

## Dispatch

### Evolution: *Endogenous Viruses Provide Shortcuts in Antiviral Immunity*

Aris Katzourakis\* and Amr Aswad

Department of Zoology, University of Oxford, OX1 3PS, Oxford, UK

\*Correspondence: [aris.katzourakis@zoo.ox.ac.uk](mailto:aris.katzourakis@zoo.ox.ac.uk)

**Endogenous viruses are occasionally co-opted by their hosts to combat other viruses. The discovery of the widespread recruitment of endogenous viruses as regulatory elements for immune genes, points to a systematic evolutionary process in their co-option for host immunity.**

Viruses are incredibly diverse, but share a replication strategy that shifts the burden of replication onto the hosts that they infect. One consequence of this extreme association with cells is that viruses sometimes integrate into their hosts' genomes, with the potential to become fixed in the population and persist for millions of years. Most of these endogenous viral elements (EVEs) appear to be non-coding 'viral fossils', but an increasing number of examples in the literature suggest an active role for EVEs in host physiology [1,2]. In a recent report in *Science*, Edward Chuong and colleagues add to the list of virally-derived contributions to host physiology, showing that viruses have been co-opted to serve as regulators of innate immunity [3].

An early example of the unusual phenomenon of a virus being put to work by a host was the discovery that an endogenous retrovirus (ERV) envelope (*env*) gene may have been co-opted for placentation [4,5]. Furthermore, the capture of these *env*-derived 'syncytin' genes has occurred independently in many mammalian lineages, indicating that a general evolutionary process is at play, rather than a biological accident. One intriguing explanation is that the gene offers the foetus immunological protection from the mother [6], explaining the need to recruit new 'versions' of syncytin after adaptation to the effect. Indeed, this is not the only example of viruses being exapted for an immune function; there are several EVE-derived immunity genes (EDIs) that interfere with distinct stages of viral replication [2].

The abundance of immune roles for co-opted EVEs suggests that their viral origin is a driving force behind this specialization. In the evolutionary arms race between viruses and hosts, each side is evolving to adapt to innovations of the other. By recruiting a viral gene, the host is introducing a new 'weapon' into the conflict that the virus has (ironically) never encountered. The elegance of this solution is that the virus simultaneously provides the selective pressure necessary to adapt, as well as the tool to do so, and potentially the means to its own extinction. Since viruses have a higher evolutionary rate than their hosts, EDIs may offer a crucial edge in this high-paced molecular race [2]. Hosts are effectively exploiting the product

of the unique evolutionary dynamics of the arms race itself, as viruses already encode the necessary functions.

Most EDIs that have been described are repurposed viral genes that once interfaced with the host. However, in the new study, Chuong *et al.* [3] revealed evidence for a widespread role of ERV-derived regulatory elements in response to interferon- $\gamma$  (IFN- $\gamma$ ), a proinflammatory cytokine that promotes expression of immune genes during infection [3]. The authors identified 20 ERV families that are bound by IFN- $\gamma$ -stimulated transcription factors in human cells and showed that they are disproportionately located near immune genes. The study goes further to prove a relationship between ERVs and the regulation of an immune gene by focusing on the MER41B ERV family, which was known to be bound by STAT1 [7] (an IFN- $\gamma$ -stimulated transcription factor) and to exhibit the H3K27 histone acetylation that occurs in latent genetic enhancers [8].

One particular MER41B element is located 220 base pairs upstream of absent in melanoma 2 (AIM2), an IFN- $\gamma$ -stimulated gene that encodes a foreign-DNA sensor that activates inflammatory pathways [9]. The authors showed that AIM2 transcripts were absent in IFN- $\gamma$ -stimulated human cells lacking the MER41.AIM2 ERV. They also used a luciferase reporter assay to investigate the specific binding site, showing that point mutations reduced expression, and that the same sites were conserved across primates. The experimental *coup de grâce* of the paper was demonstrating the loss of an inflammatory response to infection by vaccinia virus in MER41.AIM2-negative cells due to the absence of AIM2 expression (Figure 1).

Unlike other described EDIs, the MER41.AIM2 ERV regulates an existing immune gene. Consistent with the idea of an evolutionary shortcut, the ability to bind an IFN- $\gamma$ -stimulated transcription factor already existed in the retrovirus that infected our ancestors, since the STAT1-binding site is present in the MER41B consensus sequence. The authors speculate that STAT1 binding may have helped the retrovirus exploit immune signaling to promote transcription, as is the case in HIV-1 [10]. This would mean that MER41.AIM2 arrived and directly interfaced with our own immunity. Therefore, the same selective forces that allow a virus to better exploit host networks also facilitated exaptation by the host. Conversely, the nearly identical MERV41A group has a deletion in its long terminal repeat that eliminates the STAT1-binding site. Perhaps the same interferon-inducible enhancement exhibited by MER41.AIM2 proved to be detrimental to the hosts and was selected against in MER41A.

One of the least studied aspects of ERVs is the early stage of endogenization, when inherited retroviral integrations spread through the population. ERVs with negligible phenotypic consequences, as well as those that are slightly harmful, will occasionally reach fixation. We presume that most ERVs observed today are the result of these events. However, rare examples of ERVs that are co-opted by the host would facilitate a rapid sweep to fixation, and this presumably was the case with MER41.AIM2. Chuong *et al.* [3] showed that nearly 1,000 of the MER41 elements in humans contain binding sites for STAT1. Moreover, MER41-like retroviruses with STAT1-binding sites have independently colonized several other mammalian lineages (lemurs, vesper bats, carnivores and artiodactyls). The authors also confirmed that consensus sequences of dog and cow binding motifs can drive IFN- $\gamma$ -inducible

activity in HeLa cells. Together, these results suggest a larger-than-appreciated role of positive selection in the colonization of the genome by MER41.

AIM2 is part of the PYHIN gene family of DNA sensors, common to marsupials and placental mammals. In line with its antiviral activity, AIM2 has experienced diversifying selection in the primate lineage in which it acquired interferon-inducible regulation [11]. Orthologs of AIM2 are found in different species, and the PYHIN family has undergone variable expansion, deletion and rearrangements in different lineages; such diversification could offer a route to differential expression and/or function. This phenotypic novelty could be especially advantageous in the arms race, in which different hosts need to rapidly respond to different viral threats at different times. The repertoire of PYHIN genes is even known to vary amongst strains of mice (for example, [12]), and there are copy number variants within humans [13]. Moreover, expression patterns also vary in different species, and some mammals have lost the gene entirely [14,15]. This suggests that AIM2 is not always beneficial, explaining the recruitment of a regulatory ERV. Consistent with this, we also know that, although mouse AIM2 does not have an equivalent ERV regulator, the inflammasome response can be regulated by p202, another double-stranded DNA binding protein [20]. The dynamics responsible for this genotypic diversity are governed by the arms race between viruses and their hosts and, in the case of MER41.AIM2, are accelerated by the co-option of a rapidly evolving virus–host molecular interface.

The dynamic evolutionary history of PYHIN genes is better contextualized in light of recent work on the relationship between ERVs and the cell. Evidence is accumulating for a complex interplay between ERV expression, foreign-DNA sensing and transcriptional regulation. In B cells, ERVs might play a role in triggering RNA-sensing and DNA-sensing pathways that lead to antibody production [16]. Conversely, the involvement of ERVs in transcriptional regulation is not limited to STAT1 binding for innate immunity, and it has been shown that one third of the binding sites for the tumor suppressor p53 are found within ERVs [17]. All together, these results suggest that understanding the dynamics of co-option requires us to examine the balance of positive and negative consequences of the influence of ERVs.

The evolutionary dynamics surrounding EDI recruitment are neither straightforward nor limited to the adaptive consequences for the virus and host. For instance, viruses themselves are in competition with one another, engaged in an arms race of their own. For example, retroviruses and herpesviruses have been co-infecting mammals for millions of years, and herpesviruses have repeatedly captured retroviral immunomodulatory genes, much like how hosts have captured EDIs [18]. Similarly, another linked dynamic is that genomes themselves are involved in an arms race against ERVs that invade and attempt to proliferate [19]. It would be interesting to test the influence of these various dynamics on the co-option of MER41.AIM2 by reconstructing the ancestral virus. However, given that MER41B is highly degraded compared with retroviruses, we cannot determine the ancestral sequence to experimentally test its involvement in the initial co-option. The broad activity of AIM2 against foreign DNA further complicates inference of the ancient epidemic that posed the selective pressure to co-opt MER41.AIM2.

The remarkable variety of EDIs emphasizes the notion of the evolutionary shortcut as an underlying mechanism for host co-option of EVEs. A common feature behind their success is the fact that, as viruses, they had evolved self-serving compatibility to the immune system that is easily usurped by hosts. Above all, the results of Chuong *et al.* [3] are formative because they reveal a systematic and widespread evolutionary process, which includes multiple ERVs being co-opted by several genes across the mammalian phylogeny, including those that are IFN- $\gamma$ -inducible. What such discoveries show is that incorporating the unique dynamics involved behind co-opting products of the arms race will be necessary to elucidate the web of interactions that drive the coevolution of viruses and hosts.

## REFERENCES

1. Feschotte, C., and Gilbert, C. (2012). Endogenous viruses: insights into viral evolution and impact on host biology. *Nat. Rev. Genet.* 13, 283–296.
2. Aswad, A., and Katzourakis, A. (2012). Paleovirology and virally derived immunity. *Trends Ecol. Evol.* 27, 627–636.
3. Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* 351, 1083–1087.
4. Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S. *et al.* (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403, 785–789.
5. Dupressoir, A., Vernochet, C., Bawa, O., Harper, F., Pierron, G., Opolon, P., and Heidmann, T. (2009). Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. *Proc. Natl. Acad. Sci. USA* 106, 12127–12132.
6. Malik, H.S. (2012). Retroviruses push the envelope for mammalian placentation. *Proc. Natl. Acad. Sci. USA* 109, 2184–2185.
7. Schmid, C.D. and Bucher, P. (2010). MER41 repeat sequences contain inducible STAT1 binding sites. *PLoS One* 5, e11425.
8. Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. *Cell* 152, 157–171.
9. Hornung, V., Ablasser, A., Charrel-Dennis, M., Bauernfeind, F., Horvath, G., Caffrey, D.R., Latz, E., and Fitzgerald, K.A. (2009). AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 458, 514–518.
10. Sgarbanti, M., Remoli, A.L., Marsili, G., Ridolfi, B., Borsetti, A., Perrotti, E., Orsatti, R., Ilari, R., Sernicola, L., Stellacci, E., *et al.* (2008). IRF-1 is required for full NF-kappaB transcriptional activity at the human immunodeficiency virus type 1 long terminal repeat enhancer. *J. Virol.* 82, 3632–3641.
11. Cagliani, R., Forni, D., Biasin, M., Comabella, M., Guerini, F.R., Riva, S., Pozzoli, U., Agliardi, C., Caputo, D., Malhotra, S. *et al.* (2014). Ancient and recent selective pressures shaped genetic diversity at AIM2-like nucleic acid sensors. *Genome Biol. Evol.* 6, 830–845.
12. Zhang, K., Kagan, D., DuBois, W., Robinson, R., Bliskovsky, V., Vass, W.C., Zhang, S., and Mock, B.A. (2009). Mndal, a new interferon-inducible family member, is highly polymorphic, suppresses cell growth, and may modify plasmacytoma susceptibility. *Blood* 114, 2952–2960.
13. Fernando, M.M.A., de Smith, A.J., Coin, L., Morris, D.L., Froguel, P., Mangion, J., Blakemore, A.I., Graham, R.R., Behrens, T.W., and Vyse, T.J. (2011). Investigation of the HIN200 locus in UK SLE families identifies novel copy number variants. *Ann. Hum. Genet.* 75, 383–397.
14. Cridland, J.A., Curley, E.Z., Wykes, M.N., Schroder, K., Sweet, M.J., Roberts, T.L., Ragan, M.A., Kassahn, K.S., and Stacey, K.J. (2012). The mammalian PYHIN gene family: phylogeny, evolution and expression. *BMC Evol. Biol.* 12, 140.
15. Ahn, M., Cui, J., Irving, A.T., and Wang, L.-F. (2016). Unique loss of the PYHIN gene family in bats amongst mammals: implications for inflammasome sensing. *Sci. Rep.* 6, 21722.
16. Zeng, M., Hu, Z., Shi, X., Li, X., Zhan, X., Li, X.-D., Wang, J., Choi, J.H., Wang, K.W., Purrington, T. *et al.* (2014). MAVS, cGAS, and endogenous retroviruses in T-independent B cell responses. *Science* 346, 1486–1492.
17. Wang, T., Zeng, J., Lowe, C.B., Sellers, R.G., Salama, S.R., Yang, M., Burgess, S.M., Brachmann, R.K., and Haussler, D. (2007). Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc. Natl. Acad. Sci. USA* 104, 18613–18618.
18. Aswad, A. and Katzourakis, A. (2015). Convergent capture of retroviral superantigens by mammalian herpesviruses. *Nat. Comm.* 6, 8299.

19. Jacobs, F.M.J., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R., and Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* 516, 242–245.
20. Roberts, T.L., Idris, A., Dunn, J.A., Kelly, G.M., Burnton, C.M., Hodgson, S., Hardy, L.L., Garceau, V., Sweet, M.J., Ross, I.L. *et al.* (2009). HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA. *Science* 323, 1057–1060.

**Figure 1. MER41 endogenous retroviral elements regulate the response to viral infection.**

The vaccinia virus triggers IFN- $\gamma$  to stimulate the transcription factor STAT1. Binding of STAT1 to MER41 results in the upregulation of AIM2, which detects the cytosolic DNA genome of vaccinia and initiates an inflammatory response.

**In Brief:**

Endogenous viruses are occasionally co-opted by their hosts to combat other viruses. The discovery of the widespread recruitment of endogenous viruses as regulatory elements for immune genes, points to a systematic evolutionary process in their co-option for host immunity.

