HIV and innate immunity – a genomics perspective

Antonio Rausell[1,2], Paul J. McLaren[1,2,3] and Amalio Telenti[1]*

Addresses: [1]Institute of Microbiology, University Hospital of Lausanne and University of Lausanne, Switzerland; [2]Swiss Institute of Bioinformatics, Lausanne, Switzerland; [3]School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

* Corresponding author: Amalio Telenti (amalio.telenti@chuv.ch)

Abstract

Innate immunity is a theme of increasing interest for HIV research. However, the term is overstretched to cover biological barriers, cellular systems, soluble factors, signaling pathways, and effectors and is inconsistently applied. A clearer semantic classification of the components of innate immunity is needed, which will have direct relevance to the interpretation of human genome variation. Here, we discuss genomic approaches that can assist in re-defining the perimeter of innate immunity. We place particular emphasis on the characteristics of effectors of the intracellular defense against HIV and other pathogens.

Introduction

Due to their collective significance in mediating the host response against pathogens, the genes of the interferon response have been an area of particular focus in the field of antiviral defense. This system includes the induction of several hundred interferon-stimulated genes. Understanding the biology of interferon-stimulated genes is challenging because of the diversity in their specificity and breadth of action against pathogens. Among them, and of considerable interest in the field of HIV research, are the paradigmatic retroviral restriction factors TRIM5α, APOBEC3G, and BST2/Tetherin [1], as well as newly identified factors such as SAMHD1 [2,3] and SLFN11 [4]. A second challenge in studying interferon-stimulated gene biology is understanding the *apparent lack of efficacy* against HIV infection. During chronic infection, a strong interferon response does not correlate with lower levels of HIV viral load [5]. What's more, persistent production of interferon during chronic infection is thought to be deleterious [5-7]. This stands in contrast with the efficacy of exogenous administration of interferon that contributes to active control of HIV infection *in vivo* [8]. Adding to the challenge, antiviral responses can also be triggered by interferon-independent paths [9]. Understanding the protective and deleterious contribution of the innate cellular response to HIV needs a more complete understanding of the components of such defense machinery.

In this report, we highlight approaches from genomics that can help in these endeavors. The emphasis is on the components of intracellular defense; thus, we do not discuss non-cell autonomous systems that are typically considered to be within the innate framework (e.g. NK cells, macrophages, dendritic cells, etc.).

Genes comprising innate immunity?

There are multiple resources that list and categorize components of innate immunity. The Gene Ontology project (http://www.geneontology.org/) [10] includes the term "innate immunity response" (GO:0045087); Innate-DB (http://www.innatedb.ca) [11] identifies curated genes, experimentally verified protein interactions, and signaling pathways involved in innate immunity; and the interferon-stimulated gene database [12] identifies interferon-stimulated genes through expression analyses. Additionally, recent work compiled a list of interferon-stimulated genes that were used for extensive functional analyses in the context of viral infection, including HIV [13]. However, the overlap across these various lists is limited: of 1492 genes included in one or more of the above databases, only 25 are common to all four sets (Figure 1). The reasons for this lack of consensus are many fold: diversity of biology (innate immunity refers to activities spanning many molecular and cellular functions), methodological approaches for gene identification

Figure 1. Overlap of four innate immunity gene sets
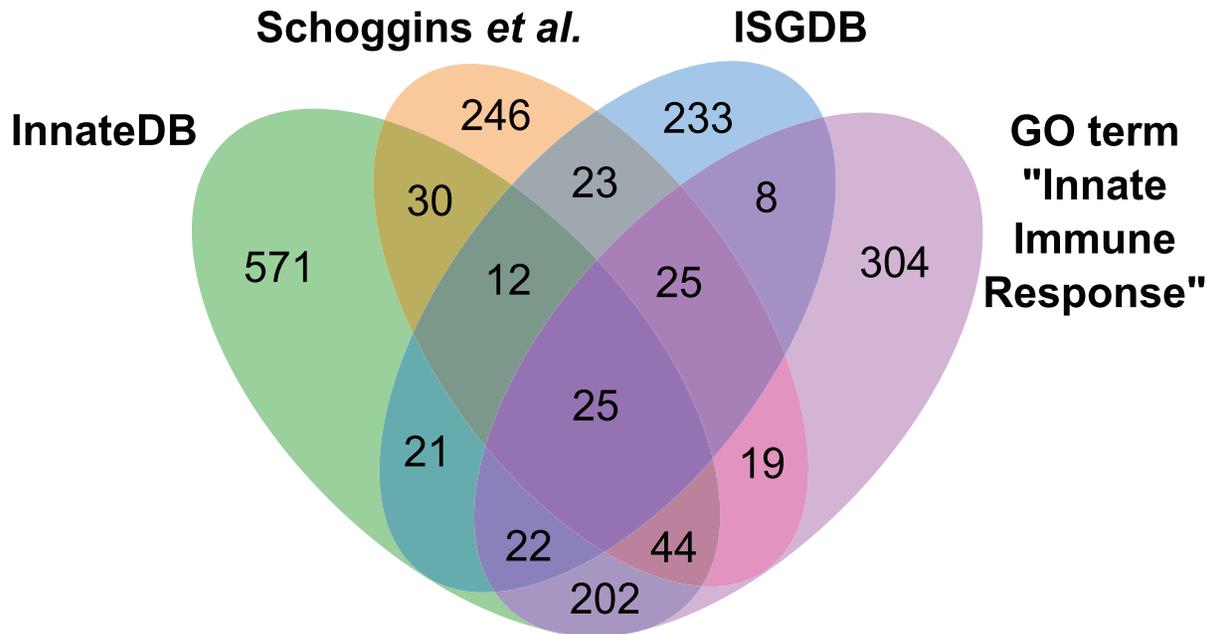
Venn diagram representing four sets of innate immunity human genes: i) In purple, 649 genes associated to GO term "innate immune response" (GO:0045087); ii) in green, 927 manually annotated genes from InnateDB; iii) In blue, 369 interferon-stimulated genes in the interferon-stimulated gene database; iv) in yellow, 424 interferon-stimulated genes genes compiled by Schoggins et al. [13].

(e.g. microarray, functional assays), experimental setups (e.g. different cell lines, stimuli, pathogens), and diverse levels of confidence in annotation. More recently, a number of initiatives have aimed at defining restricted sub-fields, such as that of the intrinsic cellular defense [14] and of cell-autonomous immunity [15]. Clearly, an effort of convergence is needed as experts agree that the "function of the several hundred genes has been comprehensively summarized to only limited extents" [16,17]. A number of recent papers have advanced evolutionary genomics, human genetic approaches, and large scale functional genomic screens that dissect components of innate immunity.

The evolutionary view
It is well accepted that genes and cellular pathways enriched for signals of positive selective pressure are frequently involved in the immune response [18,19]. The underlying concept is that evasion from, and co-evolution with, pathogens is one of the strongest evolutionary pressures, resulting in signals identifiable through comparative genomics. It is expected that genes with such characteristics have an effector role, and that the signals will be most pronounced at domains of direct interaction with a pathogen [20]. Indeed, signatures of positive selection are enriched in the various sets of innate immunity genes (Figure 2A). The HIV restriction factors TRIM5α

APOBEC3G, BST2 and SAMHD1 are relevant examples of genes that have undergone positive selection [20].

Gene expansion (Figure 2B) and, in particular in primates, segmental duplications [21,22] are prominent features of innate immunity genes. The resulting gene duplications may lead to increased gene dosage, and neo- or sub-functionalization [23]. The current state of functional annotation suggests that characterization of duplicated innate immunity genes is largely incomplete. For example, 83% of the 927 genes in InnateDB have paralogs (genes emerging from duplication events). However, most of these paralogs have not themselves been annotated as part of innate immunity, despite many of them showing high levels of sequence similarity. Therefore, estimates of positive selection and patterns of duplication can help establish categories within innate immunity genes. As a corollary, these metrics could serve to annotate genes that have not been previously considered part of innate immunity.

A human genetics view
Exome and whole-genome sequencing in thousands of individuals have revealed large numbers of variants that change amino acid sequences [24]. Increasing numbers of non-synonymous variants are a feature of genes under positive selection, and of genes of innate immunity

Figure 2. Evolutionary pattern of innate immunity gene sets
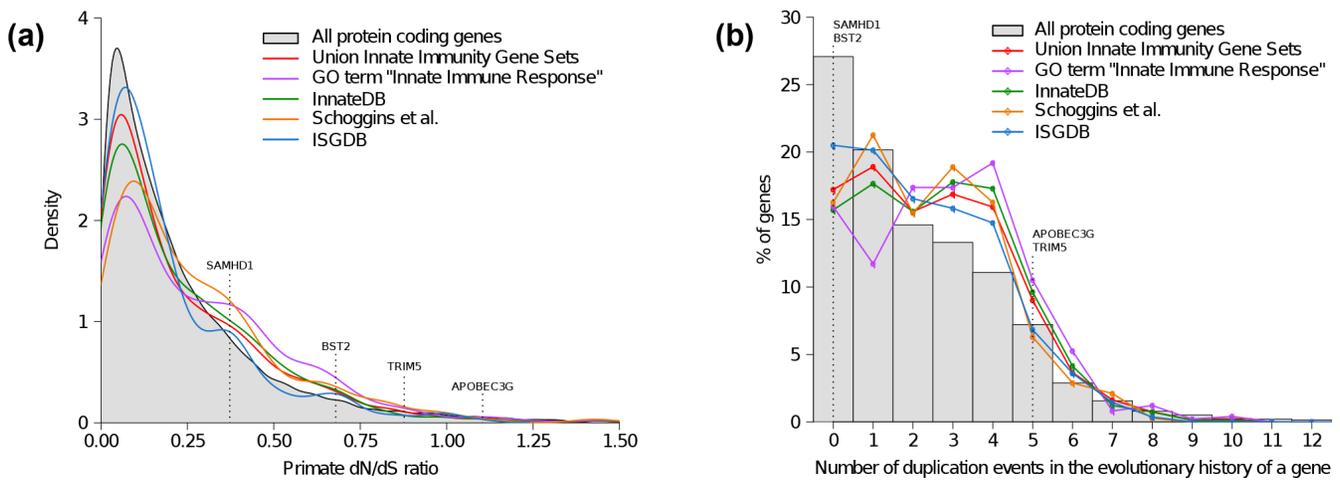
Panel **(a)** In grey, the genome-wide distribution (density) of dN/dS values – a measure of positive selection [34] - for 19252 protein coding genes in primates. Lines depict the distribution of dN/dS values for genes associated with the various innate immunity sets discussed in Fig 1. Panel **(b)** Distribution of duplication events occurring during the evolutionary history of a gene across the various innate immunity gene sets. The histogram depicts the proportion of genes that have none, one or more duplications in the human genome. Dotted lines represent the duplication events for genes associated with the innate immunity gene sets. The values of those measurements for the prototypical innate immunity genes SAMHD1, BST2, TRIM5 and APOBEC3G are indicated.

Figure 3. Burden of human genetic variation in innate immunity genes
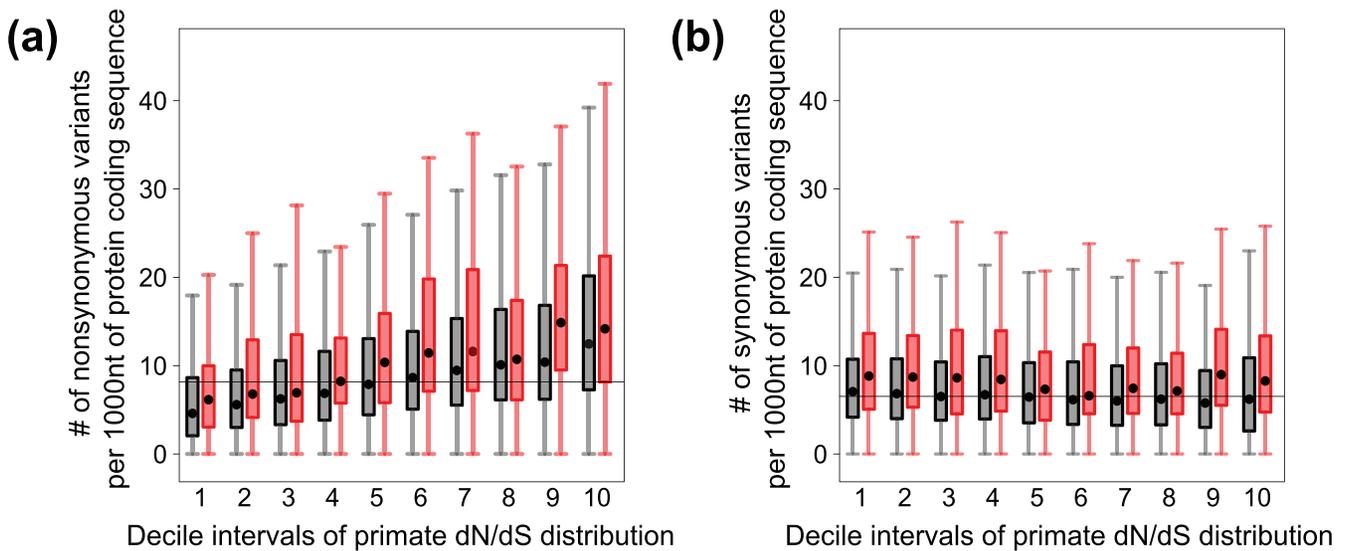
(a) Non-synonymous coding variants in 14213 human genes in the 1000 Genome Project are plotted according to estimates of positive selection in primates. The x-axis distributes genes from the most conserved (lower decile intervals) to the genes under positive selection (higher decile intervals). In red, innate immunity genes (n=1143): the greater the signal of positive selection, the more frequent the identification of non-synonymous variants. In grey, the rest of human genes. **(b)** This trend is not observed for synonymous variants. Horizontal black lines represent median values for the protein coding genome.

(Figure 3). In some instances, variants code for nonsense mutations that, when homozygous, may result in natural knock-outs. We estimate, on the basis of 1092 exome sequences from the 1000 Genomes Project [25], that around 10% of innate immunity genes carry a homozygous stop codon or frameshift variant that may lead to a loss of function. As a correlate to the discussion on gene expansion in the evolution of genomes, there is also

interest in the presence of copy number variation and, in particular, deletions – also enriched in human genes involved in innate immunity and the inflammatory response [26]. The interpretation of these data includes the possibility that greater genetic diversity provides a benefit to the species, i.e. through balancing selection. However, the high frequency of functional variation in innate immune genes could also represent the substrate of human susceptibility to infection – including the possibility of selective immunodeficiency [27]. Exome and whole-genome sequencing to understand rare human variation in the setting of HIV is an important research avenue that will complement the various genome-wide association studies that have been published in the field [28,29].

A functional view

There is significant room for characterization of innate immune genes through the iterative combination of genomic and functional assays. Some commonly applied tools include silencing RNA and gain-of-function screens, large-scale co-immunoprecipitation of interacting host and pathogen proteins in cell lines, and phosphoproteome studies [30,31]. However, it is broadly acknowledged that the interferon response is deficient in many laboratory cell lines – which explains their utility in pathogen research. This observation notwithstanding, the underlying integrity of the cellular innate immune system is rarely considered. For example, RNA sequencing of SupT1 or 293T cells, highly permissive cell lines used in HIV research, shows that they are poorly equipped to respond to the incoming virus [32]. Between 25 and 50% of innate immunity genes are downregulated or not expressed in these cell lines, which stands in contrast with their level of expression in primary CD4+ T cells. Thus, analysis of expression in multiple cellular systems generates a checkerboard of innate immunity genes that are absent in one or more susceptible cell types, but present in cell lines or primary cells that do not support pathogen replication. This fact can be leveraged to further define the perimeter of innate cellular defense. Thus, genes of innate immunity that are missing in susceptible cell lines, and present in primary cells, can be considered as candidates for further investigation.

Conclusions

From a genomic perspective, a number of approaches are useful to characterize innate immunity and thus important to characterize the first barrier of defense against HIV. Positive selection, gene duplication, human genetic diversity, and differential expression across cell lines and primary cells are quantifiable features that point to a differential genomic landscape of effector genes participating in protection against HIV and other pathogens [33].

References

1. Blanco-Melo D, Venkatesh S, Bieniasz PD: **Intrinsic cellular defenses against human immunodeficiency viruses.** *Immunity* 2012, **37**:399-411.
F1000Prime RECOMMENDED

2. Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, Yatim A, Emiliani S, Schwartz O, Benkirane M: **SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx.** *Nature* 2011, **474**:654-7.
F1000Prime RECOMMENDED

3. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, Florens L, Washburn MP, Skowronski J: **Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein.** *Nature* 2011, **474**:658-61.
F1000Prime RECOMMENDED

4. Li M, Kao E, Gao X, Sandig H, Limmer K, Pavon-Eternod M, Jones TE, Landry S, Pan T, Weitzman MD, *et al.*: **Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11.** *Nature* 2012, **491**:125-8.
F1000Prime RECOMMENDED

5. Rotger M, Dang KK, Fellay J, Heinzen EL, Feng S, Descombes P, Shianna KV, Ge D, Gunthard HF, Goldstein DB, *et al.*: **Genome-wide mRNA expression correlates of viral control in CD4+ T-cells from HIV-1-infected individuals.** *PLoS Pathog* 2010, **6**:e1000781.
F1000Prime RECOMMENDED

6. Teijaro JR, Ng C, Lee AM, Sullivan BM, Sheehan KC, Welch M, Schreiber RD, de la, Torre JC, Oldstone MB: **Persistent LCMV infection is controlled by blockade of type I interferon signaling.** *Science* 2013, **340**:207-11.
F1000Prime RECOMMENDED

7. Wilson EB, Yamada DH, Elsaesser H, Herskovitz J, Deng J, Cheng G, Aronow BJ, Karp CL, Brooks DG: **Blockade of chronic type I interferon signaling to control persistent LCMV infection.** *Science* 2013, **340**:202-7.
F1000Prime RECOMMENDED

8. Pillai SK, Abdel-Mohsen M, Guatelli J, Skasko M, Monto A, Fujimoto K, Yukl S, Greene WC, Kovari H, Rauch A, *et al.*: **Role of retroviral restriction factors in the interferon-alpha-mediated suppression of HIV-1 in vivo.** *Proc Natl Acad Sci U S A* 2012, **109**:3035-40.
F1000Prime RECOMMENDED

9. Hasan M, Koch J, Rakheja D, Pattnaik AK, Brugarolas J, Dozmorov I, Levine B, Wakeland EK, Lee-Kirsch MA, Yan N: **Trex1 regulates lysosomal biogenesis and interferon-independent activation of antiviral genes.** *Nat Immunol* 2013, **14**:61-71.
F1000Prime RECOMMENDED

10. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.*: **Gene ontology: tool**

for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**:25-9.

11. Lynn DJ, Chan C, Naseer M, Yau M, Lo R, Sribnaia A, Ring G, Que J, Wee K, Winsor GL, *et al.*: **Curating the innate immunity interactome.** *BMC Syst Biol* 2010, **4**:117.

    F1000Prime RECOMMENDED

12. Der SD, Zhou A, Williams BR, Silverman RH: **Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays.** *Proc Natl Acad Sci U S A* 1998, **95**:15623-8.

    F1000Prime RECOMMENDED

13. Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, Rice CM: **A diverse range of gene products are effectors of the type I interferon antiviral response.** *Nature* 2011, **472**:481-5.

    F1000Prime RECOMMENDED

14. Yan N, Chen ZJ: **Intrinsic antiviral immunity.** *Nat Immunol* 2012, **13**:214-22.

15. MacMicking JD: **Interferon-inducible effector mechanisms in cell-autonomous immunity.** *Nat Rev Immunol* 2012, **12**:367-82.

16. Borden EC, Williams BR: **Interferon-stimulated genes and their protein products: what and how?** *Journal of interferon & cytokine research: the official journal of the International Society for Interferon and Cytokine Research* 2011, **31**:1-4.

17. Randow F, MacMicking JD, James LC: **Cellular self-defense: how cell-autonomous immunity protects against pathogens.** *Science* 2013, **340**:701-6.

18. Kosiol C, Vinar T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A: **Patterns of positive selection in six Mammalian genomes.** *PLoS Genet* 2008, **4**:e1000144.

19. Daub JT, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L, Robinson-Rechavi M, Excoffier L: **Evidence for polygenic adaptation to pathogens in the human genome.** *Mol Biol Evol* 2013, **30**:1544-58.

20. Daugherty MD, Malik HS: **Rules of engagement: molecular insights from host-virus arms races.** *Annu Rev Genet* 2012, **46**:677-700.

    F1000Prime RECOMMENDED

21. Bailey JA, Eichler EE: **Primate segmental duplications: crucibles of evolution, diversity and disease.** *Nat Rev Genet* 2006, **7**:552-64.

22. Lorente-Galdos B, Bleyhl J, Santpere G, Vives L, Ramirez O, Hernandez J, Anglada R, Cooper GM, Navarro A, Eichler EE, *et al.*: **Accelerated exon evolution within primate segmental duplications.** *Genome Biol* 2013, **14**:R9.

    F1000Prime RECOMMENDED

23. Innan H, Kondrashov F: **The evolution of gene duplications: classifying and distinguishing between models.** *Nat Rev Genet* 2010, **11**:97-108.

    F1000Prime RECOMMENDED

24. Gonzaga-Jauregui C, Lupski JR, Gibbs RA: **Human genome sequencing in health and disease.** *Annu Rev Med* 2012, **63**:35-61.

    F1000Prime RECOMMENDED

25. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA: **An integrated map of genetic variation from 1,092 human genomes.** *Nature* 2012, **491**:56-65.

    F1000Prime RECOMMENDED

26. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, *et al.*: **Mapping copy number variation by population-scale genome sequencing.** *Nature* 2011, **470**:59-65.

    F1000Prime RECOMMENDED

27. Quintana-Murci L, Alcais A, Abel L, Casanova JL: **Immunology in natura: clinical, epidemiological and evolutionary genetics of infectious diseases.** *Nat Immunol* 2007, **8**:1165-71.

    F1000Prime RECOMMENDED

28. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, *et al.*: **A Whole-Genome Association Study of Major Determinants for Host Control of HIV-1.** *Science* 2007, **317**:944-7.

29. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, Ripke S, Brumme CJ, Pulit SL, Carrington M, *et al.*: **The major genetic determinants of HIV-1 control affect HLA class I peptide presentation.** *Science* 2010, **330**:1551-7.

    F1000Prime RECOMMENDED

30. Telenti A: **HIV-1 host interactions – integration of large scale datasets.** *F1000 Biology Reports* 2009, **1**:71.

31. Wojcechowskyj JA, Didigu CA, Lee JY, Parrish NF, Sinha R, Hahn BH, Bushman FD, Jensen ST, Seeholzer SH, Doms RW: **Quantitative Phosphoproteomics Reveals Extensive Cellular Reprogramming during HIV-1 Entry.** *Cell Host Microbe* 2013, **13**:613-23.

    F1000Prime RECOMMENDED

32. Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Munoz M, Gunthard H, Beerenwinkel N, Telenti A, Ciuffi A: **24 hours in the life of HIV-1 in a T cell line.** *PLoS Pathog,* **PPATHOGENS-D-12-02353**.

33. Patel MR, Loo YM, Horner SM, Gale M Jr, Malik HS: **Convergent evolution of escape from hepaciviral antagonism in primates.** *PLoS Biol* 2012, **10**:e1001282.

    F1000Prime RECOMMENDED

34. Kimura M: **Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution.** *Nature* 1977, **267**:275-6.