UC Berkeley
UC Berkeley Previously Published Works

Title

Astrovirology: Viruses at Large in the Universe

Permalink

Journal

ISSN

Authors

Berliner, Aaron J
Mochizuki, Tomohiro
Stedman, Kenneth M

Publication Date

DOI

Peer reviewed

Review Article

Astrovirology:
Viruses at Large in the Universe

Aaron J. Berliner,[1] Tomohiro Mochizuki,[2] and Kenneth M. Stedman[3]

Abstract

Viruses are the most abundant biological entities on modern Earth. They are highly diverse both in structure and genomic sequence, play critical roles in evolution, strongly influence terran biogeochemistry, and are believed to have played important roles in the origin and evolution of life. However, there is yet very little focus on viruses in astrobiology. Viruses arguably have coexisted with cellular life-forms since the earliest stages of life, may have been directly involved therein, and have profoundly influenced cellular evolution. Viruses are the only entities on modern Earth to use either RNA or DNA in both single- and double-stranded forms for their genetic material and thus may provide a model for the putative RNA-protein world. With this review, we hope to inspire integration of virus research into astrobiology and also point out pressing unanswered questions in astrovirology, particularly regarding the detection of virus biosignatures and whether viruses could be spread extraterrestrially. We present basic virology principles, an inclusive definition of viruses, review current virology research pertinent to astrobiology, and propose ideas for future astrovirology research foci. Key Words: Astrobiology—Virology—Biosignatures—Origin of life—Roadmap. Astrobiology 18, 207–223.

1.1. Where Are Viruses in Astrobiology?

W E CONSIDER VIRUSES to be central to multiple aspects of astrobiology, but in the current NASA Astrobiology Strategy (Hays *et al.,* 2015), viruses are mentioned only six times in over 250 pages. The European AstRoMap is slightly more inclusive, with nine mentions of viral elements (two of them parenthetical) in 42 pages but none of viruses (Horneck *et al.,* 2016). Perhaps the astrobiology community is not yet fully aware of the amazing ubiquity, diversity, and roles of viruses in biogeochemistry, evolution, and the origins of life. Therefore, we provide an annotated version of the NASA Astrobiology Strategy Goals with a focus on viruses and emphasize their critical and underappreciated roles in astrobiology. By way of background we also provide an overview of viruses and their replication, while focusing on aspects most relevant to astrobiology.

1.2. An Astrovirology Strategy

The 2015 NASA Astrobiology Strategy provides a list of goals for the astrobiology community which aim to improve research products on the origins, evolution, and future of life on Earth and beyond, while also aiming to develop stron-ger communication across disciplinary boundaries. Table 1 describes these goals and subgoals in relationship to how virology can be considered an integral component of astrobiology.

1.3. The Central Role of Viruses in Astrobiology

The extracellular form of viruses, virions, are the most abundant biological particles on Earth, with an estimated 10^{31} virions in the oceans alone (Suttle, 2005). Whether this amazing virion abundance is true in extraterrestrial oceans or was true in primordial Earth oceans are open questions. Unfortunately, the lack of validated virus biosignatures makes detection of extraterrestrial or ancient viruses challenging. Virus infections profoundly influence both global and local biogeochemical cycling (Suttle, 2009) yet have not been included in ancient climate models or models for extraterrestrial climates. Viruses are critical drivers of current evolution (Enard *et al.,* 2016) and were probably extremely important for the early evolution of life on Earth (Koonin *et al.,* 2006). These roles of viruses in astrobiology are described in more detail in Section 3. For readers who are less familiar with viruses and virus replication cycles, we describe the history of astrovirology, virus replication, virion structures, and genomes in Section 2.

[1]University of California, Berkeley, California, USA.
[2]Earth Life Sciences Institute, Tokyo Institute of Technology, Japan.
[3]Center for Life in Extreme Environments and Biology Department, Portland State University, Oregon, USA.

TABLE 1. ASTROVIROLOGY GOALS IN THE FRAMEWORK OF THE 2015 NASA ASTROBIOLOGY STRATEGY

Original goal	Subgoals	Astrovirology subgoals
Foster Interdisciplinary Science	• Understand linkages between geology, climate, and life.	• Understand the linkages between virology, geology, climate, and life.
	• Coordinate field/mission observations, laboratory experimentation, and theory.	• Coordinate field/mission observations, laboratory experimentation, and theory.
	• Enhance effective communication between scientists and technologists.	• Enhance communication with virologists.
	• Promote collaborations between government offices and agencies.	• Promote collaborations with virologists.
	• Enhance the ability of society to address important challenges that require interdisciplinary approaches.	• Enhance societal understanding of the roles of viruses on our planet.
Enhance NASA's Missions	• Research analog environments to understand processes underlying habitability.	• Develop broader proxies for viruses as indirect evidence of life in analog environments.
	• Develop more effective proxies for life (biosignatures) and environments ("envirosignatures").	• Develop and validate virus biosignatures.
Promote Planetary Stewardship	• Understand interactions between biological activity and environmental change at several spatial scales.	• Understand virus-environment interactions at multiple spatial scales.
	• Support planetary protection efforts.	• Evaluate the roles of viruses in planetary protection.
	• Constrain human exploration (minimize adverse impacts on other planetary environments).	
Enhance Societal Interest and Relevance	• Act upon the public interest to understand our place in the Universe and broaden understanding of life.	• Increase understanding of our place in the Universe through understanding our relationships with viruses and their past and present effects on evolution and life.
	• Sustain the human imperative to explore the unknown.	• Explore the mostly unknown world (or worlds) of viruses.
Inspire Future Generations	• Enhance the interest in science and technology to the benefit of our society.	• Increase interest in science and technology by including (astro)virologists.
	• Enhance critical and creative thinking in all fields of inquiry.	• Increase benefits to society through application of astrovirology technologies.
		• Incorporate virology perspectives to enhance critical and creative thinking.

1.4. The History of "Astrovirology"

The first to publicly connect the study of viruses and astrobiology was the late Baruch Blumberg, discoverer of the hepatitis B virus, Nobel laureate, and first director of the NASA Astrobiology Institute (NAI), who made the connection at the Astrobiology Science Conference in 2002. Thereafter, NAI formed the Virus Focus Group, NAIViFoG. NAIViFoG has organized a number of workshops and symposia mostly focusing on virus origins, ecology, and evolution (Stedman and Blumberg, 2005). The first use of the term "astrovirology" to our knowledge was as a session and abstract title at AbSciCon 2004 (Blumberg, 2004). Dale W. Griffin used "Astrovirology" as a subheading in his 2013 essay in *Astrobiology:* "The Quest for Extraterrestrial Life: What about the Viruses?" We completely agree with his statement: "We should be looking for viruses in our quest for extraterrestrial life" (Griffin, 2013), but, to our knowledge, very few astrobiologists have even started looking.

2.1. What Is a Virus?

Before we start "looking for viruses in our quest for extraterrestrial life" (Griffin, 2013), however, we need to agree on what a virus actually is. The original virus definition was practical: a disease-causing particle that passed through a filter that retains all known bacteria *"contagium vivum fluidum"* (contagious living fluid) (Beijerinck, 1898). The standard definition of viruses for many decades thereafter was "a very small disease-causing agent." Later, there were debates whether they were infectious liquid or particles, but the development of the electron microscope in the 1930s revealed that all viruses were found to have an inert extracellular particle state, the virion. Virions consist of a protein shell (also known as a capsid) that packages the viral genomic nucleic acid, which can be composed of DNA or RNA. In addition, some viruses can have a lipid-containing outer layer, called an envelope. Sir Peter Medawar, Nobel laureate and pioneering immunologist, is said to have called viruses "a piece of bad news wrapped up in a protein." All viruses must transfer their genome inside a host cell and reprogram that cell to produce more virus. This discovery led to the expansion of the standard virus definition to "a very small *obligate intracellular* parasite" (Acheson, 2011). However, as discussed below, other than "obligate intracellular" this definition is incomplete. A more inclusive definition of viruses, that we prefer, was proposed by Salvador Luria and coauthors: "Viruses are entities whose

genomes are elements of nucleic acid that replicate inside living cells using the cellular synthetic machinery and causing the synthesis of specialized elements that can transfer the viral genome to other cells'' (Luria *et al.,* 1978).

2.2. Viruses Are Not Just Virions: Are Viruses Alive?

It is important to make the distinction between a virus and a virion. The inert virion, the form visible under the electron microscope, is analogous to a seed or a spore that can only replicate in an appropriate environment, in the case of a virus inside a host cell. *Virus* refers to the whole virus replication cycle (see Luria's definition above). Unfortunately, most people, including many virologists, conflate the two. Considering the virion to be a virus is analogous to considering an acorn an oak tree or a spermatozoid a human (Claverie, 2006; Forterre, 2010). When one considers the whole virus replication cycle (Fig. 1A), it comes close to NASA's working definition of life: ''A self-sustaining chemical system capable of Darwinian evolution.'' Viruses are chemical, they are capable of Darwinian evolution, but they are not ''self-sustaining'' per se, as they require infection of a living cell for their replication. For further discussion of the vital nature of viruses, see the works of Schulz *et al.* (2017) and Koonin and Dolja (2014). Whether viruses are alive or not may be a moot question, but if a virion (or a virus-like particle) were to be unequivocally detected in an extraterrestrial sample, very few people would claim that this would not be evidence for life—wherever that sample was from.

2.3. Virus Replication Cycle

Following Luria's description, a basic virus replication cycle consists of (1) binding of a virion to a host cell, (2) delivery of the viral genome into the host cell, (3) reprogramming of the cellular machinery to produce more virions, and (4) release of virions from the host cell to infect other cells. Then the process repeats. As a result, while cellular organisms generally replicate by binary fission, with one mother cell resulting in two progeny cells either by division or budding, one virus replication cycle can result in tens or hundreds of progeny virions. A schematic virus replication cycle is shown in Fig. 1A. Viruses thus present a different replication mode than cellular life, prompting possible reevaluation of life-detection efforts.

2.4. Virions, Virus-Specific Biomarkers?

Luria's ''specialized element,'' the virion, is the defining characteristic of viruses (Fig. 2). The exterior of a virion is the virus capsid, which in most cases is made up of repeating substructures termed capsomeres. Virions vary in size from a diameter of approximately 20 nm to over 1 μm for the giant viruses. Some virions have an additional lipid layer or membrane, the envelope, outside the protein capsid. Virus capsids generally have very simple elegant geometries, either with polyhedral symmetry (often icosahedral), similar to a soccer ball (*e.g.,* rhinovirus, known to cause the common cold), or helical symmetry (*e.g.,* tobacco mosaic virus). Another common virion structure is the pleomorphic structure, often found with enveloped viruses (*e.g.,* influenza virus) (Fig. 2). In addition, many virions have additional proteins on their surface that are used to recognize and bind to the host cell surface. These proteins can often be found at the vertices of polyhedral virions, either one or both ends of helical virions, or on the surface of enveloped virions.

Polyhedral and helical structures are widespread in the viral world, primarily for geometrical reasons. Both of these structures can be formed with a limited number of capsid protein subunits and can increase their volume relatively easily by using the same proteins repeatedly. This limits the number of different capsid protein genes required to form the virion and thus maintains a compact viral genome. Particles made of substructures arranged with icosahedral symmetry are highly efficient ways of forming large-volume three-dimensional structures with repeating subunits. Helical symmetry not only matches nucleic acid symmetry but also allows capsid protein subunits to have identical interactions with each other, except at their ends (Crick and Watson, 1956). The geometrical efficiency of virus capsids is thought to be due to the fact that nucleic acid is relatively inefficient for encoding proteins (three nucleotides required for each amino acid) and so is relatively large. Therefore, using one or a few gene products multiple times and arranging them symmetrically is an optimal strategy for limiting the genome size. Interestingly, some of the viruses with the largest genomes also have icosahedral virions (Wilson *et al.,* 2009; Fischer *et al.,* 2010; Aherfi *et al.,* 2016). There are some notable exceptions to this geometric conservation ''rule,'' particularly those of the ''extreme'' viruses of Archaea (see Section 2.5 below).

Given that virion structures are unique and distinctive, it is tempting to use virions as biosignatures. However, to visualize most virion morphologies, a transmission electron microscope (TEM) is required. While this may not be difficult on Earth, it seems unlikely that a TEM will be put on a spacecraft in the foreseeable future. There has been progress with scanning electron microscope (SEM) technology for space exploration, but currently this SEM does not have high-enough resolution to identify virion structures (Edmunson *et al.,* 2016). There has also been progress in enumerating virus-sized particles using dynamic light scattering, nanoparticle tracking analysis, and promising nanopore technology (Schmidt and Hawkins, 2016; Yang and Yamamoto, 2016), but, to our knowledge, none of these techniques have been applied to natural samples containing unknown virions. Moreover, these techniques provide minimal shape information, limiting their use for definitive virion identification. However, these techniques could be useful to pre-sort virus-sized particles for downstream processing possibly including sequencing. They are also more adaptable to mission instrumentation than a TEM.

2.5. Astonishing Virion Diversity in Extremophiles

Viruses of bacteria are dominated by so called head-tail bacteriophages (Fig. 2J). They are taxonomically classified into three viral families, Myoviridae, Podoviridae, and Siphoviridae (Fig. 3), differentiated by their tail structures, and together form the order Caudovirales (King *et al.,* 2011). These head-tail viruses have an icosahedral ''head'' filled with a double-stranded DNA (dsDNA) genome, connected to a linear cylindrical ''tail'' which is used to attach and inject the viral genome into the host bacterium
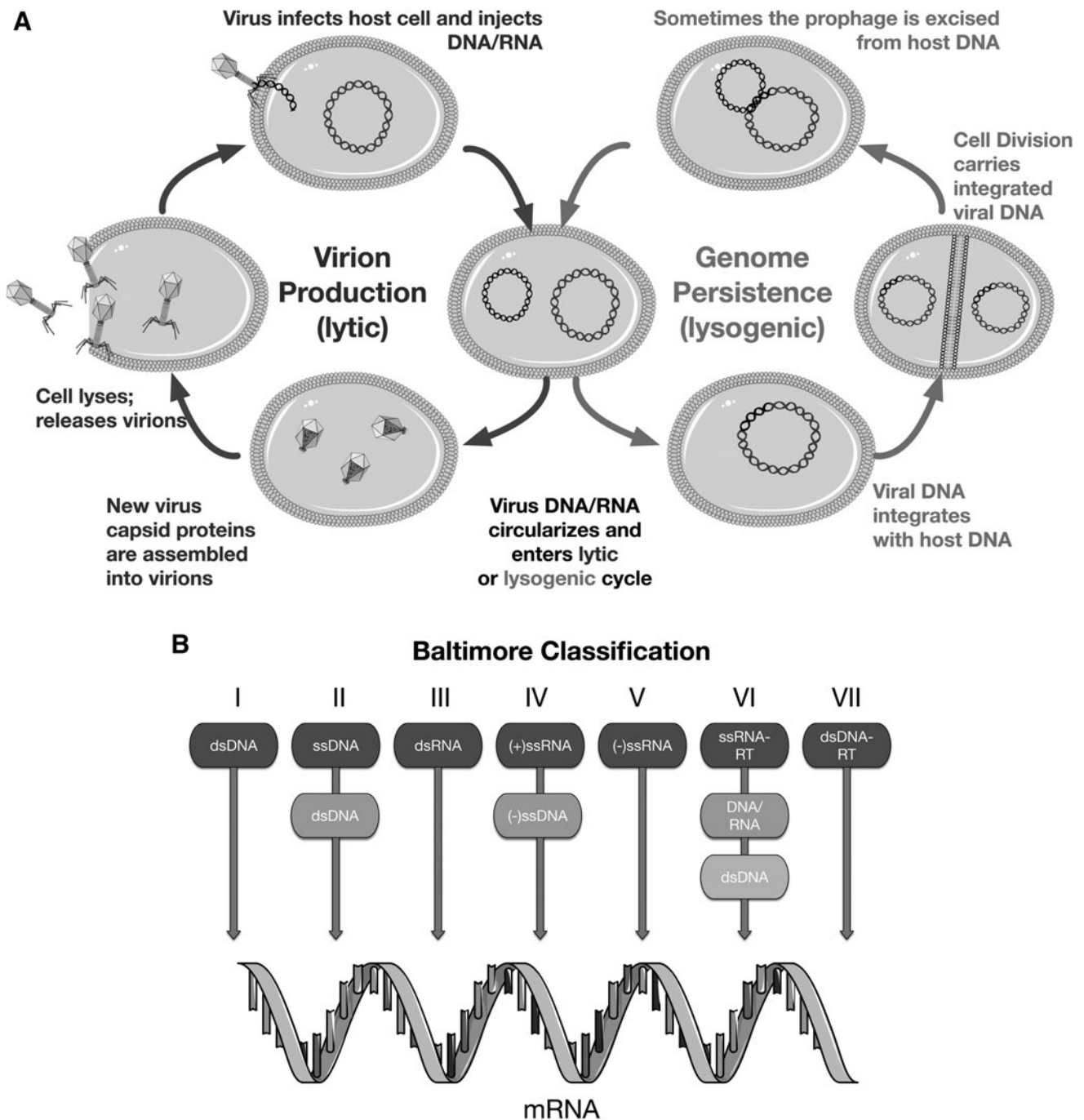
FIG. 1. (**A**) Schematic virus replication cycle. Upper left: virion inserts the viral genome into a host cell. In many cases (left arrows), the virus genome is replicated and packaged in newly formed virions, often lysing the host cell in the process. In other cases (right arrows), the virus genome persists in the host cell during cellular replication. Sometimes the viral genome is integrated into the host genome (circle), forming a provirus or prophage. The integrated virus genome replicates with the cellular replication cycle. The provirus (prophage) can be induced to produce new virions, often entering the lytic cycle. (**B**) The Baltimore classification of viruses by packaged nucleic acid type and route for producing RNA (Baltimore, 1971). The type of nucleic acid genome packaged in virions, RNA or DNA, and single stranded (ss) or double stranded (ds), is indicated in upper ovals. Positive (+) strand refers to the coding sequence of the mRNA. Negative (−) refers to the complement of the mRNA. Nucleic acid intermediates used to form mRNA are shown in lower ovals. Group I comprises dsDNA viruses; group II comprises ssDNA viruses; group III, dsRNA viruses; group IV, positively stranded (+)ssRNA viruses; group V, negatively stranded (−)ssRNA viruses. These five groups correspond to non-retroviral viruses. The discovery of retroviruses (RT) that replicate their genomes through both DNA and RNA intermediates led to the establishment of group VI, which includes (+)ssRNA-RT viruses, and group VII, which includes dsDNA-RT viruses which package ssRNA or dsDNA in their virions, respectively.
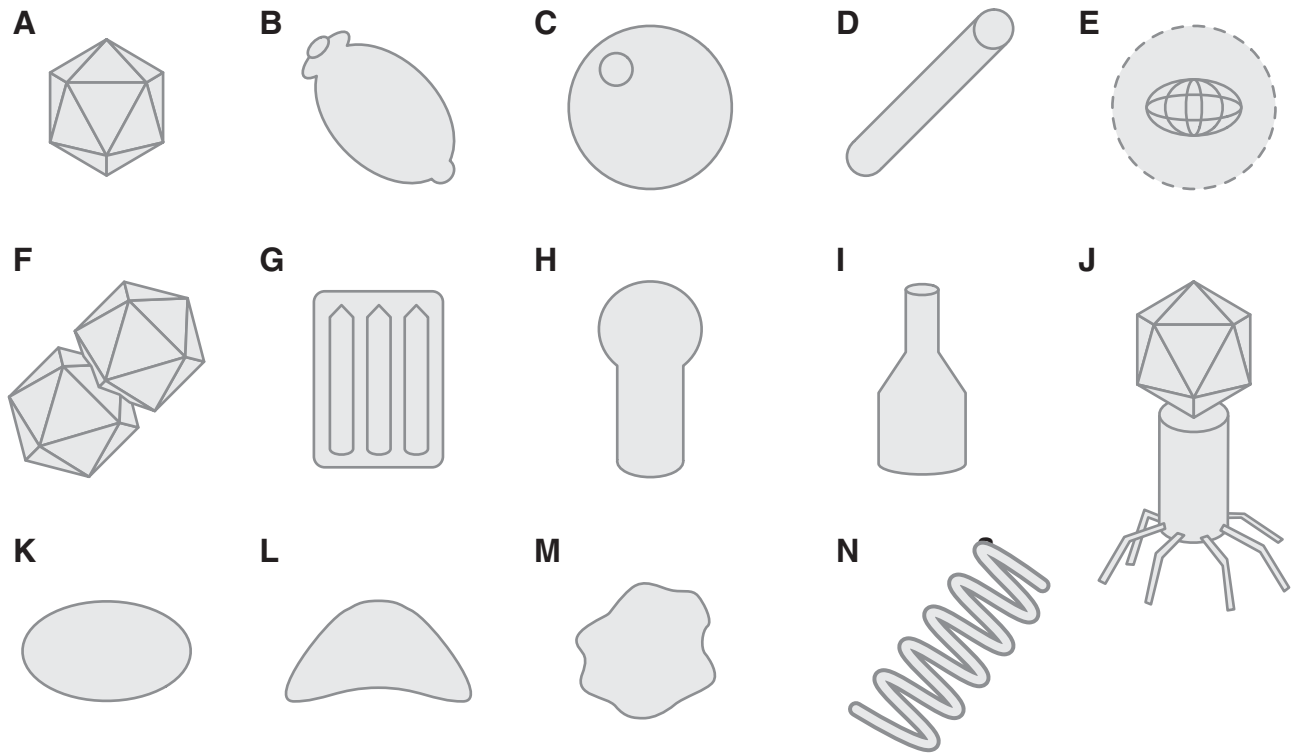
FIG. 2. Generalized representations of virion morphologies: (**A**) icosahedral, (**B**) lemon- or spindle-shaped, (**C**) spherical, (**D**) filamentous or helical or rod-shaped, (**E**) prolate ellipsoid with or without envelope (dashed line), (**F**) twinned icosahedra, (**G**) occluded, (**H**) budded, (**I**) bottle-shaped, (**J**) head-tail, (**K**) ovoid, (**L**) allantoid, (**M**) pleomorphic, (**N**) spring- or coil-shaped.

upon infection. Direct TEM observation of virus-like particles (VLPs) from soil or aquatic environmental samples shows that over 95% of these nano-sized (>500 nm in diameter) particles have the head-tail morphology (Ackermann and Prangishvili, 2012). A few bacterial viruses with other virion structures are known, such as filamentous or tailless polyhedral structures. Bacterial viruses are presently classified into nine taxonomic families (King *et al.,* 2011). Most eukaryotic virions contain some variation of the icosahedral (Fig. 2A) or helical (Fig. 2D) morphology, and many are enveloped with rounded or pleomorphic morphology (Fig. 2E, 2K, 2M).

On the other hand, archaeal viruses have surprisingly diverse morphology. Some archaeal viral morphotypes are found in bacterial or eukaryal viruses, such as polyhedral or icosahedral (Turriviridae, Sphaerolipoviridae) (Fig. 2A) (Rice *et al.,* 2004; Bamford *et al.,* 2005), helical (Rudiviridae, Lipothrixviridae, Clavaviridae, and recently proposed "Tristromaviridae") (Fig. 2D) (Janekovic *et al.,* 1983; Zillig *et al.,* 1993; Arnold *et al.,* 2000a; Mochizuki *et al.,* 2010; Rensen *et al.,* 2016), globular/spherical (Globuloviridae) (Fig. 2C) (Häring *et al.,* 2004), and pleomorphic (Pleolipoviridae) (Fig. 2M) (Pietilä *et al.,* 2009). In addition, many archaeal viruses also display morphologies that were never observed previously. These include lemon- or spindle-shaped (Fuselloviridae) (Fig. 2B) (Martin *et al.,* 1984: Wood *et al.,* 1989; Bath and Dyall-Smith, 1998), tailed spindle-shaped (Bicaudaviridae) (Häring *et al.,* 2005a), droplet-shaped (Guttaviridae) (Fig. 2B) (Arnold *et al.,* 2000b), bottle-shaped (Ampullaviridae) (Fig. 2I) (Häring *et al.,* 2005b), and coil- or spring-shaped (Spiraviridae) (Fig. 2N) (Mochizuki *et al.,*

2012) viruses. Together, despite the much smaller number of isolates, viruses of the Archaea are currently classified into 15 ICTV-approved families—well exceeding the nine bacterial virus families (Fig. 3). Environments where Archaea dominate, such as hot springs (Rice *et al.,* 2001; Prangishvili, 2003; Häring *et al.,* 2005b; Bize *et al.,* 2008; Mochizuki *et al.,* 2010) or salt lakes (Sime-Ngando *et al.,* 2011) are filled with remarkably diverse VLPs. The Archaea-specific lemon- or spindle-shaped viruses (Fig. 2B) are frequently isolated from similar Archaea-rich environments, from acidic (Martin *et al.,* 1984; Prangishvili, 2003) and neutral hot springs (Mochizuki *et al.,* 2011), deep sea hydrothermal vents (Geslin *et al.,* 2003: Gorlas *et al.,* 2012) and salt lakes, and isolated from both crenarchaeal and euryarchaeal hosts. Archaeal viruses with virions similar to the classical head-tail bacterial viruses (Fig. 2J) have also been observed and isolated, mainly from euryarchaeal halophiles (Atanasova *et al.,* 2012), but no members of the Caudovirales have been isolated from thermophilic archaea. However, Caudovirales have been found in thermophilic bacteria, infecting hosts such as *Thermus* growing around 60–80°C (Yu *et al.,* 2006). Any astrovirological virion detection method must encompass this diversity.

2.6. Even Giant Viruses Have Icosahedral Virions

The "small" size of viruses, more specifically virions, has been challenged with the discovery of the so-called giant viruses, also known as "giruses" or Megavirales. This started with the startling discovery of the Mimivirus that infects amoeba in 2003 (La Scola *et al.,* 2003). The Mimivirus has a virion and genome as large as some bacteria. Due
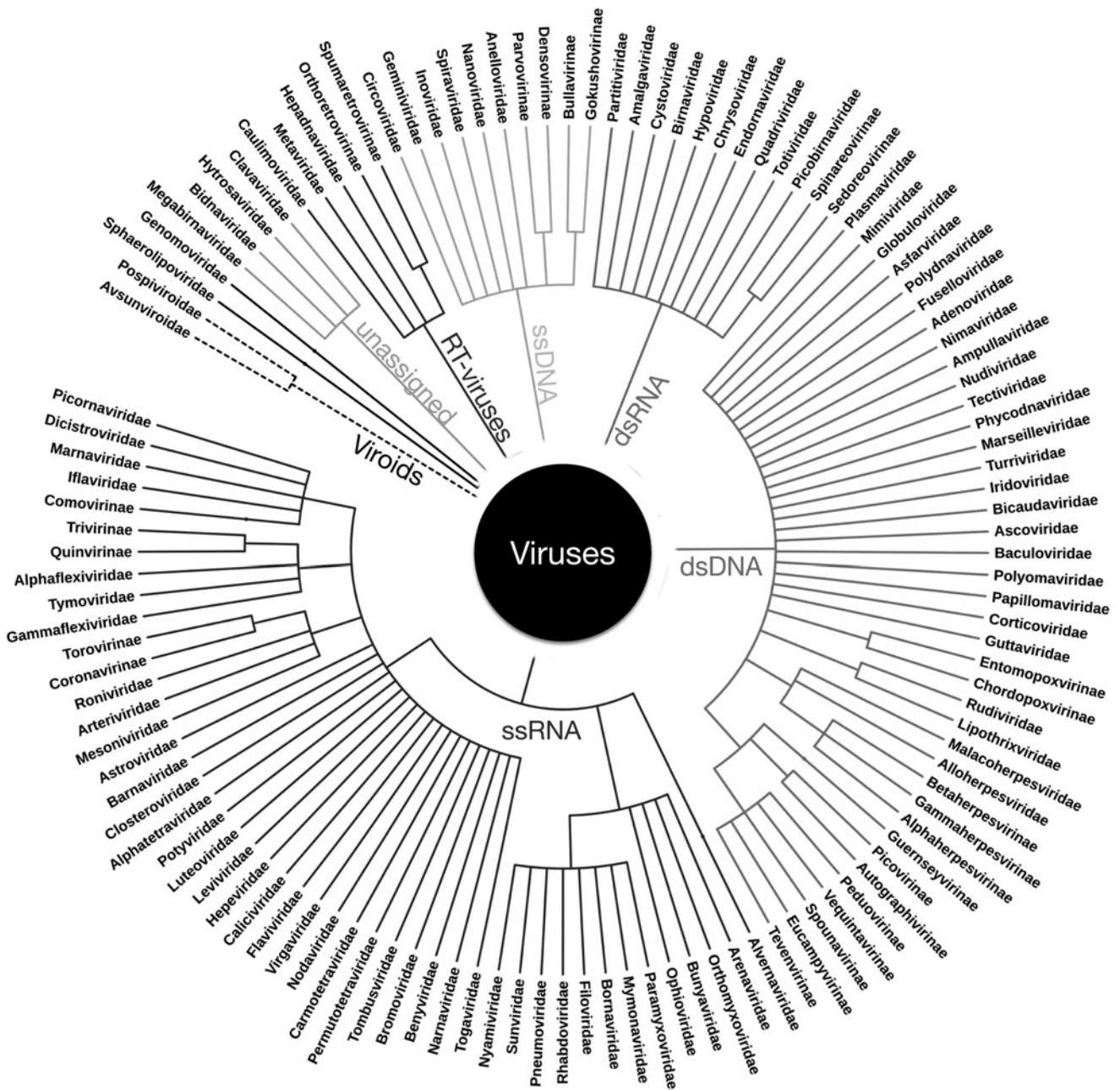
FIG. 3. Virus taxonomy: In order to provide a uniform virus nomenclature, viruses have been classified taxonomically since the 1970s by the International Committee on Taxonomy of Viruses (ICTV) (King *et al.,* 2011). The fundamental ICTV definition of a virus species is ''a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche where the polythetic class is one whose members have several properties in common.'' Virus species differ from the higher viral taxa, which are universal classes defined by strict inclusion requirements. The ICTV classifies viruses in species, genus, (subfamily), family, and in some cases order (King *et al.,* 2011). It is not clear how closely ICTV taxonomic classification reflects phylogeny, particularly given rampant recombination between genes in different viruses. This figure represents the ICTV taxonomic grouping of viruses with viral families and subfamilies, and they are grouped by their packaged nucleic acid and replication as shown in Fig. 1B.

to its large size, Mimivirus was mistakenly identified as a bacterium, *''Bradfordococcus,''* for many years, and only later determined to be a virus. This virus was named Mimivirus because some of its features, including its size, ''mimicked'' bacteria, yet most of these giant viruses still have icosahedrally symmetric virions. The discovery of giant viruses continues, with the recent findings of Pandoravirus with a 1 µm

virion and 1.9–2.5 million base pair (bp) genome and Pithovirus with a virion over 1.5 µm and a 600 thousand bp genome (Legendre *et al.,* 2014). Surprisingly, the genome of the Pandoravirus and the virion of the Pithovirus are larger than many prokaryotic cells and even some small eukaryotes (Philippe *et al.,* 2013). Some giant viruses are parasitized by other viruses, the virophage. The first of these to be discovered

was Sputnik (La Scola *et al.,* 2008) Other virophage are closely related to and may be derived from transposable elements (Fischer and Hackl, 2016). The presence of virophage and their relationships to capsidless transposable elements further blurs the distinction between viruses and capsidless parasitic replicators (Koonin and Dolja, 2014).

2.7. Viruses Are Astronomically Abundant, Particularly in Terrestrial Oceans

Regardless of the debate whether or not viruses are to be considered alive, viruses are critical for structuring Earth's biosphere. They are extremely abundant, diverse, and play important roles in global biogeochemical cycles. Only in the last few decades have scientists realized that viruses outnumber cellular life-forms on our planet by at least 10 fold, likely even more (Suttle, 2005). A number of researchers have shown, based on nucleic acid content or TEM enumeration, that in 1 mL of seawater, in general, there are from 1–10 million viral particles (Bergh *et al.,* 1989; Proctor and Fuhrman, 1990; Brum *et al.,* 2013). Viruses have been found in all environments where life has been found on Earth, for example, the Arctic sea ice (Wells and Deming, 2006), boiling acid hot springs (Rachel *et al.,* 2002; Prangishvili, 2013), deep sea hydrothermal vents (Geslin *et al.,* 2003) and the submarine mantle (Nigro *et al.,* 2017). All cells are expected to have infectious viruses associated with them. The latest estimation is that one bacterial cell can be infected by 10 viral species on average. Based on these numbers, estimated total numbers of virions on planet Earth are truly astronomical, with a calculated 10^{31} virions in the oceans alone (Rohwer, 2003; Suttle, 2005). Given the extremely large numbers of viral particles in Earth's oceans, it would be very interesting to determine if there are virions in recently observed water plumes from Europa (Sparks *et al.,* 2016) and previously observed plumes from Enceladus (Hansen *et al.,* 2008). Unfortunately, to our knowledge, no mission is currently planned that will screen for extraterrestrial viral particles in these plumes. The techniques to be used to screen these plumes for virions are unclear at this time, but nanopore and nanoparticle imaging or novel microscopy techniques may be applicable to this endeavor.

2.8. Virus Genomes—The Dark Matter of Sequence Space

Viruses are unique in the current biosphere in that they use either DNA or RNA in single- or double-stranded forms for their genetic material. Without exception, cellular genomes are exclusively composed of dsDNA. Some viruses only use RNA for their entire replication cycle, leading some researchers to postulate that they are remnants of, or direct descendants from, the hypothetical "RNA world" (Koonin *et al.,* 2006). Virus genomes can range in size from less than 2000 bases of single-stranded DNA (ssDNA) in some circoviruses to almost 2.5 million bp of double-stranded DNA (dsDNA) in Pandoravirus (Aherfi *et al.,* 2016). The smallest virus genomes encode only two genes, a capsid protein gene and a protein required for virus genome replication. By contrast, some of the largest virus genomes encode genes previously thought to be only present in cellular genomes, for example aminoacyl tRNA synthetases (Raoult *et al.,* 2004). However, no virus genome isolated to

date contains a ribosomal subunit protein or ribosomal RNA gene (Schulz *et al.,* 2017). All viruses are thus absolutely dependent on cellular translation machinery to produce virus proteins. This ribosome dependence led to a recently suggested viro-centric definition of cells and viruses; cells are ribosome-encoding organisms, whereas viruses are capsid-encoding organisms (Raoult and Forterre, 2008; Forterre, 2010).

Plants are known to harbor many so-called viroids, which are small structured pathogenic RNA molecules that are made by cellular DNA-dependent RNA polymerase enzymes that do not encode capsids. Current viroids are unlikely to be derived from ancestral RNA parasitic replicators, as they are confined to plants and depend on cellular RNA polymerases, but they could serve as models for archaic parasitic RNA replicators (Koonin and Dolja, 2014).

Many viral genomes contain genes that have no homologs except in closely related viruses. When a viral genome of an unknown group (*i.e.,* a member of a new viral family) is sequenced, functions can be annotated for a very few genes, and often far more than 50% of the putative genes can only be described as "hypothetical protein" (Pedulla *et al.,* 2003; Mochizuki *et al.,* 2010; Pope *et al.,* 2015).

The vast majority of unknown viral sequence is assumed to be of viral origin, but the lack of corresponding reference sequences in the public database does not allow definitive determination. These observations have led a number of researchers to make the analogy between these putative viral sequences and the presence of dark matter or dark energy in the Universe, by calling these sequences "biological dark matter" (Pedulla *et al.,* 2003) or "the dark matter of the biosphere" (Filée *et al.,* 2005; Hatfull, 2015).

To date, all isolated viruses of the Archaea have DNA genomes often containing more than 80% of putative genes lacking homologs in the sequence databases (Snyder *et al.,* 2015; Prangishvili *et al.,* 2016). For a long time, archaeal viruses were only found to contain dsDNA genomes, and it was speculated that, owing to the extreme physicochemical conditions under which many archaea thrive, the only possible genome type is dsDNA. However, this view has been overturned with ssDNA archaeal viruses found in halophilic archaea (Pleolipoviridae) (Pietilä *et al.,* 2009), and in hyperthermophilic archaea (Spiraviridae) (Mochizuki *et al.,* 2012). The hyperthermophilic spiravirus *Aeropyrum* coil-shaped virus (ACV) has the largest known genome of all ssDNA viruses, with 24 thousand bases, which is more than double the size of the previously largest known ssDNA genome. This is particularly surprising given that ACV was isolated from a host organism that grows at 90°C (Mochizuki *et al.,* 2012).

2.9. Virus Metagenomes

The recent development of high-throughput DNA sequencing has accelerated studies on genomic diversity of environmental samples in a culture-independent manner, known as metagenomics. However, unlike for cellular samples, where 95% of sequences are similar to previously known sequences, metagenomic analysis targeting virus-sized particles (metaviromes) in many natural environments generates 60–90% of sequences that do not have clear homologs (Edwards and Rohwer, 2005; Rosario and Breitbart, 2011; Hurwitz and Sullivan, 2013; Yoshida *et al.,* 2013; Simmonds *et al.,* 2017). For a number of reasons, marine viromes have been among the most intensively

studied. Early metavirome studies first indicated an over-whelming diversity of virus sequences with possibly billions of virus species (Rohwer, 2003). More recent studies indicate that there may only be up to a million virus species (Brum *et al.,* 2015). Nonetheless, there are still many new groups of viruses that are known only by their sequence (Breitbart, 2012; Brum *et al.,* 2015; Roux *et al.,* 2016). A recent analysis of all published viral DNA metagenomes found that nearly 70% of virus-associated genes were novel (Paez-Espino *et al.,* 2016). Thus there is still a great deal to discover in our virus world.

2.10. CRISPR—A Virus Defense Mechanism and Historical Record

Most archaeal and some bacterial cellular genomes contain so-called CRISPRs (Clustered Regularly Interspersed Palindromic Repeats), a previously cryptic sequence structure that is now known to provide acquired immunity to virus infection (reviewed in Koonin, 2017). CRISPR sequences consist of repeated DNA sequences with short so-called spacer sequences between the repeats that are identical to parts of virus genome sequences or other mobile DNA (Koonin, 2017). One of these systems, the CRISPR/Cas9 system, has been widely used for genome engineering (Wright *et al.,* 2016). However, more importantly for astrovirology, CRISPR sequences also contain a record of past virus infections. This has proved extremely useful for determining past infections (Wilmes *et al.,* 2009), reconstructing extinct viruses (Andersson and Banfield, 2008), inferring hosts for unknown viruses (Anderson *et al.,* 2011), and virus biogeography (Held and Whitaker, 2009).

2.11. Viruses Change Host Metabolism, Sometimes to the Host's Advantage

In many known cases virus infection is deleterious to the host cell or the organism. However, more recently, it has been realized that some, if not most, viruses coexist symbiotically with their hosts (reviewed in Roossinck, 2011; Metcalf and Bordenstein, 2012; Roossinck, 2015). Interactions between viruses and their hosts can range from parasitism to mutualism (*e.g.,* Hamelin *et al.,* 2017). There are a growing number of examples of virus infections that clearly benefit their hosts, mostly by contributing genes that allow their hosts to survive under conditions under which they would otherwise perish (reviewed in Roossinck, 2011; Roossinck, 2016; Roux *et al.,* 2016).

Upon viral infection, virus-encoded genes are expressed. Most of these genes are specific for virus replication. However, in many cases, viral genomes also encode genes whose products change their host's metabolism. The best-characterized of these genes are components of photosystem II, the machinery for oxygenic photosynthesis. These genes are found in viruses that infect photosynthetic cyanobacteria and not only allow the infected cyanobacterium to undergo photosynthesis under conditions that uninfected bacteria could not but also enable them to thrive in parts of the ocean that uninfected cells cannot (Mann *et al.,* 2003; Lindell *et al.,* 2005; Sullivan *et al.,* 2006). Moreover, other so-called auxiliary metabolic genes (AMG), important for nucleotide, phosphate, carbon, sulfur, nitrogen, and ammonium metabolism, have been reported in virus genomes and metagenomes (Breitbart, 2012; Anantharaman

et al., 2014; Roux *et al.,* 2016). It would be very interesting to determine if these virus-encoded gene products lead to isotopic fractionation that is different from cellular enzymes, which could lead to the development of virus-specific isotopic biomarkers. It is probable that extraterrestrial viruses—if they exist—will also coexist with host cells on a spectrum between mutualism and parasitism.

3.1. Viruses Are Critical for Biogeochemical Cycles

Due to their extremely large numbers, in addition to allowing hosts to adapt to new environments and changing their metabolism, viruses also contribute to global nutrient cycling. Viruses are assumed to cause 50% of bacterial death in the modern oceans, with a calculated 10^{28} virus infections/day (Suttle, 2007; Breitbart, 2012). Each time a viral infection causes cell lysis, cellular organic matter is released to the surrounding environment. These burst cells provide nutrients for other organisms and maintain the food chain, via the so-called viral shunt (Wommack and Colwell, 2000; Weinbauer, 2004). Up to 25% of fixed carbon is predicted to be recycled in the upper ocean water column due to the activity of these marine viruses (Wilhelm and Suttle, 1999). If there were no virus-mediated cell lysis, cells would gradually fall to the seafloor, thereby not only removing nutrients from the upper ocean but also lowering the carbon content of the oceans. This process of carbon recycling is an essential component of global climate regulation (Suttle, 2007). Viral lysis of abundant coccolithophore algae may influence climate by release of cloud-nucleating chemicals (Wilson *et al.,* 2009). Moreover, about 5% of the oxygen on Earth is estimated to be produced by virus-infected cells using virally encoded photosynthesis genes (Suttle, 2007). It is intriguing to speculate on the role of viruses in ''the great oxygenation event'' and the amount of oxygen and organic matter in the ancient oceans (Holland, 2006). To our knowledge viruses have not been incorporated in such models. Microbial ecologists are beginning to include viruses in terran biogeochemical cycle models (Stec *et al.,* 2017), and we suggest that modelers of exoplanets and their satellites also consider including viruses in their models.

3.2. Viruses Drive Evolution

It has been said that evolution and possibly life as we know it would not be possible without viruses (Greene and Reid, 2014). There are two major ways that viruses drive evolution: first, by providing strong selective pressure for resistance to viral infection; and second, by moving exogenous genetic materials into cells. Evolution of host resistance to viral infection, particularly against a virus that kills its host, provides exceptionally strong evolutionary pressure. On the other hand, since viruses require hosts for their replication, there is also strong evolutionary pressure on the virus to overcome host defenses. This interplay of host and viral evolution has been demonstrated both experimentally and theoretically (Turner and Chao, 1999; Wichman and Brown, 2010; Elena, 2016). Viruses and their hosts are thus often said to be in an ''arms race'' or ''running to stand still,'' fulfilling the ''Red Queen hypothesis'' of evolution (Van Valen, 1973). In simulations of early life, parasitic replicators, similar in concept to what we now know as viruses, inevitably emerge (Koonin, 2011).

Less appreciated but probably no less important for evolution is that viruses drive evolution by the introduction of new genes to cells and by transferring genes from one cell to another. This acquisition of genes from a source other than a direct ancestor is termed horizontal gene transfer (HGT). Viruses are known to be major agents, if not the most important agent, of HGT. Horizontally transferred genes are often acquired by the host genome, particularly if they confer the recipient with an evolutionary advantage. Since genetic diversity in the viral gene pool is very high and viruses are agents of HGT, the successful transfer of novel viral genes into host genomes is likely commonplace. Virus-mediated HGT is particularly important for evolution in species that reproduce asexually but is also widespread in both vertebrates and invertebrate animals (Koonin, 2016). For example, the human genome contains between 8% and 40% virus-derived sequences, some of which have been beneficial for human development (Lander *et al.,* 2001; Feschotte and Gilbert, 2012). One well-studied example is the syncytin gene that is essential for mammalian placental development. Syncytin is derived from not one but multiple apparently independent ancient retrovirus infections (Lavialle *et al.,* 2013).

Some viruses insert their genomes into their host's genomes (Fig. 1A). The best-studied of these are the retroviruses that require insertion of their genome into their host genome as part of their replication cycle. This integration step generates a copy of the virus genome within the host genome, in the form termed a provirus. Under certain conditions, proviruses can be triggered to replicate and produce new virions. As their proviruses replicate, retroviruses can acquire cellular genes. In fact this is how the first oncogenes (cellular genes that cause cancer when misregulated) were discovered (Stehelin *et al.,* 1976). Many bacterial and archaeal DNA viruses also integrate their genomes into the host genome, generating proviruses (or prophages, in the case of bacterial viruses). Upon replication, these prophages can also incorporate host genes and transfer them to other cells in a process called transduction. Some viruses and defective viruses, known as generalized transducing bacteriophage and gene transfer agents (GTAs) respectively, incorporate random DNA into their virions and inject these DNAs into host cells, leading to massive amounts of HGT. Generalized transducing bacteriophage have been widely used in bacterial genetics (Bachmann, 1990) and seem to be important for the development of both bacterial pathogenesis and antibiotic resistance (Moon *et al.,* 2015). Even viruses that use ssDNA or RNA as their genetic material are known to be incorporated, at least partially, into cellular dsDNA genomes (Stedman, 2015). Due to transduction and promiscuous provirus gene acquisition, the amount of horizontal gene flow via viral infection is currently considered much higher than once expected (McDaniel *et al.,* 2010) and is hypothesized to have been even higher in primordial environments (Woese, 2002; Koonin *et al.,* 2006). It is thus highly likely that viruses played a major role in the evolution of early life.

3.3. Viruses as Drivers of Major Evolutionary Transitions

Viruses have been hypothesized to be responsible for the evolution of DNA as the genetic material, the three domains of cellular life, the development of the eukaryotic nucleus, and even multicellularity. One of the first proposed virus-driven evolutionary events is that viruses may have "invented" DNA genomes in an ancestral RNA-protein world, and that DNA was adopted later by cellular organisms (Villarreal and DeFilippis, 2000; Forterre, 2002, 2005). In order to convert RNA into DNA, three steps are required: a ribonucleotide reductase that transforms ribose to deoxyribose, a thymidylate kinase that adds a phosphate group to thymidine, and a reverse-transcriptase that makes a DNA copy from an RNA template. Interestingly, all of these genes have been found in virus genomes, and these viral genes have much greater sequence diversity than the corresponding cellular genes. The most parsimonious explanation for this diversity is that these genes first arose in viruses, evolved therein, and some genes were transmitted to, and retained by, cellular genomes. Moreover, phylogenetic analysis of DNA polymerase genes often places viral genes at the base of the eukaryotic clade, indicating that the viral polymerases are ancestral (Filée *et al.,* 2002; Koonin *et al.,* 2006; Villarreal and DeFilippis, 2000).

An extended version of this hypothesis is that different virus infections led to the formation of the modern three cellular domains of life (Forterre, 2006a). The argument is that DNA polymerases in the three cellular domains are very different from each other, but each is clearly related to extant viruses, which also have orthogonal DNA replication systems (Woese, 2002). Engulfment of a large DNA virus has also been proposed as the origin of the eukaryotic nucleus (Bell, 2001; Takemura, 2001). The recent discovery of the giant Mimiviruses has caused these theories to be revisited (Bell, 2009). The intracellular compartment formed in the Mimivirus' host amoeba, termed the virus factory (or virion factory), closely resembles the eukaryotic nucleus in both structure and function (Raoult and Forterre, 2008). Thus an ancestral virus infection that caused the formation of a virus factory may have led to the development of the nucleus. A modern infection of a bacteria with a bacterial virus leads to the development of a remarkably nucleus-like structure in the infected bacterium (Chaikeeratisak *et al.,* 2017).

The remarkable morphological and sequence diversity observed in archaeal viruses provides another clue regarding cellular (and virus) evolution. A naive assumption would be that the archaeal virosphere would have limited diversity, due to the harsh living conditions of most archaeal virus hosts. But the complete opposite, high diversity observed in archaeal viruses and limited diversity in bacterial viruses, is indeed puzzling (Rachel *et al.,* 2002: Ackermann and Prangishvili, 2012: Prangishvili *et al.,* 2016). One recently proposed hypothesis incorporates ancient viruses as a direct evolutionary driving force, which eventually led to the divergence of the Bacteria and Archaea (Forterre and Prangishvili, 2009: Prangishvili, 2013). They suggest that ancient viruses were already rich in morphological diversity at the time of the Last Universal Common Ancestor (LUCA). Later, one group of cellular organisms, the ancestors of bacteria, developed a cell wall to protect themselves from viral infections. In order to overcome this rigid defense system, bacterial viruses evolved a physical penetration mechanism to inject their genetic material, the tail, which led to the predominance of head-tail viruses in the bacterial domain. On the other hand, other primordial cells, the ancestors of the Archaea, altered their cell surface proteins by

glycosylation to escape from viral recognition. Their viruses then glycosylated their capsid proteins in order to attach to the hosts. This explains the high level of glycosylation on archaeal cell surface proteins and archaeal virus capsid proteins; such glycosylation is not frequently observed in bacterial cells and viruses.

Finally, viruses may have been involved in the evolution of multicellularity. In the presence of excess resources, multi-cellularity could provide protection from virus infection, as cells on the outside of an aggregate would be more suscep-tible to virus infection than those inside (Ruardij *et al.,* 2005; Greene and Reid, 2014). Moreover, the ''invention'' of apo-ptosis or programmed cell death, which is directly coupled to the evolution of multicellularity, could be due to virus in-fection pressures (Iranzo *et al.,* 2014).

3.4. Diversity of Viruses and Their Evolutionary Relationships

Due to the lack of ''universal'' genes, such as 16/18S rRNA genes, in viruses, it is nearly impossible to place all viruses into one phylogeny to represent their evolutionary relationships. Nucleotide polymerases are often (but not nec-essarily always) encoded in virus genomes, and they have been used frequently to reveal the history of some distantly related viruses (Koonin, 1991; Koonin *et al.,* 2008; Koonin and Dolja, 2014). But the polymerases themselves are highly diverse, both in terms of sequence and substrates used. Most viruses encode a nucleotide polymerase, but some, including most hyperthermophilic archaeal viruses, do not encode a recognizable polymerase gene in their genomes.

However, more recently, comparison of the three-dimensional structure of major capsid proteins indicates that several viruses that infect archaea, bacteria, and eukarya have a common double beta-barrel capsid protein structure and are thus extremely distantly related, even though se-quence similarity cannot be detected (Rice *et al.,* 2004; Khayat *et al.,* 2005; Krupovič and Bamford, 2008). Similar results were obtained with a different capsid protein struc-ture found in a haloarchaeal virus, bacterial viruses, and herpesviruses (Pietilä *et al.,* 2013). Thus classification of different viral groups based on capsid protein structure seems to be an effective way to discern phylogenetic rela-tionships between distantly related viruses infecting viruses of all three cellular domains (Krupovič and Bamford, 2011; Abrescia *et al.,* 2012). Major capsid protein structures could potentially become the viral equivalent to the cellular 16S/ 18S rRNA gene, especially since all viruses encode capsid proteins. However, identification of the capsid protein gene is not straightforward, particularly when analyzing new viral groups from metagenomic analysis (Mokili *et al.,* 2012) and determining capsid protein structures is highly problematic in astrobiological samples.

However, even if capsid protein structures could provide a comprehensive virus phylogeny, there appears to be ram-pant recombination in virus genomes, such that many of the genes in a virus genome do not have the same phylogeny. An extreme example is the chimeric viruses that appear to have undergone recombination between RNA (Baltimore Class IV) and DNA (Baltimore Class II) viruses (Diemer and Stedman, 2012; Roux *et al.,* 2013). A recent metagenomic survey of invertebrate animals discovered hundreds of new

RNA viruses, indicated that recombination between different viral genes is rampant, and confounds virus phylogeny (Shi *et al.,* 2016). Thus it may be impossible to represent virus phylogeny by a tree-like structure, as viruses probably have a more mosaic or modular genome evolution dominated by HGT (Botstein, 1980; Hendrix *et al.,* 1999; Koonin and Dolja, 2014; Koonin *et al.,* 2015; Iranzo *et al.,* 2016).

3.5. What Is the Origin of Viruses?

The evolution of cellular and viral life may not easily be separable (Szathmary and Smith, 1997; Villarreal, 2004; Koonin *et al.,* 2006; Jalasvuori and Bamford, 2008; Forterre, 2010; Koonin and Dolja, 2013). There are three predomi-nant theories for virus origins: First is that viruses are es-caped cellular genes. Second is that viruses are degenerated cells. Both of these theories indicate a relatively recent origin for viruses. The third theory is that viruses could have an independent origin or multiple independent origins (Koonin *et al.,* 2006). Increasingly, virus researchers think that the latter is true (Forterre, 2006b; Koonin *et al.,* 2006; Witzany, 2008; Forterre and Prangishvili, 2009; Forterre and Krupovič, 2012; Nasir and Caetano-Anollés, 2015). There are also a few researchers who controversially claim that viruses are remnants of an otherwise extinct fourth domain of cellular life (Philippe *et al.,* 2013), but see the works of Schulz *et al.* (2017) and Moreira and López-García (2015). Nevertheless, if viruses were involved in the origin of DNA, the transition between an RNA and DNA world, the origin of the three domains of life, the development of the bacterial cell wall, or the origin of the eukaryotic nucleus, they must be very ancient.

One piece of evidence is that viruses are the only extant terran life-forms that use RNA genomes, and could thus be descendants of viruses or similar replicative entities that ex-isted in the hypothetical RNA world (Sankaran, 2016). An active area of origins-of-life research is the encapsidation of self-replicating RNAs in vesicles, not unlike enveloped RNA viruses (Robertson and Joyce, 2014; Deamer, 2017). There are also many genes that are found in virus genomes but not in any cellular genomes, further supporting the independent origin hypothesis (Koonin *et al.,* 2006; Krupovič and Bam-ford, 2010). One of the great mysteries of the origin of viruses is the origin of the defining characteristic of viruses, the vi-rion. Recent analyses indicate that there may have been multiple origins of capsids and thus multiple virus origins (Krupovič and Koonin, 2017). Unfortunately, there is no di-rect evidence for ancient viruses in the geological record, but there is considerable indirect evidence for virus antiquity.

3.6. What Is the Evidence for Ancient Viruses?

Indirect evidence for ancient viruses comes from genomic and structural studies. The presence of common viral genes in genomes of organisms known to have diverged tens of millions of years ago from both fossil and genomic data, together with common antiviral genes, indicates that specific viruses and viral defenses are at least tens to hundreds of millions of years old (Emerman and Malik, 2010; Aiewsa-kun and Katzourakis, 2017). Unfortunately, sequence simi-larity becomes undetectable for more ancient divergences. The presence of conserved capsid protein structures is

indirect evidence for ancient viruses' lineages (see Section 3.4 above). Moreover, when virus capsids are classified based on structures of their capsid proteins, there are a limited number of clades. The most parsimonious explanation for these data is that these virion protein structures predated the divergence of the three domains of cellular life (Rice *et al.,* 2004; Krupovič and Bamford, 2011). However, there may have been multiple, independent acquisitions of virus capsid proteins throughout both viral and cellular evolution (Krupovič and Koonin, 2017).

There is almost no evidence for viruses in the rock record. Structures resembling virus inclusions formed by current insect viruses have been detected in *ca.* 100-million-year-old amber (Poinar and Poinar, 2005). Laboratory silicification of the well-studied bacteriophage T4 and extremophile viruses led to the determination that the spatial resolution of current techniques is insufficient to determine whether organic inclusions in rock are viral in origin (Laidler and Stedman, 2010; Orange *et al.,* 2011). In both analyses, detectable virion morphology was lost over periods of days to months (Laidler and Stedman, 2010; Orange *et al.,* 2011).

3.7. Could Viruses Be Spread Extraterrestrially?

At present there is no evidence for virions in any extraterrestrial samples, mainly due to lack of biomarkers. However, considering the high abundance in the modern biosphere and due to the small particle size and mass of most virions, it is likely that meteoritic bombardment or possibly volcanic activity would have ejected viral particles from Earth. A few studies have examined virus stability in the space environment. Tobacco mosaic virus, a plant virus that is highly resistant to desiccation and can be crystallized, was exposed to simulated interstellar radiation and low temperature (77 K) high vacuum (10^{-8} torr) with minimal infectivity loss (Koike *et al.,* 1992). Poliovirus and bacteriophage T1 were exposed to the space environment on high-altitude balloons and rockets for short times without significant loss of activity (Parvenov and Lukin, 1973). Dried bacteriophage T7 was used as a biological dosimeter for UV irradiation on the International Space Station (ISS), but infectivity was not tested thereafter (Bérces *et al.,* 2015). A survey of virion stability studies was published (Griffin, 2013), but none of these studies addresses possible survival of viruses to interplanetary transit. It is theoretically possible that endolithic viruses or viruses encased in salt crystals could survive space exposure, but for how long is not clear. We found that silica-coated bacteriophage T4 could survive a month of drying, but infectivity was completely lost thereafter (Laidler *et al.,* 2013). Samples have just been returned from the Tanpopo mission on the ISS (Kawaguchi *et al.,* 2016). The Tanpopo mission collected particles present in low Earth orbit on the ISS in aerogel for over a year and also exposed terrestrial bacteria and yeast to the space environment for up to 2 years. Whether any micrometeoritic samples collected in the Tanpopo aerogel contain virions would be very interesting but unlikely to be unambiguous due to the lack of validated virus biosignatures.

3.8. Virus Biosignature Detection

What might virus biosignatures be? The 2015 NASA Astrobiology Strategy describes the breakdown of current techniques and strategies for life detection into remote and direct detection of extinct and existing life. Although viruses have been shown to impact life at an ecosystem level, remote detection of individual virions is implausible; however, macroscopic effects of virus infection could be detectable, given changes to Earth's biogeochemical cycles (see Section 3.1). Direct detection of viruses is more realistic given the average size of a virion.

The standard detection strategies for extinct life, including chemical signatures and isotopic fractionation, also serve as detection strategies for extinct viruses. While the standard array of chemical signatures for detecting extinct life includes hopanoids, sterols, cyclic alkanes, isoprenoids, and carotenoids, specific virus chemical signatures are still under investigation. Enveloped viruses containing lipids may provide a detectable chemical signature (Kyle *et al.,* 2012). Unlike most host-derived viral components, some viral lipids exhibit a distinct chemical signature from the lipids of host organisms; such a distinction may prove critical in the chemical detection of viruses in the rock record. Moreover, some viruses also have what was thought to be virus-specific nucleotide modifications, for example, 5-hydroxy-methyl cytosine in bacteriophage T4. However, this modified nucleotide has recently been found in human DNA (Bachman *et al.,* 2014).

Common techniques for biosignature detection, such as extraction of rocks with organic solvents followed by liquid chromatography and demineralization, can be adapted for use in virus detection by simple manipulations of equipment and technique. Laboratory methods to identify the isotopic abundances (high-resolution mass spectrometry), chemical species (GC-MS, NMR, or FTIR spectroscopy), and elemental composition (inductively coupled argon plasma mass spectrometry) all can be used with little to no modification to identify and characterize extant virus samples along with samples of potential life. However, differentiating viruses from potential cellular material in bulk samples is problematic. More promising is detection of a virus-mediated event, such as massive lysis or death of host organisms. Such an event has been proposed to have led to the famous chalk deposits of the white cliffs of Dover. These deposits appear to be due to massive lysis of coccolithophore algae, possibly due to virus infection (Wilson *et al.,* 2009).

Because virions do not metabolize directly, no by-products such as methane or N_2O should be detectable remotely; however, if host metabolism is changed by virus genes, this could be detectable remotely. Similarly, disequilibria such as redox, chemical, and mineralogical chemical signatures should not be used as direct detection mechanisms for viruses but could be useful as indirect measurements. It is not known to our knowledge if virus-specific AMG fractionate isotopes differently than cellular enzymes, but this work should be of high priority.

Standard detection strategies for existing life, including biomolecules and chiral excess, serve as detection strategies for existing viruses. As opposed to cellular life, fully assembled virions are composed of three of the four standard cellular macromolecules including nucleotides (DNA or RNA), amino acids, proteins, (viral capsids), and in some cases lipids (viral envelopes). Virions are smaller than their host; thus the proportional relationship between components in detectable biomolecules of viruses could provide a virus biosignature if high-enough (<10 nm) spatial resolution

could be obtained. Many viruses produce complex capsid geometries from the arrangement of repeated proteins; in some cases, a single protein is sufficient for assembly of a complete capsid (Acheson, 2011). Detection of repeated proteins could provide insight into the sample identity. Sequence identification of the amino acids or nucleotides of a virion provide immediate insight into identity of the virus. In addition to the common techniques applied to extinct viruses described above, techniques such as melting and filtration of ice followed by ultrasensitive analysis using various chemical techniques can be applied for the detection and characterization of known existing viruses. These include, but are not limited to, PCR and phylogenetic analyses, immunoassays, lab-on-a-chip methods using capillary electrophoresis and laser-induced fluorescence, and ultrasensitive mass spectrometry. These modern techniques, such as high throughput sequencing or high density microarrays (ViroChip), can identify known viruses (*e.g.*, Yozwiak *et al.*, 2012), and some bioinformatics techniques can identify unknown but prevalent viruses (Roux *et al.*, 2016). However, the identification of truly novel viruses, such as potential extraterrestrial viruses, remains problematic. Nonetheless, we find that the development of virus-specific biomarkers should be an active area of astrovirology research and is highly relevant to the detection of viruses in extraterrestrial material.

4.1. Future Outlook

More than a century has passed since the discovery of the first viruses. Entering the second century of virology, we can finally start focusing beyond our own planet. Priorities for future astrovirology research in the short term should include (1) validation of virus biosignatures, (2) consideration of virus-detection experiments to be used for missions that sample water plumes from Enceladus and Europa, (3) inclusion of viruses in models for ancient oceans and extraterrestrial systems, (4) determination whether virus-encoded and transferred AMG fractionate isotopes differently than their cellular counterparts.

In the longer term, astrovirology objectives should include (1) more exposure studies of viruses, particularly endolithic viruses, to the space environment; (2) characterization of highly abundant cosmopolitan viruses (Roux *et al.*, 2016; Urayama *et al.*, 2016); (3) greater research into the roles of viruses in the origin and evolution of early life; and (4) more outreach to astrobiologists and the general public regarding the ubiquity and role of viruses in Earth's ecosystems.

Viruses are an integral, highly abundant yet underappreciated part of life on Earth. They play a critical role in biogeochemical cycles and evolution today and may have been intimately involved in the origin of life. We do need to learn much more about viruses on modern Earth before we look elsewhere, but let's start looking.

Acknowledgments

Author Disclosure Statement

Author K.S. is founder and Chief Scientific Officer of StoneStable, Inc., and owns StoneStable, Inc., stock. Authors A.B. and T.M. declare no competing financial interests exist.

References

Abrescia, N.G.A., Bamford, D.H., Grimes, J.M., and Stuart, D.I. (2012) Structure unifies the viral universe. *Annu Rev Biochem* 81:795–822.

Acheson, N.H. (2011) *Fundamentals of Molecular Virology,* 2nd ed., Wiley, New York.

Ackermann, H.W. and Prangishvili, D. (2012) Prokaryote viruses studied by electron microscopy. *Arch Virol* 157:1843–1849.

Aherfi, S., Colson, P., La Scola, B., and Raoult, D. (2016) Giant viruses of amoebas: an update. *Front Microbiol* 7, doi:10.3389/fmicb.2016.00349.

Aiewsakun, P. and Katzourakis, A. (2017) Marine origin of retroviruses in the early Palaeozoic Era. *Nat Commun* 8, doi:10.1038/ncomms13954.

Anantharaman, K., Duhaime, M.B., Breier, J.A., Wendt, K.A., Toner, B.M., and Dick, G.J. (2014) Sulfur oxidation genes in diverse deep-sea viruses. *Science* 344:757–760.

Anderson, R.E., Brazelton, W.J., and Baross, J.A. (2011) Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol Ecol* 77:120–133.

Andersson, A.F. and Banfield, J.F. (2008) Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320:1047–1050.

Arnold, H.P., Ziese, U., and Zillig, W. (2000a) SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon *Sulfolobus. Virology* 272:409–416.

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., Weidmann, J.F., Kristjanson, J.K., Klenk, H.P., and Nelson, K.E. (2000b) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus. Virology* 267:252–266.

Atanasova, N.S., Roine, E., Oren, A., Bamford, D.H., and Oksanen, H.M. (2012) Global network of specific virus-host interactions in hypersaline environments. *Environ Microbiol* 14:426–440.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014) 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem* 6:1049–1055.

Bachmann, B.J. (1990) Linkage map of *Escherichia coli* K-12, edition 8. *Microbiol Rev* 54:130–197.

Baltimore, D. (1971) Expression of animal virus genomes. *Bacteriol Rev* 35:235–241.

Bamford, D.H., Ravantti, J.J., Rönnholm, G., Laurinavičius, S., Kukkaro, P., Dyall-Smith, M., Somerharju, P., Kalkkinen, N., and Bamford, J.K.H. (2005) Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica. J Virol* 79:9097–9107.

Bath, C. and Dyall-Smith, M.L. (1998) His1, an archaeal virus of the *Fuselloviridae* family that infects *Haloarcula hispanica*. *J Virol* 72:9392–9395.

Beijerinck, M.W. (1898) Über ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblätter. *Verhandelingen der Koninklijke akademie van Wetenschappen te Amsterdam* 65:1–22.

Bell, P.J.L. (2001) Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? *J Mol Evol* 53:251–256.

Bell, P.J.L. (2009) The viral eukaryogenesis hypothesis: a key role for viruses in the emergence of eukaryotes from a prokaryotic world environment. *Ann NY Acad Sci*, 1178:91–105.

Bérces, A., Egyeki, M., Fekete, A., Horneck, G., Kovács, G., Panitz, C., and Rontó, G. (2015) The PUR experiment on the EXPOSE-R facility: biological dosimetry of solar extraterrestrial UV radiation. *International Journal of Astrobiology* 14:47–53.

Bergh, O., Børsheim, K.Y., Bratbak, G., and Heldal, M. (1989) High abundance of viruses found in aquatic environments. *Nature* 340:467–468.

Bize, A., Peng, X., Prokofeva, M., MacLellan, K., Lucas, S., Forterre, P., Garrett, R.A., Bonch-Osmolovskaya, E.A., and Prangishvili, D. (2008) Viruses in acidic geothermal environments of the Kamchatka Peninsula. *Res Microbiol* 159:358–366.

Blumberg, B.S. (2004) Astrovirology. In *Abstracts from the Astrobiology Science Conference 2004:* Special Supplement. *International Journal of Astrobiology* 3:26–27.

Botstein, D. (1980) A theory of modular evolution for bacteriophages. *Ann NY Acad Sci* 354:484–490.

Breitbart, M. (2012) Marine viruses: truth or dare. *Ann Rev Mar Sci* 4:425–438.

Brum, J.R., Schenck, R.O., and Sullivan, M.B. (2013) Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. *ISME J* 7:1738–1751.

Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulcier, G., Acinas, S.G., Alberti, A., Chaffron, S., Cruaud, C., de Vargas, C., Gasol, J.M., Gorsky, G., Gregory, A.C., Guidi, L., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Poulos, B.T., Schwenck, S.M., Speich, S., Dimier, C., Kandels-Lewis, S., Picheral, M., Searson, S.; Tara Oceans Coordinators, Bork, P., Bowler, C., Sunagawa, S., Wincker, P., Karsenti, E., and Sullivan, M.B. (2015) Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498.

Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A.F., Erb, M.L., Coker, J.K.C., Vavilina, A., Newton, G.L., Buschauer, R., and Pogliano, K. (2017) Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355:194–197.

Claverie, J.-M. (2006) Viruses take center stage in cellular evolution. *Genome Biol* 7:110.

Crick, F.H.C. and Watson, J.D. (1956) Structure of small viruses. *Nature* 177:473–475.

Deamer, D. (2017) The role of lipid membranes in life's origin. *Life* 7, doi:10.3390/life7010005.

Diemer, G.S. and Stedman, K.M. (2012) A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol Direct* 7, doi:10.1186/1745-6150-7-13.

Edmunson, J., Gaskin, J.A., Jerman, G.A., Harvey, R.P., Doloboff, I.J., and Neidholdt, E.L. (2016) A miniaturized variable pressure scanning electron microscope (MVP-SEM) for in-situ Mars surface sample analysis [abstract 2301]. In *47th Lunar and Planetary Science Conference Abstracts,* Lunar and Planetary Institute, Houston.

Edwards, R.A. and Rohwer, F. (2005) Viral metagenomics. *Nat Rev Microbiol* 3:504–510.

Elena, S.F. (2016) Evolutionary transitions during RNA virus experimental evolution. *Philos Trans R Soc Lond B Biol Sci* 371, doi:10.1098/rstb.2015.0441.

Emerman, M. and Malik, H.S. (2010) Paleovirology—modern consequences of ancient viruses. *PLoS Biol* 8, doi:10.1371/journal.pbio.1000301.

Enard, D., Cai, L., Gwennap, C., and Petrov, D.A. (2016) Viruses are a dominant driver of protein adaptation in mammals. *eLife* 5, doi:10.7554/eLife.12469.

Feschotte, C. and Gilbert, C. (2012) Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* 13:283–296.

Filée, J., Forterre, P., Sen-Lin, T., and Laurent, J. (2002) Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *J Mol Evol* 54:763–773.

Filée, J., Tétart, F., Suttle, C.A., and Krisch, H.M. (2005) Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proc Natl Acad Sci USA* 102:12471–12476.

Fischer, M.G. and Hackl, T. (2016) Host genome integration and giant virus-induced reactivation of the virophage Mavirus. *Nature* 540:288–291.

Fischer, M.G., Allen, M.J., Wilson, W.H., and Suttle, C.A. (2010) Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc Natl Acad Sci USA* 107:19508–19513.

Forterre, P. (2002) The origin of DNA genomes and DNA replication proteins. *Curr Opin Microbiol* 5:525–532.

Forterre, P. (2005) The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 87:793–803.

Forterre, P. (2006a) Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc Natl Acad Sci USA* 103:3669–3674.

Forterre, P. (2006b) The origins of viruses and their possible roles in major evolutionary transitions. *Virus Res* 117:5–16.

Forterre, P. (2010) Defining life: the virus viewpoint. *Orig Life Evol Biosph* 40:151–160.

Forterre, P. and Krupovič, M. (2012) The origin of virions and virocells: the escape hypothesis revisited. In *Viruses: Essential Agents of Life,* edited by G. Witzany, Springer, Dordrecht, The Netherlands, pp 43–60.

Forterre, P. and Prangishvili, D. (2009) The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann NY Acad Sci* 1178:65–77.

Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, ''*Pyrococcus abyssi.*'' *J Bacteriol* 185:3888–3894.

Gorlas, A., Koonin, E.V., Bienvenu, N., Prieur, D., and Geslin, C. (2012) TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ Microbiol* 14:503–516.

Greene, S.E. and Reid, A. (2014) *Viruses Throughout Life & Time: Friends, Foes, Change Agents,* a report from the American Academy of Microbiology, American Society for Microbiology, Washington, DC.

Griffin, D.W. (2013) The quest for extraterrestrial life: what about the viruses? *Astrobiology* 13:774–783.

Hamelin, F.M., Hilker, F.M., Sun, T.A., Jeger, M.J., Hajimorad, M.R., Allen, L.J.S., and Prendeville, H.R. (2017) The evolution of parasitic and mutualistic plant-virus symbioses through transmission-virulence trade-offs. *Virus Res* 241, doi: 10.1016/j.virusres.2017.04.011.

Hansen, C.J., Esposito, L.W., Stewart, A.I.F., Meinke, B., Wallis, B., Colwell, J.E., Hendrix, A.R., Larsen, K., Pryor, W., and Tian, F. (2008) Water vapour jets inside the plume of gas leaving Enceladus. *Nature* 456:477–479.

Häring, M., Peng, X., Brügger, K., Rachel, R., Stetter, K.O., Garrett, R.A., and Prangishvili, D. (2004) Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus:* a novel virus family, the Globuloviridae. *Virology* 323:233–242.

Häring, M., Rachel, R., Peng, X., Garrett, R.A., and Prangishvili, D. (2005a) Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the Ampullaviridae. *J Virol* 79:9904–9911.

Häring, M., Vestergaard, G., Rachel, R., Chen, L.M., Garrett, R.A., and Prangishvili, D. (2005b) Virology: independent virus development outside a host. *Nature* 436:1101–1102.

Hatfull, G.F. (2015) Dark matter of the biosphere: the amazing world of bacteriophage diversity. *J Virol* 89:8107–8110.

Hays, L., editor in chief. (2015) *2015 Astrobiology Strategy,* NASA, Washington, DC.

Held, N.L. and Whitaker, R.J. (2009) Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ Microbiol* 11:457–466.

Hendrix, R.W., Smith, M.C.M., Burns, R.N., Ford, M.E., and Hatfull, G.F. (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci USA* 96:2192–2197.

Holland, H.D. (2006) The oxygenation of the atmosphere and oceans. *Philos Trans R Soc Lond B Biol Sci* 361:903–915.

Horneck, G., Walter, N., Westall, F., Grenfell, J.L., Martin, W.F., Gomez, F., Leuko, S., Lee, N., Onofri, S., Tsiganis, K., Saladino, R., Pilat-Lohinger, E., Palomba, E., Harrison, J., Rull, F., Muller, C., Strazzulla, G., Brucato, J.R., Rettberg, P., and Capria, M.T. (2016) AstRoMap European Astrobiology Roadmap. *Astrobiology* 16:201–243.

Hurwitz, B.L. and Sullivan, M.B. (2013) The Pacific Ocean Virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS One* 8, doi:10.1371/journal.pone.0057355.

Iranzo, J., Lobkovsky, A.E., Wolf, Y.I., and Koonin, E.V. (2014) Virus-host arms race at the joint origin of multicellularity and programmed cell death. *Cell Cycle* 13:3083–3088.

Iranzo, J., Krupovič, M., and Koonin, E.V. (2016) The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. *mBio* 7, doi:10.1128/mBio.00978-16.

Jalasvuori, M. and Bamford, J.K.H. (2008) Structural co-evolution of viruses and cells in the primordial world. *Orig Life Evol Biosph* 38:165–181.

Janekovic, D., Wunderl, S., Holz, I., Zillig, W., Gierl, A., and Neumann, H. (1983) TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaebacterium *Thermoproteus tenax. Mol Gen Genet* 192:39–45.

Kawaguchi, Y., Yokobori, S.-I., Hashimoto, H., Yano, H., Tabata, M., Kawai, H., and Yamagishi, A. (2016) Investigation of the interplanetary transfer of microbes in the Tanpopo mission at the Exposed Facility of the International Space Station. *Astrobiology* 16:363–376.

Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M., and Johnson, J.E. (2005) Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc Natl Acad Sci USA* 102:18944–18949.

King, A.M.Q., Lefkowitz, E., Adams, M.J., and Carstens, E.B. (2011) *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses,* Academic Press, New York.

Koike, J., Oshima, T., Koike, K.A., Taguchi, H., Tanaka, R., Nishimura, K., and Miyaji, M. (1992) Survival rates of some terrestrial microorganisms under simulated space conditions. *Adv Space Res* 12:271–274.

Koonin, E.V. (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol* 72:2197–2206.

Koonin, E.V. (2011) *The Logic of Chance: The Nature and Origin of Biological Evolution,* FT Press Science, Upper Saddle River, NJ.

Koonin, E.V. (2016) Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Res* 5, doi:10.12688/f1000research.8737.1.

Koonin, E.V. (2017) Evolution of RNA-and DNA-guided antivirus defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Biol Direct* 12, doi:10.1186/s13062-017-0177-2.

Koonin, E.V. and Dolja, V.V. (2013) A virocentric perspective on the evolution of life. *Curr Opin Virol* 3:546–557.

Koonin, E.V. and Dolja, V.V. (2014) Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol Mol Biol Rev* 78:278–303.

Koonin, E.V., Senkevich, T.G., and Dolja, V.V. (2006) The ancient Virus World and evolution of cells. *Biol Direct* 1:29.

Koonin, E.V., Wolf, Y.I., Nagasaki, K., and Dolja, V.V. (2008) The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat Rev Microbiol* 6:925–939.

Koonin, E.V., Dolja, V.V., and Krupovič, M. (2015) Origins and evolution of viruses of eukaryotes: the ultimate modularity. *Virology* 479:2–25.

Krupovič, M. and Bamford, D.H. (2008) Virus evolution: how far does the double β-barrel viral lineage extend? *Nat Rev Microbiol* 6:941–948.

Krupovič, M. and Bamford, D.H. (2010) Order to the viral universe. *J Virol* 84:12476–12479.

Krupovič, M. and Bamford, D.H. (2011) Double-stranded DNA viruses: 20 families and only five different architectural principles for virion assembly. *Curr Opin Virol* 1:118–124.

Krupovič, M. and Koonin, E.V. (2017) Multiple origins of viral capsid proteins from cellular ancestors. *Proc Natl Acad Sci USA* 114, doi:10.1073/pnas.1621061114.

Kyle, J., Jahnke, L., and Stedman, K.M. (2012) Preservation potential of lipid-containing viruses under silicifying conditions. [abstract 2228]. In *43rd Lunar and Planetary Science Conference Abstracts,* Lunar and Planetary Institute, Houston.

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., Birtles, R., Claverie, J.-M., and Raoult, D. (2003) A giant virus in amoebae. *Science* 299:2033.

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P., Koonin, E., and Raoult, D. (2008) The virophage as a unique parasite of the giant mimivirus. *Nature* 455:100–104.

Laidler, J.R. and Stedman, K.M. (2010) Virus silicification under simulated hot spring conditions. *Astrobiology* 10:569–576.

Laidler, J.R., Shugart, J.A., Cady, S.L., Bahjat, K.S., and Stedman, K.M. (2013) Reversible inactivation and desiccation tolerance of silicified viruses. *J Virol* 87:13927–13929.

Lander, E.S., and the International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921.

Lavialle, C., Cornelis, G., Dupressoir, A., Esnault, C., Heidmann, O., Vernochet, C., and Heidmann, T. (2013) Paleovirology of *'syncytins',* retroviral *env* genes exapted for a role in placentation. *Proc R Soc Lond B Biol Sci* 368, doi:10.1098/rstb.2012.0507.

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M., Poirot, O., Bertaux, L., Bruley, C., Couté, Y., Rivkina, E., Abergel, C., and Claverie, J.-M. (2014) Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc Natl Acad Sci USA* 111:4274–4279.

Lindell, D., Jaffe, J.D., Johnson, Z.I., Church, G.M., and Chisholm, S.W. (2005) Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* 438:86–89.

Luria, S.E., Darnell, J.E., Baltimore, D., and Campbell, A. (1978) *General Virology,* John Wiley and Sons, New York.

Mann, N.H., Cook, A., Millard, A., Bailey, S., and Clokie, M. (2003) Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* 424, doi:10.1038/424741a.

Martin, A., Yeats, S., Janekovic, D., Reiter, W.-D., Aicher, W., and Zillig, W. (1984) SAV 1, a temperate UV-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. *EMBO J* 3:2165–2168.

McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science* 330, doi:10.1126/science.1192243.

Metcalf, J.A. and Bordenstein, S.R. (2012) The complexity of virus systems: the case of endosymbionts. *Curr Opin Microbiol* 15:546–552.

Mochizuki, T., Yoshida, T., Tanaka, R., Forterre, P., Sako, Y., and Prangishvili, D. (2010) Diversity of viruses of the hyperthermophilic archaeal genus *Aeropyrum,* and isolation of the *Aeropyrum pernix* bacilliform virus 1, APBV1, the first representative of the family *Clavaviridae. Virology* 402:347–354.

Mochizuki, T., Sako, Y., and Prangishvili, D. (2011) Provirus induction in hyperthermophilic archaea: characterization of *Aeropyrum pernix* spindle-shaped virus 1 and *Aeropyrum pernix* ovoid virus 1. *J Bacteriol* 193:5412–5419.

Mochizuki, T., Krupovič, M., Pehau-Arnaudet, G., Sako, Y., Forterre, P., and Prangishvili, D. (2012) Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc Natl Acad Sci USA* 109:13386–13391.

Mokili, J.L., Rohwer, F., and Dutilh, B.E. (2012) Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2:63–77.

Moon, B.Y., Park, J.Y., Hwang, S.Y., Robinson, D.A., Thomas, J.C., Fitzgerald, J.R., Park, Y.H., and Seo, K.S. (2015) Phage-mediated horizontal transfer of a *Staphylococcus aureus* virulence-associated genomic island. *Sci Rep* 5, doi:10.1038/srep09784.

Moreira, D. and López-García, P. (2015) Evolution of viruses and cells: do we need a fourth domain of life to explain the origin of eukaryotes? *Philos Trans R Soc Lond B Biol Sci* 370, doi:10.1098/rstb.2014.0327.

Nasir, A. and Caetano-Anollés, G. (2015) A phylogenomic data-driven exploration of viral origins and evolution. *Sci Adv* 1, doi:10.1126/sciadv.1500527.

Nigro, O.D., Jungbluth, S.P., Lin, H.-T., Hsieh, C.-C., Miranda, J.A., Schvarcz, C.R., Rappé, M.S., and Steward, G.F. (2017) Viruses in the oceanic basement. *mBio* 8, doi:10.1128/mBio.02129-16.

Orange, F., Chabin, A., Gorlas, A., Lucas-Staat, S., Geslin, C., Le Romancer, M., Prangishvili, D., Forterre, P., and Westall, F. (2011) Experimental fossilisation of viruses from extremophilic archaea. *Biogeosciences* 8:1465–1475.

Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N.N., and Kyrpides, N.C. (2016) Uncovering Earth's virome. *Nature* 536:425–430.

Parfenov, G.P. and Lukin, A.A. (1973) Results and prospects of microbiological studies in outer space. *Space Life Sci* 4:160–179.

Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A., Jacobs-Sera, D., Falbo, J., Gross, J., Pannunzio, N.R., Brucker, W., Kumar, V., Kandasamy, J., Keenan, L., Bardarov, S., Kirakov, J., Lawrence, J.G., Jacobs, W.R. Jr., Hendrix, R.W., and Hatfull, G.F. (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* 113:171–182.

Philippe, N., Legendre, M., Doutre, G., Coute, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J.M., and Abergel, C. (2013) Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341:281–286.

Pietilä, M.K., Roine, E., Paulin, L., Kalkkinen, N., and Bamford, D.H. (2009) An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. *Mol Microbiol* 72:307–319.

Pietilä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.-C., Jacobs-Sera, D., Hendrix, R.W., Bamford, D.H., and Butcher, S.J. (2013) Structure of the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. *Proc Natl Acad Sci USA* 110:10604–10609.

Poinar, G. and Poinar, R. (2005) Fossil evidence of insect pathogens. *J Invertebr Pathol* 89:243–250.

Pope, W.H., Bowman, C.A., Russell, D.A., Jacobs-Sera, D., Asai, D.J., Cresawn, S.G., Jacobs, W.R., Jr., Hendrix, R.W., Lawrence, J.G., and Hatfull, G.F. (2015) Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *eLife* 4, doi:10.7554/eLife.06416.

Prangishvili, D. (2003) Evolutionary insights from studies on viruses of hyperthermophilic archaea. *Res Microbiol* 154:289–294.

Prangishvili, D. (2013) The wonderful world of archaeal viruses. *Annu Rev Microbiol* 67:565–585.

Prangishvili, D., Basta, T., Garrett, R.A., and Krupovič, M. (2016) Viruses of the archaea. *eLS* doi:10.1002/9780470015902.a0000774.pub3.

Proctor, L.M. and Fuhrman, J.A. (1990) Viral mortality of marine bacteria and cyanobacteria. *Nature* 343:60–62.

Rachel, R., Bettstetter, M., Hedlund, B.P., Häring, M., Kessler, A., Stetter, K.O., and Prangishvili, D. (2002) Remark-

able morphological diversity of viruses and virus-like particles in hot terrestrial environments. *Arch Virol* 147:2419–2429.

Raoult, D. and Forterre, P. (2008) Redefining viruses: lessons from Mimivirus. *Nat Rev Microbiol* 6:315–319.

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., and Claverie, J.M. (2004) The 1.2-megabase genome sequence of Mimivirus. *Science* 306:1344–1350.

Rensen, E.I., Mochizuki, T., Quemin, E., Schouten, S., Krupovič, M., and Prangishvili, D. (2016) A virus of hyperthermophilic archaea with a unique architecture among DNA viruses. *Proc Natl Acad Sci USA* 113:2478–2483.

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T., and Young, M.J. (2001) Viruses from extreme thermal environments. *Proc Natl Acad Sci USA* 98:13341–13345.

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T., and Young, M. (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc Natl Acad Sci USA* 101:7716–7720.

Robertson, M.P. and Joyce, G.F. (2014) Highly efficient self-replicating RNA enzymes. *Chem Biol* 21:238–245.

Rohwer, F. (2003) Global phage diversity. *Cell* 113:141.

Roossinck, M.J. (2011) The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol* 9:99–108.

Roossinck, M.J. (2015) Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *J Virol* 89:6532–6535.

Roossinck, M.J. (2016) *Virus: An Illustrated Guide to 101 Incredible Microbes,* Princeton University Press, Princeton, NJ.

Rosario, K. and Breitbart, M. (2011) Exploring the viral world through metagenomics. *Curr Opin Virol* 1:289–297.

Roux, S., Enault, F., Bronner, G., Vaulot, D., Forterre, P., and Krupovič, M. (2013) Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nat Commun* 4, doi:10.1038/ncomms3700.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., Poulos, B.T., Solonenko, N., Lara, E., Poulain, J., Pesant, S., Kandels-Lewis, S., Dimier, C., Picheral, M., Searson, S., Cruaud, C., Alberti, A., Duarte, C.M., Gasol, J.M., Vaqué, D.; Tara Oceans Coordinators, Bork, P., Acinas, S.G., Wincker, P., and Sullivan, M.B. (2016) Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537:689–693.

Ruardij, P., Veldhuis, M.J.W., and Brussaard, C.P.D. (2005) Modeling the bloom dynamics of the polymorphic phytoplankter *Phaeocystis globosa:* impact of grazers and viruses. *Harmful Algae* 4:941–963.

Sankaran, N. (2016) The RNA world at thirty: a look back with its author. *J Mol Evol* 83:169–175.

Schmidt, H. and Hawkins, A.R. (2016) Single-virus analysis through chip-based optical detection. *Bioanalysis* 8:867–870.

Schulz, F., Yutin, N., Ivanova, N.N., Ortega, D.R., Lee, T.K., Vierheilig, J., Daims, H., Horn, M., Wagner, M., Jensen, G.J., Kyrpides, N.C., Koonin, E.V., and Woyke, T. (2017) Giant viruses with an expanded complement of translation system components. *Science* 356:82–85.

Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., and Zhang, Y.Z. (2016) Redefining the invertebrate RNA virosphere. *Nature* 540:539–543.

Sime-Ngando, T., Lucas, S., Robin, A., Tucker, K.P., Colombet, J., Bettarel, Y., Desmond, E., Gribaldo, S., Forterre, P., and Breitbart, M. (2011) Diversity of virus–host systems in hypersaline Lake Retba, Senegal. *Environ Microbiol* 13:1956–1972.

Simmonds, P., Adams, M.J., Benkő, M., Breitbart, M., Brister, J.R., Carstens, E.B., Davison, A.J., Delwart, E., Gorbalenya, A.E., Harrach, B., Hull, R., King, A.M., Koonin, E.V., Krupovic, M., Kuhn, J.H., Lefkowitz, E.J., Nibert, M.L., Orton, R., Roossinck, M.J., Sabanadzovic, S., Sullivan, M.B., Suttle, C.A., Tesh, R.B., van der Vlugt, R.A., Varsani, A., and Zerbini, F.M. (2017) Consensus statement: virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* 15:161–168.

Snyder, J.C., Bolduc, B., and Young, M.J. (2015) 40 years of archaeal virology: expanding viral diversity. *Virology* 479:369–378.

Sparks, W.B., Hand, K.P., McGrath, M.A., Bergeron, E., Cracraft, M., and Deustua, S.E. (2016) Probing for evidence of plumes on Europa with HST/STIS. *Astrophys J* 829, doi:10.3847/0004-637X/829/2/121.

Stec, K.F., Caputi, L., Buttigieg, P.L., D'Alelio, D., Ibarbalz, F.M., Sullivan, M.B., Chaffron, S., Bowler, C., d'Alcalà, M.R., and Iudicone, D. (2017) Modelling plankton ecosystems in the meta-omics era. Are we ready? *Mar Genomics* 32:1–17.

Stedman, K. and Blumberg, B.S. (2005) The NASA Astrobiology Institute Virus Focus Group workshop and field trip to Mono and Mammoth Lakes, CA, June 22–24, 2004. *Astrobiology* 5:441–443.

Stedman, K.M. (2015) Deep recombination: RNA and ssDNA virus genes in DNA virus and host genomes. *Annu Rev Virol* 2:203–217.

Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260:170–173.

Sullivan, M.B., Lindell, D., Lee, J.A., Thompson, L.R., Bielawski, J.P., and Chisholm, S.W. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol* 4:1344–1357.

Suttle, C.A. (2005) Viruses in the sea. *Nature* 437:356–361.

Suttle, C.A. (2007) Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol* 5:801–812.

Suttle, C.A. (2009) Viruses: a vast reservoir of genetic diversity and driver of global processes. *Retrovirology* 6, doi:10.1186/1742-4690-6-S2-I7.

Szathmary, E. and Smith, J.M. (1997) From replicators to reproducers: the first major transitions leading to life. *J Theor Biol* 187:555–571.

Takemura, M. (2001) Poxviruses and the origin of the eukaryotic nucleus. *J Mol Evol* 52:419–425.

Turner, P.E. and Chao, L. (1999) Prisoner's dilemma in an RNA virus. *Nature* 398:441–443.

Urayama, S.-i., Takaki, Y., and Nunoura, T. (2016) FLDS: a comprehensive dsRNA sequencing method for intracellular RNA virus surveillance. *Microbes Environ* 31:33–40.

Van Valen, L. (1973) A new evolutionary law. *Evolutionary Theory* 1:1–30.

Villarreal, L.P. (2004) *Viruses and the Evolution of Life,* ASM Press, Washington, DC.

Villarreal, L.P. and DeFilippis, V.R. (2000) A hypothesis for DNA viruses as the origin of eukaryotic replication proteins. *J Virol* 74:7079–7084.

Weinbauer, M.G. (2004) Ecology of prokaryotic viruses. *FEMS Microbiol Rev* 28:127–181.

Wells, L.E. and Deming, J.W. (2006) Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environ Microbiol* 8:1115–1121.

Wichman, H.A. and Brown, C.J. (2010) Experimental evolution of viruses: *Microviridae* as a model system. *Philos Trans R Soc Lond B Biol Sci* 365:2495–2501.

Wilhelm, S.W. and Suttle, C.A. (1999) Viruses and nutrient cycles in the sea viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* 49: 781–788.

Wilmes, P., Simmons, S.L., Denef, V.J., and Banfield, J.F. (2009) The dynamic genetic repertoire of microbial communities. *FEMS Microbiol Rev* 33:109–132.

Wilson, W.H., Van, Etten, J.L., and Allen, M.J. (2009) The Phycodnaviridae: the story of how tiny giants rule the world. *Curr Top Microbiol Immunol* 328:1–42.

Witzany, G. (2008) The viral origins of telomeres and telomerases and their important role in eukaryogenesis and genome maintenance. *Biosemiotics* 1:191–206.

Woese, C.R. (2002) On the evolution of cells. *Proc Natl Acad Sci USA* 99:8742–8747.

Wommack, K.E. and Colwell, R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* 64: 69–114.

Wood, A.G., Whitman, W.B., and Konisky, J. (1989) Isolation and characterization of an archaebacterial virus-like particle from *Methanococcus voltae* A3. *J Bacteriol* 171: 93–98.

Wright, A.V., Nuñez, J.K., and Doudna, J.A. (2016) Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164:29–44.

Yang, L. and Yamamoto, T. (2016) Quantification of virus particles using nanopore-based resistive-pulse sensing techniques. *Front Microbiol* 7, doi:10.3389/fmicb.2016.01500.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., and Takai, K. (2013) Metagenomic analysis of viral communities in (hado)pelagic sediments. *PLoS One* 8, doi:10.1371/journal .pone.0057271.

Yozwiak, N.L., Skewes-Cox, P., Stenglein, M.D., Balmaseda, A., Harris, E., and DeRisi, J.L. (2012) Virus identification in unknown tropical febrile illness cases using deep sequ-encing. *PLoS Negl Trop Dis* 6, doi:10.1371/journal.pntd .0001485.

Yu, M.X., Slater, M.R., and Ackermann, H.W. (2006) Isolation and characterization of *Thermus* bacteriophages. *Arch Virol* 151:663–679.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., Lanzendörfer, M., and Kristjansson, J.K. (1993) Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfataras. *Syst Appl Microbiol* 16:609–628.

Address correspondence to:
Kenneth M. Stedman
Biology Department
Portland State University
P.O. Box 751
Portland, OR 97207-0751

E-mail: kstedman@pdx.edu

Abbreviations Used

ACV = *Aeropyrum* coil-shaped virus
AMG = auxiliary metabolic genes
CRISPRs = Clustered Regularly Interspersed Palindromic Repeats
dsDNA = double-stranded DNA
HGT = horizontal gene transfer
ICTV = International Committee on Taxonomy of Viruses
ISS = International Space Station
NAI = NASA Astrobiology Institute
NAIViFoG = NASA Astrobiology Institute Virus Focus Group
SEM = scanning electron microscope
ssDNA = single-stranded DNA
TEM = transmission electron microscope
VLPs = virus-like particles