Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

The depths of virus exaptation

Eugene V Koonin[1] and Mart Krupovic[2]

Viruses are ubiquitous parasites of cellular life forms and the most abundant biological entities on earth. The relationships between viruses and their hosts involve the continuous arms race but are by no account limited to it. Growing evidence shows that, in the course of evolution, viruses and their components are repeatedly recruited (exapted) for host functions. The functions of exapted viruses typically involve either defense from other viruses or cellular competitors or transfer of nucleic acids between cells, or storage functions. Virus exaptation can reach different depths, from recruitment of a fully functional virus to exploitation of defective, partially degraded viruses, to utilization of individual virus proteins.

Addresses
[1] National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, United States
[2] Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Department of Microbiology, Institut Pasteur, 25 rue du Docteur Roux, Paris 75015, France

Corresponding authors: Koonin, Eugene V (koonin@ncbi.nlm.nih.gov), Krupovic, Mart (mart.krupovic@pasteur.fr)

Current Opinion in Virology 2018, **31**:1–8

This review comes from a themed issue on **Viral evolution**

Edited by **Gustavo Palacios** and **Jens H Kuhn**

https://doi.org/10.1016/j.coviro.2018.07.011

Introduction

Parasitic genetic elements are ubiquitous companions of cellular life forms. Theoretical argument and empirical evidence strongly suggest that emergence of parasites is inevitable in replicator systems [1–3]. Moreover, most cellular organisms are hosts to multiple types of genetic parasites that differ with respect to their degree of autonomy and impact on the host [4–6]. Some of the parasitic elements, such as plasmids and transposons, are primarily commensals that reproduce at a low cost to the host [7]. Others are virulent viruses that kill the host. The key difference between viruses and other genetic parasites is that most viruses form virions, specialized particles that encapsidate the viral genome and serve as transmission devices [8••,9]. Many viruses are evolutionarily related to non-viral, particle-less genetic elements such as plasmids or transposons, and transitions between viruses and particle-less elements apparently have occurred in both directions on many occasions [10–14]. Furthermore, genomes of all types of parasitic elements integrate into host genomes, either as part of their life cycle or sporadically. In many animals and plants, integrated parasitic elements, mostly, inactivated ones, account for the majority of the genomic DNA [15,16].

Given the ubiquity of viruses in the biosphere and the billions of years of virus–host coevolution [17,18], it is obvious that the relationships between viruses and hosts cannot be limited to the proverbial arms race. Far from that, genetic material of viruses and other parasitic elements is repeatedly recruited by hosts for various functional roles. The domestication of viruses and exaptation (a concept and term introduced by Gould and Vrba to denote recruitment of a biological entity for a new function unrelated to the original one [19]) of viral genomes, individual genes or smaller fragments for host functions takes many different forms and reaches different depths, with respect to what remains of the virus genome. In this brief review, we discuss several distinct cases of virus exaptation that have been illuminated by recent discoveries.

Proviruses, virophages, and antivirus defense

Viruses with an essential stage of integration into the host genome include several families of reverse-transcribing viruses [20]. Many viruses, especially those with DNA genomes, use a 'bet hedging' strategy whereby a virus switches from a lytic to a lysogenic reproduction mode in which viral genome integrates into the host chromosome and is vertically inherited. Such is the lifestyle of numerous bacteriophages, particularly, those in the family *Siphoviridae*, as exemplified by Enterobacteria phage Lambda, the classic model of molecular biology [21,22]. Most bacterial genomes carry prophages, often, several ones. Are prophages a form of exaptation? Although prophages have been studied for decades, there is still no general answer to this question. Estimation of the fitness values for all microbial genes shows that, on average, prophages are costly [7]. Nevertheless, prophages can protect the host from superinfection by other phages via a variety of mechanisms ranging from modification of the host cell surface or masking phage receptors to active blockage of genome injection of superinfecting phages to repressor-based immunity at the transcription level [23•,24]. Notably, deletion of all prophages from *Escherichia coli* genome has led to decreased fitness of

the bacterium under various environmental conditions [25]. Provirus-mediated superinfection exclusion (also referred to as superinfection immunity or superinfection resistance) is not restricted to bacteriophages but also takes place in the case of retroviral infections [26].

Recently, a group of eukaryotic viruses has been discovered that mimic the prophage life style but shows clear evidence of exaptation. These are the virophages (family *Lavidaviridae* [27]), dsDNA viruses with small genomes that parasitize on giant viruses of the family *Mimiviridae* [28]. When virophage Mavirus infects the flagellate host of the giant helper virus, Cafeteria roenbergensis virus (CroV), it integrates into the cellular genome, without causing any tangible harm to the cell, and remains there 'in waiting' for the infection with CroV [29••,30]. Once CroV infects, the expression of the integrated Mavirus is induced, and the propagating Mavirus abrogates the giant virus reproduction. The infected cell dies nevertheless but the surrounding ones are protected and can be infected by the released Mavirus, perpetuating the protection [29••]. The striking feature of the Mavirus genome that underlies this mechanism of defense is the identity of the Mavirus gene promoters to those of CroV [30]. This is a double adaptation that both allows the virophage to utilize the transcription machinery of the giant virus and to safeguard the uninfected cells in the host population. Thus, the virophage seems to represent a 'transient exaptation' (Figure 1a): it is adopted by the host for a distinct function but remains a full-fledged virus.

Some protist genomes contain multiple insertions of intact and degraded virophage genomes which appear to be traces of an active exaptation process [31]. The phenomenon of virophage exaptation might be much more general. Polintons, self-synthesizing transposons that are common in many diverse eukaryotes [32,33], are clearly related to virophages [30]. Although, originally, polintons have not been considered viruses, it has been shown that they encode major and minor capsid proteins homologous to those of adenoviruses and virophages [34,35]. Virions of the polintons so far have not been observed but the conservation of all the structural elements in the encoded capsid proteins implies that such particles exist. An attractive hypothesis, then, is that all polintons are actually integrated virophages that have been exapted as a mechanism of adaptive immunity against large virulent viruses that, in most cases, remain unknown [14,36]. This hypothesis is compatible with the findings indicating that at least one member of the recently discovered, poorly characterized family of polinton-like viruses (PLV) is a virophage that is associated with *Pheocystis globosa* virus [37,38], and multiple copies of some of the PLV are integrated in algal genomes [38].

Gene transfer agents, polydnaviruses and contractile injection systems: virus-derived vehicles for DNA and proteins

The virophages discussed in the preceding section apparently were exapted by the host for an antivirus defense function but remain viruses that are competent for replication, even if only in the presence of the supporting giant virus. A deeper level of exaptation, where viruses are more highly derived, includes the Gene Transfer Agents (GTAs) (Figure 1b). The GTAs are highly derived, defective prophages that have been studied in greatest detail in the α-proteobacterium *Rhodobacter capsulatus* but subsequently have been identified in diverse bacteria and some archaea [39••]. As shown by genetic methods, the genes of the GTAs are dispersed in microbial chromosomes although the cluster of genes encoding the protein subunits of the phage head and tail stays compact [40•]. The GTAs have been exapted on at least five independent occasions, from different viral lineages [39••]. The key feature of the GTAs is that they generally do not package the prophage genes and instead encapsidate fragments of the host bacterial DNA. However, depending on the GTA family, the encapsidated DNA varies from essentially random DNA fragments of ~4 kb, as in the case of the *R. capsulatus* GTAs [41], to semi-specific packaging of larger genomic fragments of ~14 kb in *Bartonella* GTAs [42]. In the latter case, GTAs preferentially package genes encoding host interaction factors, including secretion systems and putative secretion substrates such as cholera-like toxins, that are amplified from a nearby phage-derived origin of replication [43]. It has been proposed that the specific encapsidation of host-adaptation systems facilitated adaptive evolution and explosive radiation of *Bartonella*, an emerging pathogen [44]. Whereas the heads of the *R. capsulatus* GTAs are too small for encapsidation of the entire phage morphogenetic module (~14 kb), the *Bartonella* GTAs potentially could be self-transmissible [39••]. A recent phylogenomic analysis of the *R. capsulatus* GTAs has shown that the GTA organization was fixed at the base of one of the α-proteobacterial branches, and further, that the GTA genes evolve much slower than the corresponding prophage genes [45•]. These findings indicate that the GTAs are an exaptation that was fixed in bacterial evolution and persisted for a long evolutionary span, presumably as a dedicated device for HGT. Notably, the GTAs are beneficial only at the population level because the individual cells producing GTAs are lysed upon GTA release [46]. This 'altruistic' character of the GTAs obviously mimics the defense function of the virophages discussed above.

Polydnaviruses (PDVs) are a group of unusual insect viruses with genomes consisting of multiple, circular segments of dsDNA [47•,48] that presents a close parallel to the prokaryotic GTAs. The PDVs are mutualistic symbionts of parasitoid wasps that are stably integrated into the wasp genomes. The PDVs appear to have