. Furie R, Khamashta M, Merrill JT, Werth VP, Kalunian K, Brohawn P, et al. Anifrolumab, an Anti-Interferon-alpha Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. *Arthritis Rheumatol* (2017) 69(2):376–86. doi: 10.1002/art.39962
213. Petri M, Wallace DJ, Spindler A, Chindalore V, Kalunian K, Mysler E, et al. Sifalimumab, a human anti-interferon-alpha monoclonal antibody, in systemic lupus erythematosus: a phase I randomized, controlled, dose-escalation study. *Arthritis Rheum* (2013) 65(4):1011–21. doi: 10.1002/art.37824
214. Steinfeld SD, Demols P, Van Vooren JP, Cogan E, Appelboom T. Zidovudine in primary Sjogren's syndrome. *Rheumatol (Oxford)* (1999) 38(9):814–7. doi: 10.1093/rheumatology/38.9.814
215. Wang L, Wang FS, Gershwin ME. Human autoimmune diseases: a comprehensive update. *J Intern Med* (2015) 278(4):369–95. doi: 10.1111/joim.12395

Conflict of Interest: TM has received consulting fees from Cugene, Kiniksa, Miro Bio, and QiLu Pharmaceuticals, has an ownership share in Amdax, and has received research funding from Gilead Sciences.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Mustelin and Ukadike. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

Autoantibody	antibody reactive with an antigen encoded by the host's genome, including retroelements.
DNA transposon	genomic sequences that can move to new locations in the genome without a reverse-transcription step.
HERV	human endogenous retrovirus, the genomic sequence (in its current form) of an originally free retrovirus that infected our ancestors and inserted its reverse-transcribed genome into the germline of their host. HERVs are inherited in a Mendelian fashion. The HERVs have accumulated genetic alterations, such as point mutations, small or large deletions, insertions, etc in a time-dependent manner.
LTR	long terminal repeat, the approximately 650-bp region of the retroviral genome that exerts transcriptional control. The LTR contains numerous transcription factor binding sites and the transcriptional initiation site, but is normally silenced by DNA methylation and histone modifications and other epigenetic means. There is a single LTR in the circular RNA genome of a free retrovirus, but during reverse-transcription for genomic

(Continued)

Continued

	insertion, the LTR is reverse-transcribed twice, ending up in the 5' and 3' ends of the insert.
Retroelement	general term for all genomic sequences that can be transcribed into RNA, reverse-transcribed to DNA, and then reinserted into a new location of the genome. Endogenous retroviruses are LTR-retrotransposons, while other retrotransposons are not.
Retrotransposon	essentially a synonym of retroelement, but is often used for non-LTR retroelements.
Provirus	the genomic cDNA for a retrovirus, endogenous or exogenous.
Solo-LTR	as a result of annealing between the two identical LTRs of a retroviral provirus followed by excision repair, only one LTR sequence is left in the genome. An insertionally polymorphic HERV may be present in the human genome as a provirus, a solo-LTR, or be absent. In the latter case, the location is referred to as "pre-integration" sequence.
Reverse transcriptase	the enzyme that synthesizes a DNA strand complementary to a single-stranded RNA template. Upon completion, an RNase domain of the reverse-transcriptase degrades the RNA template, allowing the reverse transcriptase to synthesize a second complementary DNA strand.
