


The emerging field of human endogenous retroviruses: understanding their physiological role and contribution to diseases

Nicole Grandi¹, Maria Paola Pisano¹ & Enzo Tramontano*¹ 

¹Department of Life & Environmental Sciences, University of Cagliari, Cittadella Universitaria di Monserrato, SS554, 09042 Monserrato (Cagliari), Italy

*Author for correspondence: Tel.: +39 070 675 4538; Fax: +39 070 675 4536; tramon@unica.it

“in the last decades, emerging functions have been associated with ERV presence and expression, revealing intriguing evolutionary interplay between ERVs and host biology, leading to varied outcomes.”

First draft submitted: 30 May 2019; Accepted for publication: 10 June 2019; Published online: 17 July 2019

Keywords: endogenous retroviruses • ERV • evolution • genome • immunity • syncytin

It is now well established that an impressive portion of vertebrate genomes have a viral origin, harboring integrated sequences that are designated as endogenous retroviruses (ERVs). ERVs have been acquired along vertebrate evolution, through infection of germ line cells followed by stable endogenization. ERV sequences have thus been inherent within the genome, passing down generations over millions of years, while the originating exogenous retroviruses have generally become extinct, even if some examples of still ongoing endogenization have been reported [1].

For a long time, ERVs and the other repetitive elements have been considered as junk DNA having no impact on the host. However, in the last decades, emerging functions have been associated with ERV presence and expression, revealing intriguing evolutionary interplay between ERVs and host biology, leading to varied outcomes [2]. On the one side, ERVs can rapidly lose infectivity due to random mutagenesis, recombination and a variety of defensive systems developed by vertebrates to control their expression and avoid detrimental effects. On the other side, some ERV sequences can contribute new advantageous biological activities to the host physiology, hence, their preservation throughout evolution [1,2]. The most fascinating example of positive selection is represented by syncytins: envelope proteins encoded by different ERVs that have been independently exapted in different species through a process of convergent evolution, providing placental functions to eutherian mammals [3,4]. Aside from ERV protein production, which is a relatively rare event, the sole presence of ERV integrations has, too, contributed to vertebrate physiology [5]. For instance, sequence homology between ERV copies allowed these to act as a source for genome rearrangements, contributing to genome shaping [2]. In fact, retroelements are overrepresented in mammalian genes with rapid evolution, such as those involved in immune and stress responses, suggesting an active role in the diversification and expansion of these gene families [6]. In addition, ERVs are able to influence nearby cellular genes by providing regulatory activities through their long terminal repeats (LTR), often in a tissue-specific manner [5]. The ancestral amylase gene, whose expression is specifically enhanced in parotid glands by an upstream HERV-E provirus, is just one example of this [7]. Of note, growing evidence is showing that ERVs have not only acted on single gene regulation, but contributed to the evolution of complex, lineage-specific networks of gene expression that are pivotal in vertebrate physiology. Accordingly, ERV colonization has been concomitant to the development of important immune pathways, increasing the complexity and specificity of vertebrate antiviral defenses, especially the innate ones [2,8]. The latter are the first and most ancient line of resistance against infections, being based on the recognition of conserved pathogen-associated molecular patterns by cellular sensors, and rely on

a complex system that is conserved throughout the animal kingdom [2]. In this context, it has been suggested that virtually all-existing immune defense pathways against extracellular pathogens were originated through the activity of RNase H-like enzymes from retroviruses and transposable elements, for all life domains [9]. Such mechanisms range from superinfection exclusion strategies, for example, receptor block by ERV proteins protecting from related retrovirus infections – to RNA silencing and interference, to the immune-activating action of ERV-encoded RNA and proteins resembling exogenous pathogen-associated molecular patterns [2,9]. Among the best-known examples of ERV-mediated inhibition of exogenous viral infections are two mouse factors originating from defective ERVs – *Fv4* and *Fv1* – producing an Env and a Gag protein, respectively. These ERV proteins confer resistance against murine leukemia virus either by blocking viral entry, binding to the same cellular receptor (*Fv4*), or by affecting viral integration, likely interfering with its pre-integration complex (*Fv1*) [9].

As ERVs are found in all vertebrates, they are also present in humans (HERV), and have been recently classified in different groups according to retroviral taxonomy [10]. During recent decades, the discovery and study of HERVs have been boosted by the advent of genomics and transcriptomics, which has grown year by year. This has allowed some contributions of HERVs to human genomic evolution and functions to begin to be unveiled, increasing the search for their role in healthy conditions and, especially, in a number of diseased contexts [5].

With regard to the co-option for physiological activities, the human genome harbors two HERV-derived loci, namely *ERVWE1* (HERV-W group, 7q21.2) and *ERVWE2* (HERV-FRD group, 6p24.1), coding for the domesticated Env proteins: syncytin-1 and -2, respectively [11,12]. As for other mammals, both proteins are essential in placental biology: syncytin-1 has a key role in the development and homeostasis of the syncytiotrophoblast, representing the connection between mother and fetus; while syncytin-2 is thought to be involved in immune tolerance to the fetal allograft [4]. Other established contributions to our physiology are related to embryonic development, accompanied by high levels of HERV expression. In particular, the HERV-H group has been reported to drive embryonic stem cells pluripotency through the production of nuclear long noncoding RNAs that recruit specific transcriptional activators [13]. Other evidences of HERV functions adopted for early development come from the HERV-K(HML2) elements, which encode an Env-splicing variant called Rec that was shown to interact with ~1600 mRNAs in embryonic cells, influencing their ribosomal occupancy [14]. As mentioned for other vertebrates, the colonization of the primate genome by ERVs led to important evolutionary developments, providing variation among species and shaping pivotal biological pathways [2]. Indeed, the major histocompatibility complex loci in humans and other primates present a remarkable plasticity also resulting from the high density of (H)ERV insertions, which can even result in gene physical variations among species. For example, complement C4A gene holds an HERV-K(HML10) provirus integrated within its ninth intron, and this retroviral insertion is responsible for its dichotomous size in the human population [15,16]. When considering HERV-mediated regulatory networks, it has been shown that at least one-third of p53 genomic-binding sites have been scattered by species-specific HERV proviruses, which have become major components of this conserved and pleiotropic regulatory network [17]. Similarly, it has been proposed that ERV insertions have been crucial to the establishment and amplification of the interferon transcriptional network, independently dispersing interferon-inducible enhancers in several mammalian genomes [18]. In line with this, experimental deletion of these HERV sequences in human cell lines has been shown to impair the activity of neighboring interferon-stimulated genes, affecting this fundamental immune pathway [18].

In the last years, HERVs have been tentatively linked to a number of pathological conditions, even if no definitive associations have been proven yet. In general, the quest for HERV harmful roles has been primarily studied detecting an overall increased HERV expression in diseased tissues as compared with healthy controls, observed especially in autoimmunity and tumors [2,5]. Such diseases, however, are complex disorders sharing a general immune and epigenetic dysregulation, both known to strongly and nonspecifically modulate HERV expression [2,5]. Hence, HERV stable expression in physiological conditions, together with their nonspecific epigenetic dysregulation common to many disorders, clearly indicates the need for assessing the differential expression of definite HERV loci in a specific pathological context, in order to evaluate their direct involvement. In fact, these elements could contribute, with other factors, to disease progression and symptomatology, based on their immune modulatory and/or oncogenic properties, such as with the immunogenic and syncytial functions held by Env proteins, meaning HERV multifaceted elements are balancing the same biological activities serving host physiology and exerting harmful effects [4]. Alternatively, HERV altered expression could be used as diagnostic tools, even in the absence of a causal role. For these reasons, the analysis of HERV transcriptome in human disorders and the search for pathological significance cannot disregard an exhaustive knowledge of single HERV locus structure and basal expression in healthy conditions, in order to assess the specific modulation of selected members in the diseased

counterpart [4]. Such a genomic characterization is currently still ongoing, being available for a few major HERV groups only, like the HERV-W one [19,20].

In summary, viruses developed systems to escape cellular defenses, pushing host biological systems to evolve under viral pressure, occasionally taking advantage from viral activities domesticated for novel physiological functions. Hence, studying ERV distribution and expression in normal conditions can give very important insights into evolution of humans, as well all other vertebrates. In addition, while no causal associations between HERV expression and human pathogenesis has been demonstrated yet, it is likely that the loss of balance between the control of ERV expression and its contribution to host biology – such as in disorders characterized by immune and epigenetics dysregulation – can contribute to complex diseases with a poorly understood etiology, currently. Furthermore, HERV epigenetics dysregulation might be an attractive source of biomarkers. Overall, (H)ERV genomic and transcriptomic characterization are needed to individuate specific loci of viral origins significantly modulated in pathological conditions, to assess their actual influence on the disease onset and progression and their possible relevance as biomarkers and therapeutic targets.

Financial and competing interests' disclosure

This work was supported by a POR-FESR 2014-2020 grant (G87H18000020006). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

- Greenwood AD, Ishida Y, O'Brien SP, Roca AL, Eiden MV. Transmission, evolution, and endogenization: lessons learned from recent retroviral invasions. *Microbiol. Mol. Biol. Rev.* 82(1), e00044-17 (2018).
- Grandi N, Tramontano E. Human endogenous retroviruses are ancient acquired elements still shaping innate immune responses. *Front. Immunol.* 9(September), 1–16 (2018).
- Lavialle C, Cornelis G, Dupressoir A *et al.* Paleovirology of “syncytins”, retroviral env genes exapted for a role in placentation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 368, 20120507 (2013).
- Grandi N, Tramontano E. HERV envelope proteins: physiological role and pathogenic potential in cancer and autoimmunity. *Front. Microbiol.* 9, 2039 (2018).
- Grandi N, Tramontano E. Type W human endogenous retrovirus (HERV-W) integrations and their mobilization by L1 machinery: contribution to the human transcriptome and impact on the host physiopathology. *Viruses*. 9(162), (2017).
- Bannert N, Kurth R. Retroelements and the human genome: new perspectives on an old relation. *Proc. Natl Acad. Sci. USA* 101(Suppl.), 14572–14579 (2004).
- Ting CN, Rosenberg MP, Snow CM, Samuelson LC, Meisler MH. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev.* 6(8), 1457–1465 (1992).
- Villarreal LP. Viral ancestors of antiviral systems. *Viruses* 3(10), 1933–1958 (2011).
- Moelling K, Broecker F, Russo G, Sunagawa S. RNase H as gene modifier, driver of evolution and antiviral defense. *Front. Microbiol.* 8, 1745 (2017).
- Vargiu L, Rodriguez-Tomé P, Sperber GO *et al.* Classification and characterization of human endogenous retroviruses mosaic forms are common. *Retrovirology* 13(7), (2016).
- Mi S, Lee X, Li X *et al.* Syncytin is a captive retroviral envelope protein involved. *Nature* 403(February), 785–789 (2000).
- Malassiné A, Blaise S, Handschuh K *et al.* Expression of the fusogenic HERV-FRD Env glycoprotein (Syncytin 2) in human placenta is restricted to villous cytotrophoblastic cells. *Placenta* 28(2–3), 185–191 (2007).
- Lu X, Sachs F, Ramsay L *et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nat. Struct. Mol. Biol.* 21(4), 423–425 (2014).
- Grow EJ, Flynn RA, Chavez SL *et al.* Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* 522(7555), 221–225 (2015).
- Dangel AW, Mendoza AR, Baker BJ *et al.* The dichotomous size variation of human complement C4 genes is mediated by a novel family of endogenous retroviruses, which also establishes species-specific genomic patterns among Old World primates. *Immunogenetics* 40(6), 425–436 (1994).
- Grandi N, Cadeddu M, Pisano MP, Esposito F, Blomberg J, Tramontano E. Identification of a novel HERV-K(HML10): comprehensive characterization and comparative analysis in non-human primates provide insights about HML10 proviruses structure and diffusion. *Mob. DNA* 8(1), 15 (2017).

17. Wang T, Zeng J, Lowe CB *et al.* Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc. Natl Acad. Sci. USA* 104(47), 18613–18618 (2007).
18. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* (80-) 351(6277), 1083–1087 (2016).
19. Grandi N, Cadeddu M, Blomberg J, Tramontano E. Contribution of type W human endogenous retroviruses to the human genome: characterization of HERV-W proviral insertions and processed pseudogenes. *Retrovirology* 13(67), 1–25 (2016).
20. 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.* 18(6), 1–14 (2018).