The popular textbook image of viruses as noxious and selfish genetic parasites greatly underestimates the beneficial contributions of viruses to the biosphere. Given the crucial dependency of viruses to reproduce in an intracellular environment, viruses that engage in excessive killing (lysis) can drive their cellular hosts to extinction and will not survive. The lytic mode of virus propagation must, therefore, be tempered and balanced by non-lytic modes of virus latency and symbiosis. Here, we review recent bioinformatics and metagenomic studies to argue that viral endogenization and domestication may be more frequent mechanisms of virus persistence than lysis. We use a triangle diagram to explain the three major virus persistence strategies that explain the global scope of virus-cell interactions including lysis, latency and virus-cell symbiosis. This paradigm can help identify novel directions in virology research where scientists could artificially gain control over switching lytic and beneficial viral lifestyles.

Also see the Video Abstract: https://youtu.be/E1TOU1JDXo4

Keywords:
- beneficial viruses; persistence triangle; viral domestication; viral endogenization; virus-host interactions

Introduction

Viruses engage in lytic interactions that can destroy the infected cells. Lysis underpins the name “virus” (Latin, venom, poisonous emanation) and supports the dominant view of textbooks and popular press that viruses are noxious parasites, selfish genetic agents and significant threats to crops, livestock, poultry, and human life. A focus on the “panspermic” lytic mode of virus propagation, however, greatly underestimates the global scope of virus-cell interactions and any possible beneficial roles that viruses may play in the biosphere [1, 2]. Viruses can also become dormant through episomal (plasmid-like) or proviral latency mechanisms of cellular and genomic integration, exemplified by lysogeny of temperate bacterial phages [3], and can lead to symbiosis (intimate associations), which can be taken to the extreme in symbiogenesis (fusion of partners) [2]. The three distinct forms of virus-cell interactions, lysis, latency, and symbiosis, can have contrasting long-term evolutionary consequences for both cells and viruses.

Lysis, if successful, often results in cell death and/or observable cytopathic effects under the microscope (e.g. syncytia formation, budding of enveloped viruses). Viral progeny in lytic interactions can be a source of evolutionary innovations and novelties [4] (e.g. evolution of antiviral defense systems [5] and viral mimicry of cellular proteins to escape host immune system [6]), as interactions drive “evolutionary arms races” in both cells and viruses [7]. In turn, latency often results in a state of viral inactivity or cellular integration that is covert, cannot be readily observed under the microscope, and can provide fitness advantages. For example, phage (virus) integration into bacterial chromosomes is known to enhance virulence of bacterial species and is also a mechanism of phage-mediated bacterial gene regulation [8]. Latency also involves domestication of full-length viral genomes or genes for functions beneficial to cells (e.g. [9]). Lytic and latent virus-cell interactions are generally restricted to specific hosts although some viruses can cross species borders [10, 11]. In contrast, symbiosis results in partnerships that can impact organisms separated by large evolutionary distances. Examples include viruses influencing archaecal and bacterial species of the eukaryotic microbiota [12] (similar to known examples of bacterial endosymbions in plants and fungi [2]) and the use of lytic and/or latent properties of viruses by cells to gain a competitive edge against a “third” party (e.g. bacteriophages [13] providing immunity to metazoa against invading bacteria [14]).
Because of the crucial dependency of viruses to reproduce in an intracellular environment, the three forms of virus-cell interactions, lysis, latency, and symbiosis are in conflict. Inspired by a previous explanatory framework of trade-offs of engineering strategies [15], here we propose that these interactions can lead to propagation, dormancy, and dependency trade-off solutions fostering flexibility, robustness, and economy (defined in [15]), respectively, that are beneficial to the long-term evolution of viruses. This triangle of viral persistence (Fig. 1) depicts a “Janus-Faced” balance of power between the lytic pathogenic and the cooperative and more altruistic non-lytic transects. Janus is the Roman God of beginnings, transitions, and time, usually portrayed with two faces, one looking into the future and the other into the past. A global viral quasispecies locates in the trade-off triangle according to its physiology, ecology, and history. Such balance of trade-offs seeks explanation of the “evolutionary dilemma” that too much success is a potential disaster: an organism that drives its prey or hosts to extinction does not survive” [16]. That is, a long-term persistent evolutionary push towards the propagation vertex, as generally believed, could lead to extinction of virus hosts. Thus, the evolutionary opportunities for mutational innovation (flexibility) provided by viral propagation must be offset by counteracting pushes toward the dormancy and dependency vertices (Fig. 1), which foster robustness through cellular latency and economy through sharing of resources between interacting partners, respectively. We speculate that these non-lytic modes are preferred or more frequent outcomes in virus evolution than anticipated but have been greatly underestimated because: (i) they do not yield the phenotypic (cytopathic) effects of viral infection; (ii) sequencing databases hold information for only a tiny fraction of extant viruses (e.g. the human virome is far from complete [17]); and (iii) the current bias is to study viruses of clinical, economical, and agricultural importance (i.e. lytic viruses).

**Is there a preference for latency and symbiosis in the long-term evolution of viruses?**

Evidence for the evolutionary push towards the dormancy and dependency vertices of the triangle (Fig. 1) comes from the historical record. An expected long-term outcome of viral latency is cellular integration of viral genetic material that can be domesticated/coopted by cells [18, 19], suggesting this persistence mode could be a widely employed cellular mechanism. Indeed, the prevalence and abundance of endogenous retroviruses (ERVs) in mammalian genomes [20], evolution of “integration hotspots” in bacterial chromosomes (e.g. prophages) [18], plasmids co-existing harmoniously in diverse prokaryotes, and virus-derived genes in a number of cellular genomes [21–23] provide support to the idea that historically both full-length viral genomes and viral genes have either established permanent residence in hosts or were domesticated/coopted by cells. Additional support comes from the recent metagenomic analysis of diverse microbial communities revealing that increases in microbial abundance were linked to a decline in virus-to-microbe ratio and increases in abundance of hallmark genes involved in viral lysogeny [24]. Although being the first evidence of this kind, the study indicated viral preference for dormancy in special circumstances and offered unique insights into the ecological dynamics of viral lifestyles. In fact, temperate phages tend to dominate viromes extracted from fecal microbiota samples [25, 26] and ~60–70% of sequenced bacterial genomes are estimated to contain prophages [27].
Figure 2. The world of virus-host interactions. **A:** The Venn diagram shows the diversity of viruses infecting the three superkingdoms of life, archaeoviruses (a), bacterioviruses (b), and eukaryoviruses (e), grouped on the basis of common or unique virion morphotypes (modified from [41]). Note that there are only two morphotypes common to all three viral groups and that there are no bacterial-specific morphotypes. **B:** Diagram depicting lysis, latency, and symbiosis interactions of the virosphere. Viruses establish interactions with their hosts in each superkingdom, sometimes becoming latent inside cells (a, b, and e inside cell diagrams). A ring of symbiosis unifies organisms and microbiomes across all superkingdoms providing opportunities of sharing and exchange. Note that archaea and bacteria have a mobilome that is more similar to each other and different from the eukaryotic mobilome.

Virus-cell interactions that push viral persistence towards the dependency vertex can alter the definition of virus “host” and involve virus-cell symbiosis (Fig. 2). Viral host jumps are common (e.g. influenza viruses evolving to infect new species) but are mostly restricted to organisms related by taxonomy. Indeed, no virus is currently known to produce progeny (virions) in organisms belonging to more than one superkingdom. However, this should not mean that viruses do not influence cells they do not lyse. For example, a recent study reported direct virus-metazoan symbiosis limiting pathogenic bacterial growth on mucosal surfaces, a phenomenon apparently conserved from cnidarians to humans [14]. Such symbiotic relationships represent virus-host dependencies involving simultaneous interactions of viruses with organisms from more than one superkingdom (bacteria and eukarya), which in this case provide a novel non-host derived virus-based immunity to metazoa while lytic effects are observed only in bacteria. Such interactions are relatively well documented for bacterial endosymbionts and their host pathogens of other groups of organisms (e.g. [28]). For example, the plant pathogenic fungus *Rhizopus microsporus* harbors the proteobacterial endosymbiont *Burkholderia rhizoxinica*. The endosymbiont produces the virulence factor (rhizoxin) that is antimitotic in nature and arrests cell cycle in the plant host of fungi leading to rice seedling blight disease [29]. It will be intriguing to extend this tripartite interaction between the bacterial endosymbiont, pathogenic fungi, and plants to prophages inserted into the bacterial genomes. Prophages are known to enhance the virulence of their hosts [30] and in doing so can modulate animal and plant microbiomes (e.g. a 3-way virus-fungus-plant symbiosis [31]). More recently, a eukaryotic association module was detected in prophages (bacteriophage WO) inserted in the genomes of the bacterial parasite *Wolbachia* that infects arthropods reporting the first documented example of lateral gene transfers between eukaryotes and bacterioviruses [32]. Another interesting example are polydnaviruses integrated into the genomes of parasitoid wasps [33, 34]. Parasitic wasps have undoubtedly domesticated polydnaviruses [35] using them to coat wasp genes and behaving as holobionts [36]. In another example, ASPE phage genes appear to protect aphids (hosts of endosymbiotic bacterium *Hamiltonella defensa*) against parasitoid wasps [37]. In summary, virus-host interactions do not always yield hallmark phenotypic symptoms of viral infections and can influence hosts they do not lyse. Moreover, integration or domestication of viral genetic elements often benefits the cells, well illustrated by the examples of parasitoid wasps [33, 34] and prophages in bacteria [38]. These observations confront our textbook perception of viruses as selfish genetic parasites and call for a wider recognition of a multiplicity of viral roles including their utility as symbionts and beneficial drivers of host evolution [2, 39]. Their study may also reveal a preference for virus-cell dormancy and dependency [26] that is worthy of exploration.

**Lytic interactions hold deep historical accounts of how superkingdoms customized virospheres by gain-and-loss of viral lineages**

Lytic interactions that drive ongoing evolutionary arm races between cells and viruses (sensu [41]) and push viral persistence towards the propagation vertex of the triangle (Fig. 1) could have
triggered major evolutionary innovations. For example, cellular organisms could have evolved strategies that permanently block some viral infections. If true, there would be strong detectable biases in the distribution of viral replicons in major groups of cellular organisms. Indeed, RNA viruses are either absent in archaea or rare in bacteria, retrotranscribing and RNA viruses are abundant in animal and plant hosts, dsRNA viruses are abundant in Fungi, and DNA viruses are rare in plants (Table 1, see also [40]). These biases hint that virus-cell conflicts have historically led to gain/loss of viral lineages, customizing the virospheres of superkingdoms of cellular life [41]. For example, the ancestors of archaea were likely thermophilic organisms [42, 43] (see [44] for an example phylogeny). Perhaps migration to warmer habitats provided a fitness advantage to ancestral archaeal cells to get rid of the primordial RNA viruses, especially because RNA is quite unstable at extreme temperatures [45]. Similarly, the evolutionary development of a thick peptidoglycan layer of bacterial ancestors that is seemingly impermeable to many viruses could have blocked many viral interactions [46]. Viral persistence could have also driven cellular complexity. The significant abundance and diversity of RNA viruses and retroviruses in eukaryotes (i.e. 55 distinct dsRNA, ssRNA, and retrotranscribing viral families out of total 77) (Table 1) suggest they triggered arms races responsible for eukaryotic organismal complexity [4]. This is especially relevant since RNA and retroviruses are known to mediate genetic rearrangements and induce epigenetic changes [19, 45, 47].

**Viral persistence could have driven cellular diversification**

The strong bias in the distribution of viral replicon types in prokaryotes (mostly DNA) and eukaryotes (mostly RNA) (Table 1) can test scenarios of origin of superkingdoms and viruses (Fig. 3). For example, is the (near-)absence of RNA viruses in prokaryotes due to loss of viral lineages [40] or late de novo gain of viral families in eukaryotes [48]? How to reconcile loss of RNA viruses from prokaryotes under the 3-domain “Woeseian” canonical tree [49], the 2-domain archaeal-ancestor scenario (AAS) [50], or the ring of life models of evolution [51]? Or alternatively, is data compatible with a root of the ToL in the branch leading to archaea [52]?

We speculate that the late origin of a large number of eukaryotic RNA and retroviruses from mixing of prokaryotic viruses [48] seems unlikely because: (i) RNA and retroviruses are likely very ancient and mediated the transition to the DNA world via retrotranscription [53]; (ii) a total of 68 protein fold superfamilies (FSFs) [54] encoded by all seven viral replicon types are present in archaeoviruses, bacterioviruses, and eukaryoviruses (the ae group, Fig. 3) suggesting an origin of viral lineages before the origin of modern cells [55]; and (iii) under the 2-domain AAS or canonical 3-domain trees, eukaryoviruses should exhibit (at least) some overlap with archaeoviruses but only two FSFs (involved in DNA replication/repair and metabolism) and two virion morphotypes (rod-shaped and bacilliform) are shared by archaeoviruses and eukaryoviruses (ae group, Fig. 3). The ae FSFs are coded by dsDNA (and not RNA) viruses while the common morphotypes likely evolved via convergence [51]. Notably, while the archaeal and eukaryotic virospheres appear starkly different, the mobiomes of archaea and bacteria show remarkable resemblances (e.g. common viral families, abundances of plasmids, and 23 common FSFs encoded by dsDNA and ssDNA viruses) in addition to employing common CRISPR-Cas antiviral

---

### Table 1. Counts of viral replicon types (RC) and families (FC) in major host groups

<table>
<thead>
<tr>
<th>Hostb</th>
<th>dsDNA</th>
<th>ssDNA</th>
<th>dsDNA</th>
<th>ssDNA</th>
<th>Retrotranscribing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>FC</td>
<td>RC</td>
<td>FC</td>
<td>RC</td>
</tr>
<tr>
<td>Archaea</td>
<td>40</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bacteria</td>
<td>1731</td>
<td>6</td>
<td>82</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Algae</td>
<td>23</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fungi</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>Invertebrates</td>
<td>103</td>
<td>8</td>
<td>72</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>Invertebrates and plants</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Invertebrates and vertebrates</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Plants</td>
<td>1</td>
<td>1</td>
<td>402</td>
<td>2</td>
<td>62</td>
</tr>
<tr>
<td>Protozoa</td>
<td>12</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Vertebrates</td>
<td>378</td>
<td>7</td>
<td>149</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>Vertebrates and humans</td>
<td>44</td>
<td>5</td>
<td>59</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Vertebrates, invertebrates</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Vertebrates, invertebrates, and humans</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

bUnclassified and unassigned viruses, satellites, viroids, and environmental isolations were excluded from counts. Some viral families and replicons may repeat in more than one group (e.g. Retroviridae and Hepadnaviridae that infect both vertebrates and invertebrates).
Can the persistence triangle help artificial construction of beneficial virus-cell interactions?

Modifying viral persistence by changing the triangle’s trade-offs can have important medical applications. Because the differences in the cellular membrane and molecular biology of the three superkingdoms apparently block viruses from lysing organisms in more than one superkingdom, the artificial construction of virus-host alliances against a “third party” could benefit antimicrobial research. For example, the idea of virus-mediated cleansing of microbiota to treat bacterial infections has gained popularity (reviewed in [25]; see [57] for practical challenges and concerns). The example discussed above where bacteriophages residing in the mucosal membranes of metazoa kill invading bacteria lends additional support to virus-host mutualism [14]. While the idea may seem a distant engineering possibility, the social, molecular, and genetic processes behind such transitions are increasingly becoming better understood. For example, a recent study demonstrated that lytic-to-lyogenic viral switching in microbe-rich seawater samples increased significantly with increases in host density [24]. These results support a “piggyback-the-winner” model and challenge the long-held “killing the winner” model of viral switching.

In another recent study, viruses utilized small peptides to communicate and coordinate decisions about entering into lysis or lysogeny [58]. Perhaps the best-studied lytic-to-lysogenic “switch” is that of the bacteriophage λ, which is dependent on host environment and number of infecting viral particles ([59]; recently reviewed in [3]). Moreover, the virion is the crucial distinction between viruses and plasmids, which is sometimes rooted in the presence or absence of a single capsid-encoding gene [60]. Knockout of capsid genes (a new switch?), along with genes that trigger lysis, could theoretically transform viruses into plasmids and vice versa (e.g. [61]). Given the ongoing metagenomics trends towards the discovery of novel viruses in environmental samples, these distant possibilities may become practical sooner than later but will need to be evaluated on a case-by-case basis to avoid possible viral health side-effects.

Conclusions and outlook

Viruses interact with cells directly and indirectly sometimes involving multiple host layers. These interactions include virus-host symbiosis and virus domestication to pursue common objectives (similar to documented examples of bacteria-eukarya partnerships) and lead to interesting evolutionary, ecological, and social consequences for interacting partners. While much has been written about the lytic virus-cell interactions, a better understanding of the beneficial virus-host partnerships holds enormous clinical and medical value as it opens new doors for therapeutic research in microbiology. An extended survey of the virosphere will help populate the triangle of persistence. Technical demands include accurate viral detection in
metagenomic samples surveyed broadly from geographically diverse habitats. This can be problematic because viruses do not encode a universal gene marker such as ribosomal RNA. The solution could be to focus instead on in silico detection of protein folds present in the viral capsid/coat proteins because capsids have been termed the virus “self” [62] and protein folds involved in capsid assembly tend to be remarkably conserved throughout the virosphere [60]. A shift in strategy may therefore improve viral detection and discovery in environmental samples [63]. It is also important to pursue landmark-sampling efforts to complete the human virome especially from infants, frequent travelers, individuals in contact with livestock and poultry, immunocompromised individuals, and from geographically diverse regions. A long-term objective is to understand how and when viruses switch to the endogenous or endosymbiotic mode. Keeping in mind the crucial dependency of viruses on host cells, we hypothesize that this switching could perhaps be the long-term preferred evolutionary route outcome for viruses. Unlocking novel mechanisms of viral endogenization will require clever integration of bioinformatics and wet lab experiments. It will be necessary to map molecular data (e.g. capsid genes) to virus-host partnerships [55], to identify viral genome integration sites in high-throughput sequencing data (e.g. virusomes) [64] and to simulate behavior of viral particle conformational dynamics under varying conditions using atomic scale molecular dynamics (MD) simulations [66], and to trace the evolutionary spread of viral folds in cellular life [55] to better understand viral lifestyles. We hope that our arguments will encourage an updated thinking about the virosphere, increase interest and focus in discovering non-lytic and beneficial virus-cell interactions (see also [1, 2]), and inspire novel microbiological approaches to study viruses and manage viral diseases.

Acknowledgments

Research is supported by the Higher Education Commission Start-up Research Grant Program to AN (Project No. 21-519/SRGP/R&D/HEC/2014), a grant from KOPRI research program (PE17020) to KMK, and grants from the National Science Foundation (OISE-1132791) and United States Department of Agriculture (ILLU-802-909 and ILLU-483-625) and Blue Waters allocation to GCA.

The authors have declared no conflict of interest.

References


Forterre P. 2013. The common ancestor of archaea and eukarya was not an archaeon. Archites 2013: 372396.


