

## PRIMER

## Less is more: Biased loss of CpG dinucleotides strengthens antiviral immunity

Daniel Sauter<sup>1\*</sup>, Frank Kirchhoff<sup>2\*</sup>**1** Institute for Medical Virology and Epidemiology of Viral Diseases, University Hospital Tübingen, Tübingen, Germany, **2** Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany\* [daniel.sauter@med.uni-tuebingen.de](mailto:daniel.sauter@med.uni-tuebingen.de) (DS); [frank.kirchhoff@uni-ulm.de](mailto:frank.kirchhoff@uni-ulm.de) (FK)

*Viruses may not only affect our daily lives but also shape our genome evolution. A recent study shows that the zinc-finger antiviral protein (ZAP) drives CpG suppression in a biased manner. Genes involved in the defense against viral invaders are particularly CpG suppressed to avoid self-targeting and to promote an effective immune response.*



## OPEN ACCESS

**Citation:** Sauter D, Kirchhoff F (2021) Less is more: Biased loss of CpG dinucleotides strengthens antiviral immunity. *PLoS Biol* 19(9): e3001353. <https://doi.org/10.1371/journal.pbio.3001353>

**Published:** September 8, 2021

**Copyright:** © 2021 Sauter, Kirchhoff. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** DS and FK are funded in part by the Deutsche Forschungsgemeinschaft (CRC 1279, SA 2676/3-1 and SPP 1923). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** IFN, interferon; IRG, interferon-repressed gene; ISG, interferon-stimulated gene; ZAP, zinc-finger antiviral protein.

Upon viral infection, our innate immune system needs to initiate efficient defense mechanisms attacking the pathogen without markedly harming the host. Distinguishing self from nonself represents a formidable challenge since all viral components are generated by the host cell. Although viruses and their hosts use the same cellular machineries, viral and cellular RNAs show differences in codon usage, nucleotide composition, and the presence of specific secondary structures. These characteristics are exploited by antiviral host factors to preferentially target viral RNA, while sparing cellular transcripts. However, the establishment of an antiviral mechanism initiates an arms race: Viruses mimic the features of their hosts to evade restriction, while the host will evolve toward more pronounced differences to maintain discrimination between self and nonself. Over millions of years of coevolution, this interplay had a major impact on the composition of host genomes.

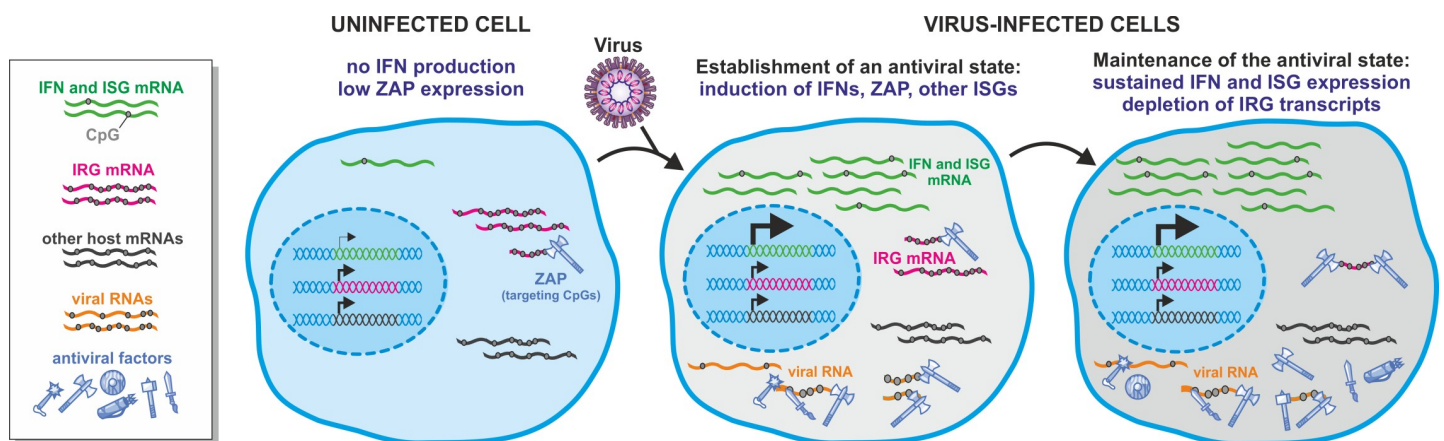
A prime example for this phenomenon is the marked suppression of CpG dinucleotides in mammalian genomes [1]. For example, CpGs are more than 4-fold less frequent in the human genome than expected. Many viruses mimic this host feature [2]. Just recently, the zinc-finger antiviral protein (ZAP) has been identified as a driving force behind this suppression of CpG dinucleotides [3]. It has been shown that ZAP specifically binds to CpG dinucleotides in HIV-1 mRNAs to target them for degradation. ZAP has an ancient origin [4] and is capable of restricting a large variety of viral pathogens including retro-, alpha-, filo-, hepadna-, picorna-, toga-, herpes-, corona-, and flaviviruses [5]. Thus, ZAP may have driven CpG suppression in many viruses and their respective host species. Notably, ZAP itself lacks nuclease activity and requires the RNase KHNYN to degrade viral RNAs [6].

Effective distinction between self and nonself is especially important in the presence of a viral pathogen, i.e., when host defense mechanisms are induced and highly active. Sensing of viral pathogens triggers the production of interferons (IFNs), which induce the expression of hundreds of interferon-stimulated genes (ISGs) and establish an antiviral state of the cell. Since ZAP is also an ISG [7], the degradation of CpG-rich RNAs increases during the innate

antiviral immune response. It would obviously be counterproductive if ZAP suppressed the expression of other antiviral defense factors. Now, Shaw and colleagues revealed how the host avoids detrimental self-targeting of ISG transcripts [8]. By performing comprehensive computational analyses, they demonstrate that on average, ISGs show particularly strong CpG suppression. Similarly, the study revealed that IFN genes are frequently extremely suppressed in CpGs, although this was not observed in all species. In contrast, interferon-repressed genes (IRGs) were frequently relatively CpG rich. Shaw and colleagues even identified CpG content as the most definitive feature distinguishing ISGs from IRGs. These findings suggest that ZAP has a major impact on the composition of IFN-modulated host transcripts, i.e., the interferome.

Shaw and colleagues provide evidence that following IFN treatment, IRGs are frequently suppressed by ZAP, while ISGs are hardly affected. Thus, particularly effective CpG suppression promotes the maintenance of effective IFN responses and antiviral gene expression upon ZAP induction (Fig 1). While the study shows that ZAP accounts for the repression of some IRGs, it remains to be determined whether this repression just represents a tolerable side effect of ZAP or whether their silencing helps to suppress viral replication. It will be a challenging task to clarify this because suppressive effects on IRGs are usually modest (<2-fold), while ISGs may be up-regulated by several orders of magnitude. Notably, ZAP depletion was not only associated with increased expression of relatively CpG-rich genes but also down-regulation of a significant number of genes. The authors speculate that this might be due to increased expression of genes suppressing transcription. However, the underlying mechanisms and the significance of ZAP-mediated effects on cellular gene expression for the antiviral immune response and beyond warrant further study.

Shaw and colleagues analyzed the CpG content of ISGs and IRGs from a variety of species covering over 300 million years of evolution. They provide compelling evidence that CpGs are more rapidly purged from ISGs than from IRGs and other cellular genes. Overall, increased CpG suppression of highly IFN-responsive ISGs was evolutionarily conserved. However, the authors also identified marked differences between different species. For example, canine ISGs



**Fig 1. Biased CpG suppression in cellular genes to maintain efficient innate antiviral immune responses in the presence of the antiviral factor ZAP.** Uninfected cells (left) express a large variety of mRNAs showing significant but varying magnitudes of CpG suppression. Virus infection (middle) triggers the IFN response associated with strong induction of IFNs and numerous ISGs, expressing a large array of structurally and functionally diverse antiviral factors. One of them is ZAP, which mediates degradation of CpG-rich viral as well as cellular mRNAs. Particularly strong CpG suppression allows mRNAs encoding IFNs and antiviral factors to evade degradation by ZAP and, thus, to maintain an effective antiviral state (right). Some cellular genes are suppressed by the IFN response, and this effect is partly dependent on ZAP expression. Whether IRGs might otherwise promote viral replication directly or by attenuating the antiviral immune response remains to be determined. Color coding of viral and cellular RNA species and symbols is explained on the left. IFN, interferon; IRG, interferon-repressed gene; ISG, interferon-stimulated gene; ZAP, zinc-finger antiviral protein.

<https://doi.org/10.1371/journal.pbio.3001353.g001>

contained 3-fold higher levels of CpGs than IRGs, while only marginal differences were found in equines. In addition, IFN genes showed extremely low CpG content in most species but not in horses and were even relatively CpG rich in chickens. Again, it will be of significant interest to determine whether this might be merely due to misannotated sequences, differences in ZAP activity, or evasion from ZAP restriction independently of CpG depletion.

The results of Shaw and colleagues agree with previous evidence that CpG numbers are an important, but are not the only, determinant of ZAP sensitivity. For example, the average CpG content of genes suppressed upon IFN stimulation was remarkably similar in the presence and absence of ZAP. It has been shown that RNA structure and/or the number of CpGs at specific locations of viral RNA rather than overall CpG frequency may determine ZAP sensitivity [9–11]. Notably, ZAP targeting is a highly plausible but not the only possible reason for CpG suppression in highly inducible genes. For example, DNA methylation at CpG dinucleotides is associated with silencing of cellular gene expression. Thus, the need for efficient induction may have contributed to the loss of CpGs in immune-associated genes.

The study by Shaw and colleagues [8] adds to a growing body of evidence that human evolution is under numerous constraints and that our genome has been shaped to a large extent by viruses. In order to maintain effective innate antiviral defense mechanisms, ZAP not only drives CpG suppression but also resulted in a significant bias in the composition of the vertebrate interferome to avoid self-targeting during the antiviral immune response. ZAP is not the only cellular factor targeting viral RNAs. Thus, further studies on how the need to distinguish self from nonself has influenced (and is continuing to influence) the composition of genes in general and may select for bias in specific genes due to differences in evolutionary pressures will be of significant interest.

## References

1. Cooper DN, Gerber-Huber S. DNA methylation and CpG suppression. *Cell Differ.* 1985; 17:199–205. [https://doi.org/10.1016/0045-6039\(85\)90488-9](https://doi.org/10.1016/0045-6039(85)90488-9) PMID: 3902251
2. Karlin S, Doerfler W, Cardon LR. Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? *J Virol.* 1994; 68:2889–97. <https://doi.org/10.1128/JVI.68.5.2889-2897.1994> PMID: 8151759
3. Takata MA, Gonçalves-Carneiro D, Zang TM, Soll SJ, York A, Blanco-Melo D, et al. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature.* 2017; 550:124–7. <https://doi.org/10.1038/nature24039> PMID: 28953888
4. Shaw AE, Hughes J, Gu Q, Behdenna A, Singer JB, Dennis T, et al. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLoS Biol.* 2017; 15:e2004086. <https://doi.org/10.1371/journal.pbio.2004086> PMID: 29253856
5. Nchioua R, Bosso M, Kmiec D, Kirchhoff F. Cellular Factors Targeting HIV-1 Transcription and Viral RNA Transcripts. *Viruses.* 2020; 12:495. <https://doi.org/10.3390/v12050495> PMID: 32365692
6. Ficarelli M, Wilson H, Galão RP, Mazzon M, Antzin-Anduetza I, Marsh M, et al. KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict HIV-1 containing clustered CpG dinucleotides. *Elife.* 2019; 8. <https://doi.org/10.7554/eLife.46767> PMID: 31284899
7. Hayakawa S, Shiratori S, Yamato H, Kameyama T, Kitatsuji C, Kashigi F, et al. ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. *Nat Immunol.* 2011; 12:37–44. <https://doi.org/10.1038/ni.1963> PMID: 21102435
8. Shaw AE, Rihn SJ, Mollentze N, Wickenhagen A, Stewart DG, Orton RJ, et al. The “antiviral state” has shaped the CpG composition of the vertebrate interferome to avoid self-targeting. *PLoS Biol.* 2021. <https://doi.org/10.1371/journal.pbio.3001352>
9. Meagher JL, Takata M, Gonçalves-Carneiro D, Keane SC, Rebendenne A, Ong H, et al. Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad Sci U S A.* 2019; 116:24303–9. <https://doi.org/10.1073/pnas.1913232116> PMID: 31719195
10. Luo X, Wang X, Gao Y, Zhu J, Liu S, Gao G, et al. Molecular Mechanism of RNA Recognition by Zinc-Finger Antiviral Protein. *Cell Rep.* 2020; 30:46–52.e4. <https://doi.org/10.1016/j.celrep.2019.11.116> PMID: 31914396

11. Chen S, Xu Y, Zhang K, Wang X, Sun J, Gao G, et al. Structure of N-terminal domain of ZAP indicates how a zinc-finger protein recognizes complex RNA. *Nat Struct Mol Biol.* 2012; 19:430–5. <https://doi.org/10.1038/nsmb.2243> PMID: 22407013