

The Origin and Evolution of Viruses as Molecular Organisms

Claudiu I. Bandea

National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333

cbandea@cdc.gov / October 20, 2009

Viruses are the most abundant life forms and the repertoire of viral genes is greater than that of cellular genes. It is also evident that viruses have played a major role in driving cellular evolution, and yet, viruses are not part of mainstream biology, nor are they included in the Tree of Life. A reason for this major paradox in biology is the misleading dogma of viruses as viral particles and their enigmatic evolutionary origin. This article presents an alternative view about the nature of viruses based on their properties during the intracellular stage of their life cycle, when viruses express features comparable to those of many parasitic cellular species. Supporting this view about the nature of viruses is a novel hypothetical evolutionary model for their origin from parasitic cellular species that fused with their host cells. By losing their membrane and cellular structure within the host cell, these new types of parasitic species gained full access to precursors for the synthesis of their specific molecules and to the host's information processing machineries, such as translation, which created unique parasitic and evolutionary opportunities. To identify viruses during their intracellular stage of their life cycle, in which their specific molecules are free or dispersed within the host cell, this paper introduces the concept of "molecular structure" and labels viruses as "molecular organisms." Among the extant viruses, the life cycle of poxviruses and other complex viruses that fuse with their host cells provides compelling evidence for the fusion model. One of the most remarkable implications of fusion model is that new viral lineages originated from parasitic cellular species throughout the history of life, and that this process might still be active. Surprisingly, it appears that several parasitic cellular species are currently evolving into molecular organisms. More remarkably though, according to this model, several parasites that are currently classified as cellular organisms are in fact genuine molecular organisms. The current evidence for the fusion hypothesis is strong and it is fully testable using both experimental and phylogenetic approaches. The academic and research implications of this model, which supports the inclusion of viruses in the Tree of Life, are highly significant. Some of these implications are discussed in more detail in two other articles of this series, which presents a unifying model for the origin and evolution of cellular and viral domains, including the origin of life.

Introduction

Viruses are the most abundant life forms and the repertoire of viral genes is greater than that of cellular genes (1-3). It is also evident that viruses have played a major role in driving cellular evolution (4-12), and yet, viruses are set outside of mainstream biology and evolutionary paradigms and are not included in the Tree of Life (TOL) (13-15). A reason for this major paradox in biology is the misleading dogma of viruses as viral particles and their enigmatic evolutionary origin.

Since their discovery a century ago, viruses have been conceptually identified with their viral particles and defined based on the properties of these particles. The viral particles are highly specialized structures that are used by many

viruses for their transmission to new host cells [for comprehensive facts about viruses and their life cycle see (16)]. This role of viral particles in the viral life cycle explains their particular properties, such as their apparent inert status or the presence of only one type of nucleic acid - DNA or RNA. Many viruses, however, do not produce viral particles, using alternative modes of transmission, such as vertical transmission from mother to daughter cells (17-19). Clearly, the fundamental biological properties of all viruses, whether they do or do not produce viral particles are expressed during the intracellular stage of the viral life cycle, when viruses replicate their genome and synthesize their specific molecules, many of which are not components of the viral particles.

Based on these facts about the viral life cycle and on a reductive model for their origin and evolution from parasitic cellular species, I suggested in the early 1980's that the dogma of viruses as viral particles misrepresents their nature, and that viruses should be identified with their intracellular forms (20). Since this early publication, it has become increasingly evident that this dogma can no longer be reconciled with modern advances and knowledge about viruses, and similar, independent views about the nature of viruses are emerging (4;21). For instance, in reflecting on the dogma of viruses as viral particles, Jean-Michel Claverie recently asked "what if we have totally missed the true nature of (at least some) viruses?"(4), and Didier Raoult and Patrick Forterre argued for a change of the way we define viruses (21).

This paper presents a hypothetical fusion model for the origin of viral lineages from parasitic cellular species that in order to gain better access to host resources fused with their host cell. According to this model, thousands of parasitic cellular species evolved into viral lineages throughout the history of life and, remarkably, there is evidence that this process is still going on. This paper outlines also the evolutionary diversification of viral lineages by reductive evolution into a myriad of viral species with novel modes of reproduction and transmission, and it concludes with a proposal for including viruses in the TOL. The evidence for the fusion model is strong and the academic and research implications could be highly significant. One of the most pragmatic research application of the new view about the nature and evolution viruses is in the field of transmissible spongiform encephalopathies (TSE), or prion diseases, which is presented in a separate article (22). This is one of a series of three articles that present a broad, unifying scenario on the origin and evolution of viral and cellular domains that could lead to a significant paradigm shift in biology (23).

The dogma of viruses as viral particles

The dogma of viruses as viral particles has a long history, which started at the turn of the last century when viruses were first recognized as a distinct group of infectious agents [reviewed in (16;24;25)]. Questioning the validity of this dogma, which has guided several generations of researchers to extraordinary discoveries and progress in virology and in related biomedical fields, is challenging. Without doubt, by focusing on the infectious stage in the viral life cycle, the dogma of viruses as viral particles has been an intuitive and practical approach for conducting research, particularly from a medical and public health perspective. However, as argued throughout this series of articles, this dogma is conceptually and evolutionarily flawed, and more importantly, it might have constrained the full potential for progress in many biomedical fields (22;23). Considering the significance of this dogma, it is essential to present its historical development, at least briefly.

In 1892, Russian botanist Dimitri Ivanovsky published

the first report showing that the agent causing tobacco mosaic disease passed through porcelain filters that were presumed to retain all cellular microbial pathogens [reviewed in (16;24;25)]. Ivanovsky proposed that the agent that passed through the very small pores of the porcelain filters and caused the diseases was not the tobacco mosaic microbial pathogen, but a toxin produced by this pathogen. A few years later, Martinus Beijerinck, a Dutch microbiologist, showed that the filterable tobacco mosaic agent reproduced in newly infected plants and concluded that the agent was a microbe, not a toxin. To explain the property of the tobacco mosaic microbe to pass through the very small pores of the porcelain filters, Beijerinck proposed that the pathogen was not a cellular microorganism, nor a "particulate" entity, but a novel type of microorganism that had a fluid structure, which allowed it to pass through very small pores. Based on this and additional experiments, Beijerinck proposed that the tobacco mosaic infectious agent was a "contagium vivum fluidum"- a contagious living fluid.

During the next few decades numerous "filterable agents," or viruses as they were eventually labeled, infecting plants, animals, and bacteria were identified. Although Ivanovsky and other researchers during that early period did study the intracellular form of these pathogens, the focus of the research and that of the developing concept about their nature, focused on the transmissible, infectious units - the viral particles [reviewed in (16;24-26)]. In 1935, American biochemist Wendell Stanley crystallized tobacco mosaic virus (i.e. the viral particles of the virus) showing that this and presumably all other viruses, although did not have a fluid structure, were fundamentally different from the conventional cellular microorganisms (27). Soon, it was discovered that viruses contained primarily proteins and nucleic acids (28), however unlike cellular microorganisms, apparently they contained only one type of nucleic acid, either DNA or RNA. This presumptive fundamental biochemical difference between viruses and cells has been one of the defining properties of viruses to this day. For example, in the 1970s in his seminal book *The Molecular Biology of the Gene*, James Watson wrote "all viruses differ fundamentally from cells, which have both DNA and RNA, in that viruses contain only one type of nucleic acid, which may be either DNA or RNA" (29). A decade later, in *A Dictionary of Virology*, viruses were defined as "Infectious units consisting of either RNA or DNA enclosed in a protective coat" (30). In the 1990s, a classical microbiology text stated that viruses "consist of a genome, either RNA or DNA, that is surrounded by a protective protein shell" (31). And, more recently, viruses have been described in the following terms "The simplest viruses consist of a protein coat made up primarily of many copies of a single polypeptide chain surrounding a small genome composed of as few as three genes. More complex viruses have larger genomes of up to several hundred genes, surrounded by an elaborate shell composed of many different proteins" (32). In addition to describing viruses as having only one type of

nucleic acid and a relatively simple structure, many publications, including most biology, microbiology, or virology textbooks published in the last half century, describe viruses as having no metabolic activity or growth in size.

Interestingly, during the 1950s and 1960s, when the emerging knowledge about the molecular biology of viruses was integrated into the dogma of viruses as viral particles, Andre Lwoff a prominent French scientist and one of the founders of modern virology proposed in his article, “The Concept of Virus,” that viruses should be defined primarily based on their properties during the intracellular, vegetative stage of their life cycle (33). Lwoff, however, defined the intracellular viruses with the same properties as those of the viral particles (Box 1) and, ironically, his paper reinforced the dogma of viruses as viral particles. By the 1970s, this dogma was well established. Even clear evidence to the contrary had little effect. For instance, at that time, it was common knowledge that during their intracellular stage, “DNA viruses” have both nucleic acids, DNA and RNA, yet because of the dogma viruses as viral particles, viruses were defined as having only one type of nucleic acid (see quotes above). This is strong example of the power of concepts in science; a concept that misrepresents the facts can still be viable for decades.

As argued throughout this and the other papers in this series (22;23), defining viruses using the properties of their viral particles misrepresents their nature. Metaphorically speaking, comparing the viral particles with ordinary cells is like comparing apples and oranges. Similar to other intracellular parasitic species, viruses pass through several stages in their life cycle and, thus, an integrated sum of all these stages defines them. Also, similar to many other intracellular parasitic species, viruses synthesize their molecules, replicate their genome, and reproduce exclusively within their unique environment – the host cell. Therefore, the structural and biological properties of viruses during this intracellular stage, which is the mature, reproductive phase in the viral life cycle, should be primarily used to define viruses, and these properties should be compared to those of other intracellular parasites. The evident problem with this approach is that, unlike parasitic cellular species, during the intracellular phase of the viral life cycle the viral molecules are not surrounded by a virus specific membrane, but are more or less free or dispersed within the host cell. Historically, this phase of the viral life cycle was referred to as an “eclipse phase” in order to denote the disappearance of viruses (i.e. the disappearance of viral particles). Although no reputable scholar in this field believed that viruses literally disappear, no clear solutions on how to identify viruses during the intracellular phase of their life cycle have been proposed [discussed in more detail in (20)], and the dogma of viruses as viral particles has remained in effect to this day. A potential solution to this major scientific issue is presented next in the context of a hypothetical fusion model for the origin of viruses as organisms with a novel type of biological

Box 1. Lowff’s definition of viruses based on their intracellular [quote (a)] or extracellular [quote (b)] properties:

(a) “*strictly intracellular and potentially pathogenic entities with an infectious phase*, and (i) possessing only one type of nucleic acid, (ii) multiplying in the form of their genetic material, (iii) unable to grow and to undergo binary fission, (iv) devoid of a Lipmann system”^a

(b) “*infectious, potentially pathogenic, nucleoproteinic entities* (i) possessing only one type of nucleic acid, which (ii) are reproduced from their genetic material, (iii) are unable to grow and to undergo binary fission, and (iv) are devoid of a Lipmann system”^b

^a italicization present in the original text; Lipmann system refers to energy metabolism.

^b italicization and (i) to (iv) were added to the original text.

organization and structure.

A fusion model for the origin of viruses

In exploring the origin and evolution of viruses, it is useful to conceive a first or ancestral virus (AV). Because viruses coevolved with their host cells, it is relevant to discuss the origin and evolution of AV in reference to the Last Universal Common Ancestor (LUCA) of the three cellular domains - Archaea, Bacteria, and Eukarya (34-38). Using LUCA as a reference, the hypotheses on the origin of AV fall within three broad evolutionary pathways (6;20;39):

- [i] Origin from pre-cellular genetic elements that evolved as parasites of the LUCA lineage and eventually as parasites of the three evolving cellular domains (referred to as Pre-cellular Theory);
- [ii] Evolution from escaped cellular genetic material of LUCA lineage or that of bacterial, archaeal and eukaryal species (Endogenous Theory);
- [iii] Origin from parasitic cellular organisms within the LUCA lineage, or from parasitic bacterial, archaeal and eukaryal species by reductive evolution (Reductive Theory).

Regarding the evolutionary pathway for the diversification of AV leading to current viral families, the Pre-cellular and Endogenous Theories imply that AV was a relatively simple parasitic genetic element that acquired new genes and evolved and diversified into more complex viruses with larger genomes. On the contrary, the Reductive Theory proposes that AV was a rather complex parasitic cellular species that, because of its unique environment (the

host cell), lost a large number of genes and evolved and diversified into less complex viruses with smaller genomes. All three evolutionary pathways allow for independent origin of multiple AVs (i.e. polyphyletic origin), but only the Endogenous and the Reductive Theories share the interesting possibility that AVs originated from the LUCA lineage and its descendants - Archaea, Bacteria, and Eukarya - throughout their evolutionary history, with the remarkable prospect that this process is still active. It should be emphasized also at this point that, in exploring the origin and evolution of viruses, it is important to distinguish between the origin of AVs and their evolutionary diversification leading to the origin of new viral families or species of various complexity and life cycle, which is discussed latter in this article.

Historically, the hypotheses on the origin and evolution of viruses followed generally the views about their nature (6;20;39). This is strong testimony to the significant influence of the view about the nature of viruses on the views about their evolutionary origin (see below). Early on, when viruses were discovered as filterable agents, it was reasonable to think that these minute parasitic species originated from more complex parasites by reductive evolution. After it was shown, however, that some viruses (i.e. viral particles) have a non-cellular structure and a relatively simple biochemical composition reminiscent of those of the hypothetical first living entities, the hypothesis on their ancient origin from pre-cellular organisms was popular. And later on, when the mobile genetic elements were discovered, the Endogenous Theory on the origin of viruses from escaped cellular genetic material became prevalent.

Until the last few decades, the knowledge about viruses and other microorganisms was limited, precluding development of comprehensive theories about their origin and evolution. Moreover, because of the misleading view about the nature of viruses, the hypotheses about their origin focused on the viral particle, a constraint that affected in particular the reductive hypothesis. This is eloquently described by Salvador Luria and James Darnell: "The strongest argument against the regressive origin of viruses from cellular parasites is the non-cellular organization of viruses. The viral capsids are morphogenetically analogous to cellular organelles made up of protein subunits, such as bacterial flagella, actin filaments, and the like, and not to cellular membranes.... This theory today has little to recommend it, at least in its original form." (39).

In the evolutionary model for the origin of viruses presented earlier (20), I proposed that the viral lineages evolved from intracellular parasitic species that gradually modified their cellular membrane to allow increased access to the host cell resources, eventually losing this membrane within the host cell. Similar views about the potential advantages gained by parasites by losing their cellular membrane within the host cell were suggested half a century earlier by Australian biologist McFarlane Burnet (40;41), and more recently by Richard Matthews (42), a

plant virologist from New Zealand.

Based on an earlier suggestion about the significance of "cellular hybridization," or cellular fusion, for early cellular evolution (20), I present here a fusion model for the origin of AVs. This evolutionary model proposes that viral lineages originated from parasitic species that fused with their host cells, by a process in which the parasite cellular membrane fused with the host's cellular membrane. By losing their cellular membrane and having their components free or dispersed within the host cell, these parasitic species had better access to the host cells' resources such as amino acids, lipids, and nucleotides for the synthesis of their specific molecules. More significantly, however, these parasites had access to the host cell's information processing machineries, particularly to translation, which created unique parasitic and evolutionary opportunities. After synthesizing their molecules and multiple rounds of replication of their genome using host resources, these viral parasitic species directed the morphogenesis of cell-like reproductive forms that differentiated into infectious, spore-like structures, which started a new life cycle by fusing with other host cells.

Among extant viruses, the life cycle of poxviruses provides compelling evidence for the fusion model (Fig. 1):

- [i] Poxviral particles, which are cell-like structures containing a nucleo-core, fuse with their host cells by a mechanism in which the pox-viral membrane fuses with the host's cellular membrane (16;43);
- [ii] After fusing with their host cell, poxviruses transcribe their early genes (approximately 100 genes) within the nucleo-core using their own transcription machinery, which includes a multi-subunit DNA-dependent RNA polymerase, transcription factors and topoisomerases, as well as enzymes for capping, methylation, and polyadenylation of the transcripts. Then, poxviruses synthesize their proteins using the host's translation apparatus, and replicate their genome using their own replication machinery. These viral activities are associated with a microscopically distinct conglomerate of viral molecules called a viroplasm (44;45) (Fig. 1-1);
- [iii] To reproduce, poxviruses induce the assembly of viral membranes at the periphery of the viroplasm, forming cell-like structures that include a genome, complete transcriptional machinery, and many other viral molecules (Fig. 1-1). These newly-formed poxviral cell-like reproductive forms differentiate by a *bona fide* cellular differentiation process into transmissible, infectious forms (Fig. 1-2a/f). The assembly of poxviral cell-like structures at the periphery of the viroplasms is fundamentally similar to assembly within parental yeast cells of progenies that eventually differentiate into yeast spores (46;47). Apparently, in poxviruses and some other complex viruses (45;48),

the entire viroplasm is enclosed in a virus-specific membrane, which is a strong indication of their ancestors' cellular structure;

[iv] To complete their life cycle, the poxviral particles exit the host cell by diverse mechanisms controlled by the parental virus. After the release of its progeny, the parental viruses as well as the host cell die.

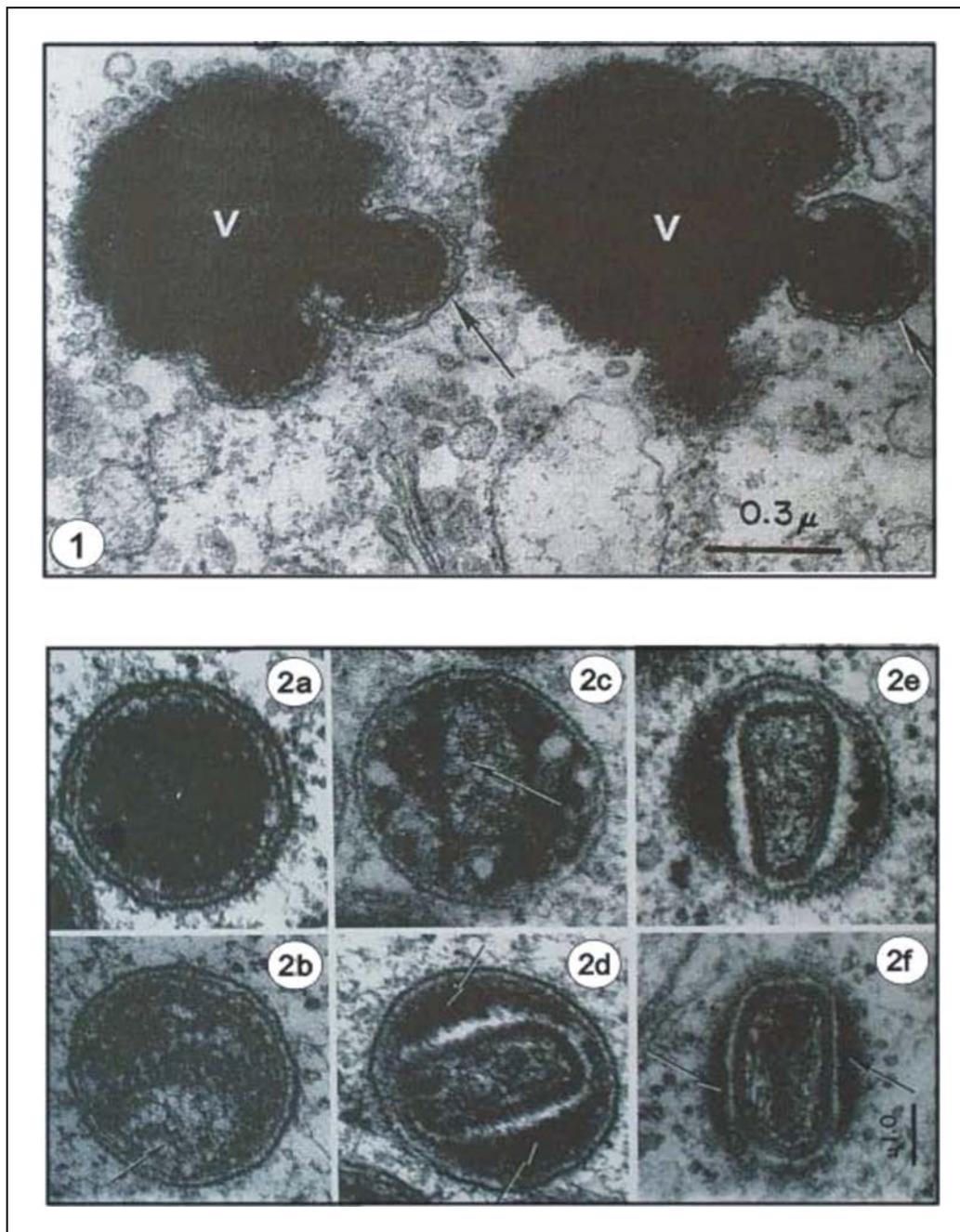


Figure 1. Morphogenesis of cell-like poxviral reproductive forms and their differentiation into viral particles [from (48); See (16;49) for similar electron micrographs]. 1. Morphogenesis of poxviral cell-like structures (arrows) at the periphery of two parental viroplasm (V). 2. Differentiation of a newly-assembled poxviral cell-like reproductive form into an infectious viral particle (2a to 2f). The differentiation of the nucleocore is evident (arrows). Reproduced by the kind permission of Elsevier from reference (48)

As mentioned in the previous section, due to the dogma of viruses as viral particles, the question of how to identify poxviruses and other viruses during their intracellular stage of their life cycle was rarely addressed. The agnostic approach of not addressing this issue notwithstanding, there are several potential solutions [see (20) for a more detailed discussion of this issue]. A limited approach would be to identify the viral organisms during their intracellular stage of their life cycle with their genome, excluding all the other specific viral molecules, such as the viral transcripts or proteins. On the contrary, taking an extreme integrative approach, viruses could be identified with the entire host cell/virus complex. I believe neither of these two approaches respects the individuality of viral organisms, and they do not facilitate a balanced comparison of viruses with other intracellular parasitic or symbiotic species in order to define their fundamental biological properties. Alternatively, I proposed to identify viruses during their intracellular development with all their specific molecules, which are more or less free or dispersed within the host cell (20). However, if at a particular stage during their intracellular development some viruses consist solely of their genome and have no other specific molecules, then obviously during that particular stage of their life cycle these viruses should be identified with their genome.

As a descriptive way to identify viruses during intracellular development with the *integrative sum of all their molecules*, this paper proposes the concept of a *molecular structure*, and it labels viruses and the related viral elements as *molecular organisms*. Although, some

viruses do develop a cell-like structure during some stages of their life cycle and, obviously, the cellular organisms are also made of molecules, this new concept is specifically intended to emphasize the fundamental difference between the viral and cellular parasitic organisms during their intracellular development, when both types of organisms express similar biological properties (Fig. 2; see also Box 2 in a later section of this paper).

The full relevance and implications of a molecular structure in the co-evolution of viruses and their host cells is further discussed here and in more detail in another paper of this series (23), but it implies that in order to take full advantage of the host resources, including the host's replication, transcription, and translation machineries, the molecular biology of the parasite and that of the host cell must be compatible. Therefore, the fusion model predicts that only parasites which infect hosts within the same cellular domain would have the opportunity to evolve into viral lineages - the AVs. Accordingly, the ancestors of poxvirus-like AVs were probably parasitic eukaryal cellular organisms, and the ancestors of the AVs infecting Bacteria or Archaea were parasitic bacterial or archaeal species, respectively. It is very likely, however, that the first parasitic viral species originated from the LUCA lineages before the origin of cellular domains, and that these early viral species played a significant role in the origin of cellular domains (23). The significance of evaluating the hypotheses on the evolutionary origin of viruses before or after the origin of cellular domains has been eloquently discussed in other publications (6;7).

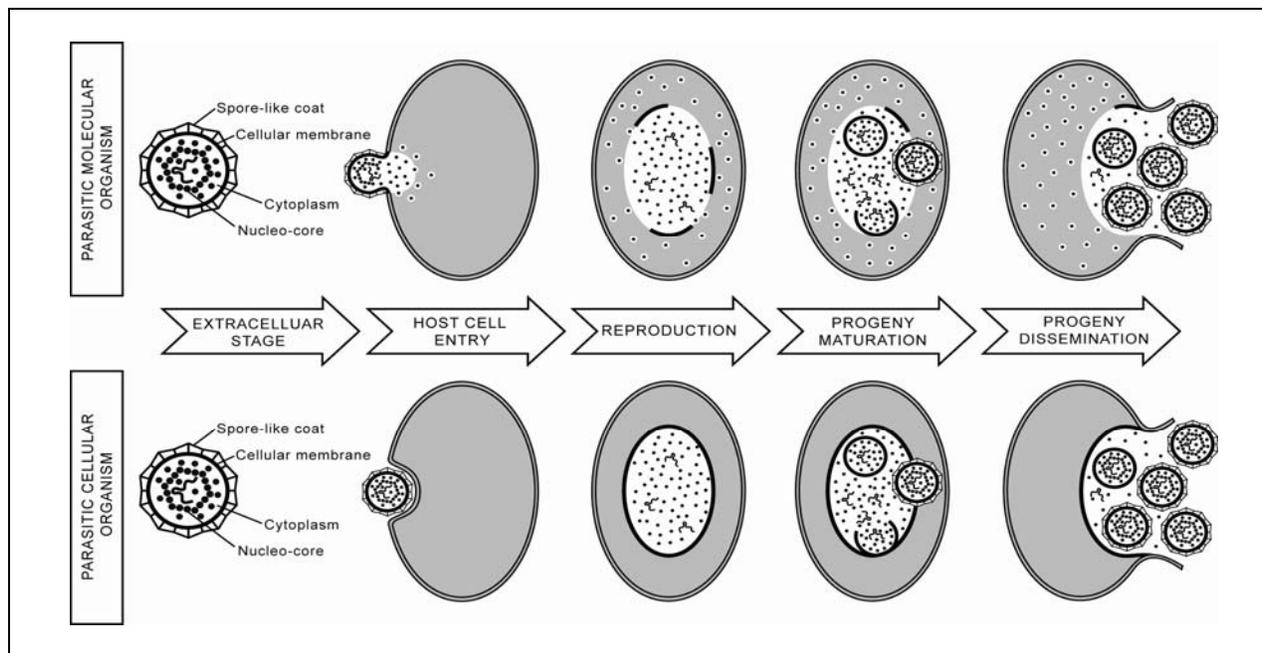


Figure 2. Comparison of the life cycle of a generic intracellular parasite that maintains a cellular structure within its host cell (lower panel) with that of a parasite that fuses with its host cell and develops a molecular structure (upper panel)

There is no doubt that the concepts of molecular structure and molecular organism are more suggestively envisioned within the framework of the fusion model for the origin of viruses from cellular parasitic species (Fig. 2), but these concepts might still apply even if viruses originated and evolved according to the other two theories. Although circumstantial, the evolutionary trends of extant viruses might be the most compelling evidence for evaluating the hypotheses on the origin of viruses. If this evidence indicates that, in general, viruses have been evolving towards more complex viruses, it would make sense to favor the view predicted by the pre-cellular and endogenous theories that the AVs were relatively simple genetic elements. On the contrary, if this evidence indicates that the general trend of viral lineages has been to evolve into less complex viruses with smaller genome, then it would make more sense to favor the Reductive Theory.

Because of their relative small genome and strong lateral gene transfer (LGT), the current general trend in evolution of viruses is difficult to assess (50). But, as more viral genomes are sequenced and the current phylogenetic analysis programs improve, this trend might become more evident. There is, however, overwhelming evidence for a reductive evolutionary trend parasitic and endo-symbiotic cellular species [reviewed in (51-55)]. The key for understanding the reductive evolutionary trend of intracellular parasitic or symbiotic species is their unique environment - the host cell. This particular environment can provide these species basically with all the precursors for the synthesis of their specific molecules, which generates the opportunity for species to lose many of their genes and evolve smaller genomes. There is no doubt that in order to overcome the host defense systems or, in the case of parasitic cellular species to evolve elaborate mechanisms for transporting precursors such as amino acids, lipids, or nucleotides across their cellular membrane, some the intracellular parasites have gained new genes. As compared to their non-parasitic or non-symbiotic relatives, however, all intracellular parasites and endosymbionts have a smaller genome [reviewed in (51-55)].

According to the Pre-cellular Theory on the origin of viruses, which currently is the prevalent view (6;7;50), viruses originated before the evolution of first cellular organisms, 3 to 4 billion years ago. This theory has been particularly attractive because, in conjunction with the presumptive RNA world (56), it provides an relatively easy solution to the question regarding the evolutionary origin of "RNA viruses" (6;7;50). In the absence of traditional fossil records, direct evidence for the pre-cellular theory would come primarily from phylogenetic analysis of extant viral genomic or protein sequences. Due to a high evolutionary rate of RNA genomes, however, the sequence homology among "RNA viruses" descending from a common ancestor are expected to vanish in a few dozens of thousands of years, at the most (8;57;58). And, although the evolutionary rate of "DNA viruses" is several orders of magnitude lower than that of "RNA viruses", their evolutionary relationship

would be difficult to recognize after a relatively short period on the geological time scale [it should be noted, however, that this rapid evolution doesn't necessarily apply to endogenous viruses whose genome is replicated as part of the host genome (59)]. Therefore, using extant viral genomic or protein sequences to infer potential evolutionary relations among of viruses going back a few billion or even a few hundred million years is questionable.¹ This leaves the Pre-cellular Theory with little sequence-based phylogenetic supporting evidence. Similarly, there is no evidence for the Endogenous Theory. Apparently, the cellular mobile genetic elements that were previously considered potential endogenous cellular precursors for the origin of viruses (61) are viral evolutionary remnants (62-65).

The Reductive Theory on the origin of AV was questioned (39) because it could not be reconciled with the misleading concept of viruses as viral particles. However, in the context of the concept of viruses as molecular organisms and that of the fusion model for the origin of AVs, the evidence for the origin of viruses from parasitic cellular species is strong. Clearly, the life cycle of poxviruses is an ideal model for the fusion model. It would be difficult to even theoretically conceive a stronger case. The life cycle of poxviruses is exactly what the fusion model for the origin of AV predicts. Moreover, the recent discovery of mimiviruses, which have several indisputable cellular remnants, including: (a) a cell-like structure during at least some stages of their intracellular development; (b) cell-like reproductive forms; (c) a genome larger than that of many parasitic or symbiotic cellular species; (c) remnants of a translation machinery, and (v) presence of both DNA and RNA in their cell-like viral particles, is strong evidence for their origin from parasitic cellular species (66-69). Particularly, the presence of remnants of translation machinery, a feature found also in other viruses such as chloroviruses (70), might be best explained by a reductive evolution from a parasitic cellular species as predicted by the fusion model.

An obvious strength of the fusion model, compared to the other hypotheses on the origin of viruses, is that it can be approached experimentally. Cellular parasitic species can be fused with their host cells in order to study their potential development as parasitic molecular organisms. Moreover, if new viral lineages originated from parasitic cellular species throughout life's history as predicted by the

¹ These sequence-based phylogenetic issues are inherent not only to viruses but also to parasitic cellular species that evolved at a different rate compared to their non-parasitic relatives. For instance, based on the phylogenetic analysis of ribosomal genes, which are among the most conserved cellular genes, the microsporidia were initially classified as pre-eukaryal species [reviewed in (60)]. Recently, however, it was shown that the ancestors of these highly reduced parasitic species were sexually reproducing fungal species that due to their unique intracellular environment have lost not only numerous genes but also their mitochondria [reviewed in (60)].

fusion model, we should expect to find extant cellular species on their way to evolving into molecular organisms. The current scientific literature abounds with examples of parasites and symbionts that have evolved by reductive evolution from more complex organisms (51-55). As presented in the next section, using the fusion model for the origin viruses as a working hypothesis, I identified several bacterial, archaeal, and eukaryal species that might be evolving into molecular organisms. Surprisingly though, for more than a century, researchers have been studying parasitic algae (71) and fungi (72) that do develop a molecular structure within their host cells (73-81). Remarkably, these parasites fuse with their host cells and develop as molecular organisms.

Extant parasitic species potentially evolving into molecular organisms

Before discussing the remarkable prospect that several extant parasitic cellular species might be evolving into molecular organisms, it is important to reiterate that only parasites that have a molecular biology compatible to that of their host cells would be able to fuse with their host cells, develop a molecular structure, and evolve as molecular organisms. This means, for instance, that the numerous parasitic and symbiotic bacteria infecting eukaryal host cells would not be able to evolve into viral species. Indeed, there are many parasitic and endo-symbiotic bacterial species, such as *Carsonella*, *Portiera*, *Tremblaya*, that have much smaller genomes than some viruses [reviewed in (82-84)]. The genome of some of these bacterial species is as small as 160 kb (coding for approximately 30 proteins), which is only a fraction of the coding capacity of the mimivirus genome, which codes for more than 900 proteins. Because the molecular biology of these bacterial species is not compatible to that of their host, however, they need to maintain a cellular membrane and a cellular structure within their eukaryal host cell and, therefore, they are unlikely to evolve into viral species by a fusion mechanism.

As proposed in another article in this series (23), the three cellular domains Archaea, Bacteria, and Eukarya originated from the LUCA lineage by evolving anti-fusion mechanisms, such as cell walls (Bacteria and Eukarya) or membranes with distinct lipid composition (Archaea), which lead to their reproductive isolation and speciation. Due to these anti-fusion devices, in order to evolve into molecular organisms (i.e. AVs), these archaeal, bacterial, and eukaryal parasites would have to develop special fusion-like mechanisms. An evolutionary pathway that could account for the origin of AVs from these non-fusing lineages would be similar to that of the external archaeal parasite *Nanoarchaeum equitans*, which has a very small genome (570 kb) about half the size of the mimivirus genome (85;86). Presumably, this external parasite of the archaeal *Ignicoccus* species lost a large number of its genes because it had access to the host cell's metabolites through a special membrane-to-membrane connection not yet

defined, but probably involving some kind of membrane fusion (85;86). These external parasites would eventually acquire the ability to inject their cytoplasm and genome into the host cell and develop as molecular organisms.

This evolutionary model for the origin of viral lineages infecting Bacteria and Archaea is strongly supported by the current life cycle of the numerous head and tail bacterial and archaeal viruses that inject their genome and other molecules into their host cell. The ancestors of these viruses were probably external parasites with a life cycle similar to that of *N. equitans* and *Vampirococcus* (87). It is also likely that some viruses infecting eukaryal species originated by a similar mechanism. For instance, the viral particles of chloroviruses, which are cell-like structures surrounded by a glycoprotein wall, attach to and degrade the cell wall of their host green alga in order to extrude their content into the host cell (70), leaving their spore-like coat outside the host cell. The life cycle of chloroviruses, which have over 300 protein-coding genes, including remnants of translation machinery, suggests that their ancestors were external parasitic cellular species.

Among Bacteria, *Bdellovibrio bacteriovorus* and many other ubiquitous bacterial parasites that infect other bacterial species [reviewed in (87-89)] are enticing examples of parasitic bacterial lineages that are potentially evolving into molecular organisms. *B. bacteriovorus* penetrates their host cell's outer membrane, hydrolyses the cell wall, and establishes as a parasite within the host periplasm. From within the periplasm, the parasite systematically degrades the host components, ultimately taking over the entire host cell. Similar to viruses, this intracellular parasite replicates its genome repeatedly and then induces formation of internal, new cell-like progeny, which differentiate into transmissible, infectious forms. Then, the parental parasite lyses the host cell and their progeny are released in search of new hosts. Apparently, *B. bacteriovorus* maintains a cellular membrane, and therefore a cellular structure, throughout their life cycle. It is relatively straightforward, however, to envision how *B. bacteriovorus* and other similar parasitic bacterial species could lose their cellular membrane within the host cell and develop as molecular organisms. Also, it would not be surprising to discover that some of the extant bacterial or archaeal parasitic species that infect bacterial or archaeal hosts, respectively, are in fact genuine molecular organisms.

Similar to archaeal and bacterial parasitic species, there are also numerous eukaryal parasitic species, such as the algae and fungi, which might be evolving into molecular organisms. Remarkably, it appears that some of these parasitic species are already true parasitic molecular organisms, although currently they are not recognized as such (73-81). As predicted by the fusion model for the origin of AVs, these fungal and algal parasites: (a) fuse with their host cells, (b) develop a molecular structure, (c) replicate their genome and synthesize their other specific molecules using the host cell resources, and (d) generate

cell-like reproductive forms that differentiate into spores, which start a new life cycle by fusing with other host cells [reviewed in (73-81)].

The idea of a highly complex parasitic cellular species evolving into viruses seems far-fetched. There are, however, several eukaryal parasites and endo-symbionts with genomes smaller than that of large viruses. So, clearly, the potential for profound reductive evolution of eukaryal cells is well supported (54;90-92). In the fusion model for the reductive evolution of parasitic eukaryal species into viral lineages, the parasites would first lose their organelles - such as mitochondria, plastids, and much of the cyto-membrane system - which are readily found in their unique environment, the eukaryal host cell. The nucleus, however, which is tightly coupled with gene expression, would be maintained evolutionarily for long periods before it could be lost through reductive evolution. It would be expected, therefore, that some complex viruses might have remnants of a nuclear membrane and nucleus. Interestingly, using this prediction as a working hypothesis, I found published data that supports this evolutionary model. As shown by electron tomography (93), the poxviral cell-like infectious particles apparently contain a genuine nuclear membrane, which disassembles during the replication of the poviral genome and the morphogenesis of new cell-like progenies and reassembles during the differentiation of these newly-formed daughter cell-like forms into infectious viral particles (see Fig. 1-2f). The lateral bodies of these particles represent a condensed viral cytoplasm (Fig.1-2d).

According to the fusion model for the origin of viruses presented here, a parasitic species that fuses with the host cell and develops a molecular structure is a genuine molecular organism. As proposed in the evolutionary model on the evolutionary diversification of viral species presented in the next section, during early reductive evolutionary stages, the AVs produced cell-like reproductive forms that differentiated into viral particles. Indeed, the viral particles produced by poxviruses, mimivirus, and other large viruses have a cell-like structure and a biochemical composition that are comparable to those of the transmissible reproductive forms of some parasitic cellular species. Many viruses, however, produce non-enveloped viral particles, which do not resemble a cellular structure.

The evolutionary origin of these particles, specifically the evolution of capsid proteins, has been considered central to understanding the origin of viruses and their evolutionary relationships (6;7;50). Because no sequence homology was found between the extant capsid genes and cellular genes, it was proposed that the capsid genes originated from a pool of genes that preceded the origin of cells (6;7;50;94;95). Based on this assumption, the capsid genes have been considered to be viral hallmark genes, or the "self" of viruses, and used as evidence for the origin of viruses from pre-cellular ancestors (6;7;50;94;95). It appears, though, that capsid-like proteins are commonly found among the extant cellular species, but because their

sequence is not homologous to that of viral capsid proteins they were not identified by the phylogenetic programs usually employed in such searches. Surprisingly, some of the best examples of capsid-like proteins are among the most abundant and ubiquitous cellular proteins - the bacterial and archaeal S-layer proteins and the pore-forming proteins across all cellular domains [reviewed in (96-99)].

The structural and functional similarity between the bacterial and archaeal S-layer proteins and some of the viral capsid proteins is striking: (a) they are all rich in β -sheet domains, (b) self-assemble into multiple configurations, including a common viral-like hexagonal symmetry, (c) form pore-like assemblies, (d) have an environmental protective role, and (e) there is little or no sequence homology among S-layer proteins from different species (96;97). This last feature is highly relevant, because it indicates that despite similar structural and functional roles, and likely common evolutionary origin, no strong sequence homology among the S-layer proteins can be detected; therefore, no sequence homology between the viral capsid genes and their presumed cellular gene ancestors should be expected. The proteins forming bacterial microcompartments are another group of proteins that not only resemble the capsid proteins, but remarkably, they assemble into viral particle-like structures (100). These proteins, however, might be genuine viral proteins of symbiotic viruses that co-evolved with their host cells, rather than hypothetical cellular precursors for viral capsid proteins.

Some other intriguing cellular capsid-like proteins are those associated with the spore and polar tube formation in microsporidia. Microsporidia is a very large group of eukaryal intracellular parasites (over 1200 species identified so far) that infect numerous eukaryal species [reviewed in (60)]. The intercellular stage and reproductive mode of these parasites are highly diverse, with some species producing dozens of internal progenies within the parental cell. Just like in viruses, the only form that is viable outside the host cells, and the only infectious form in their life cycle, is their spores, which are analogous to viral particles (see Table 2). Even more remarkable is the fact that to start a new life cycle the microsporidial spores do not enter the host cells, but similar to viral particles of some viruses, such as those of chloroviruses, they inject their genome along with the sporoplasm into the host cell through a polar tube - a very long protein-based structure that is expelled from the spore into the host cell cytoplasm with high velocity. The spore coat and other associated internal structures, including the parental cellular membrane, are left outside the host cells, a phenomenon similar to that observed during the infection of algal cells by chloroviruses (70). Unlike chloroviruses, which develop a molecular structure within the host cell, the sporoplasm of microsporidia is apparently surrounded by a specific microsporidial membrane and maintains a cellular structure during intracellular development. However, it is possible that some microsporidial species do develop a molecular

structure within the host cell, but because of conventional thinking it has been overlooked. Using the fusion model as a working hypothesis, researchers in this field might discover that some microsporidial or related species are genuine molecular organisms. Clearly, microsporidia represent one of the best experimental model systems for studying the potential evolution of eukaryal parasites into molecular organisms.

Evolutionary diversification of viruses

In the fusion model for the origin of viruses, new viral lineages originated from parasitic cellular species throughout the history of life. The vast majority of the extant viral families, however, originated from other viruses by reductive evolution. Some of the major hypothetical steps in the reductive evolutionary diversification of poxvirus-like AVs (referred to as “enveloped viruses”) into viruses that produce transmissible forms analogous to those produced by the other major classes of extant viruses are [reviewed in (16); (Fig. 3)]:

- [i] Origin of viruses that abandoned production of cell-like reproductive forms and used their nucleo-cores as transmissible infectious forms (currently, these viral species are referred to as “unenveloped viruses”). These novel transmissible forms entered the host cells by endocytosis or by injecting their genome and other necessary molecules into the host cells;
- [ii] Origin of plasmid-like viral species from viruses that evolved alternative mechanisms for transmission that did not require production of viral particles, such as endogenous or vertical transmission. In this model, the conjugative plasmids represent a transitional evolutionary stage;
- [iii] Ultimately, an extreme reductive evolutionary pathway led to the origin of viral organisms (not shown in Fig. 3) consisting of a replicating genome (*e.g.* viroids) that doesn’t code for proteins, or to germ-line, endogenous viruses that produced transmissible protein-only viral particles, which can enter new host cells and activate the parental or a related endogenous virus to produce similar transmissible, protein-only viral particles; as described in a separate paper in this series (22), this phenomenon led to the development of the misleading prion hypothesis on the nature of transmissible spongiform encephalopathies.

The evolutionary model above describes the origin of the major groups of extant viruses according to their reproductive and transmission mechanisms (Fig. 3). There are, however, a myriad of intermediate viral lineages spanning these major groups [reviewed in (16)].

The origin and evolution of viral lineages that produce viral particles containing a RNA genome is puzzling (101;102). Considering the popularity of the RNA world

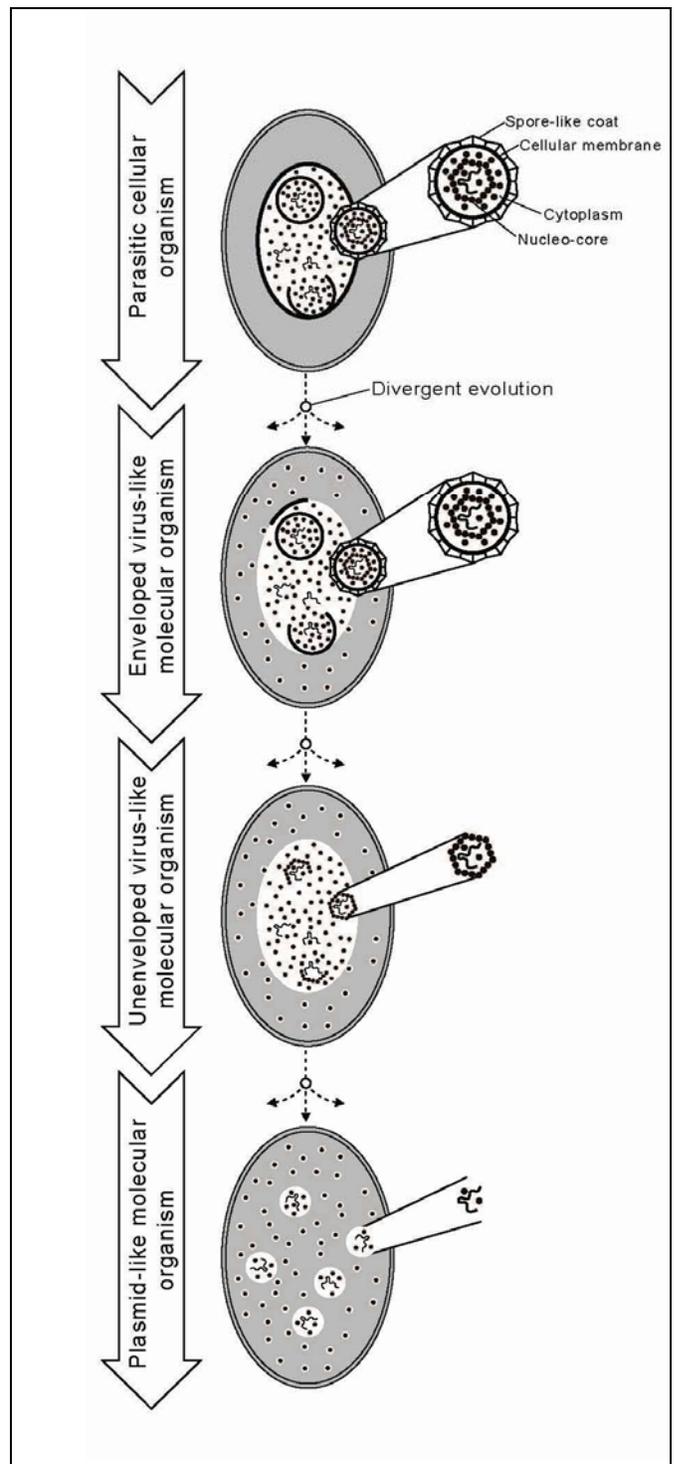


Figure 3. Reductive evolutionary pathway of a parasitic cellular species that lead to the evolution of molecular organisms producing transmissible forms (i.e. viral particles) analogous to those of the major groups of extant viral species.

model for the origin of life (56), it is not surprising that the evolutionary origin of the “RNA viruses” has been associated with this presumed ancestral period (6;7;50).

Box 2. Comparison of the biological properties of parasitic viral species that make infectious viral particles (e.g. Poxviruses) and those of cellular parasitic species that makes infectious spores (e.g. Microsporidia).

	Cellular parasitic species		Viral parasitic species	
	Extracellular	Intracellular	Extracellular	Intracellular
Metabolism ^a	-	+	-	+
Growth ^b	-	+	-	+
Reproduction	-	+	-	+
Infectivity	+	-	+	-

^a Defined as the integrated sum of activities associated with the synthesis of molecules, including carbohydrates, lipids, proteins, and nucleic acids.

^b Defined as an increase in the number of specific molecules.

The current evidence, however, supports an alternative, testable evolutionary model in which “RNA viruses” originated from “DNA viruses.” The vast majority of “RNA viruses” produce virus particles with a positive (i.e. transcript-like) RNA genome, and there is relatively strong evidence that viruses making viral particles containing double-stranded RNAs evolved from viruses producing positive single-stranded RNA containing viral particles (103;104). These findings might be clues to their evolutionary origin. In viruses that have a relatively small number of genes, the RNA transcripts can substitute for a DNA-based genome, which becomes expendable. Likely, the ancestors of “RNA viruses” were viruses with a DNA genome that produced viral particles carrying a DNA genome and RNA transcripts. These viral species discarded their genome after the RNA transcripts were able to replicate within the host cell.

Similar to “RNA viruses”, retroviruses evolved the ability to pack their whole-genome transcripts, which are capped and polyadenylated, within viral particles. However, to be able to insert their genome into the host genome for vertical transmission opportunities, retroviruses use these transcripts as a template for the synthesis of a DNA genome. Considering that only a few amino-acid changes confer a nucleic acid polymerase the property of replicating either DNA or RNA molecules (105;106), or that viral particles can encapsidate either RNA or DNA molecules (107), the evolutionary transitions from “DNA viruses” to “RNA viruses” or the evolutionary transitions among different types of “RNA viruses”, such as those with segmented and non-segmented genomes (108), appear relatively straightforward and open to laboratory testing.

The fusion model predicts that the first AVs originated from the LUCA lineage before the origin of cellular domains (23). It also predicts that thousands of new AV lineages originated from bacterial, archaeal, or eukaryal parasitic cellular species throughout the history of life, and that these evolutionary events are still going on. Clearly, the cellular and viral domains have a long and rich co-

evolutionary history that should be represented in the Tree of Life (TOL). Because of their molecular structure, viruses are prone to intense LGT [reviewed in (15;109)]. It is likely, therefore, that during extended periods of evolution most if not all the genes in some viral lineages have been replaced by genes from other viruses or other sources, possibly several times over. Moreover, owing to a high evolutionary rate of viral genes, it is possible that much of sequence homology among diverging viral lineages has been erased. Evidently, in these cases, the line of descent of extant viruses can not be established by sequence-based phylogenetic analysis. This is also true for some cellular species, particularly parasitic species, which led to open questions about the value and validity of the TOL [reviewed in (110)]. The intent of the TOL, however, is to establish the line of descent among groups of organisms or species, not necessarily the evolutionary relationships among their genes. Certainly, each of the millions of cellular and viral genes has an evolutionary history that can be revealed by a sequence-based phylogenetic tree, but many of these gene-based trees do not represent a TOL that reflects the line of descent among the species. The problem, therefore, might not be with the TOL but with the reductionist approach of generating a TOL based exclusively on sequence-based phylogenetic analysis.

As recently discussed [see (15) and the associated correspondence (111-117)], the issue of classifying viruses as living or non-living entities, which has been disputed ever since viruses were discovered at the turn of the last century (20), is highly relevant in evaluating the merit of including viruses in the TOL. There is no doubt that, by emphasizing the properties of the viral particles, the dogma of viruses as viral particles has played a significant role in portraying viruses as non-living entities [discussed in ref. (20)]. Indeed, it is difficult to consider an apparently inert entity such as the viral particle to be alive, although ultimately, defining an entity as alive, or not, depends on the definition of life. Within the conceptual framework of viruses as molecular organisms, however, their biological

properties are evident (see Box 2 and Fig. 2). And, certainly, it would be easier to accept viruses as organisms and to include them in the TOL in context of the fusion model for the origin of viruses from parasitic cellular species, than within the framework of the pre-cellular or endogenous theories. If the ancestors of viruses were living cellular species as predicted by the fusion model, then it would be difficult to rationalize that viruses, which are their descendents, are not living entities. On the other hand, if the ancestors of viruses were relatively simple replicating genetic elements then it might be easier to argue for classifying them as non-living entities.

As discussed earlier in this section, most, if not all extant viruses originated from other viruses by divergent reductive evolution. Therefore, the issue of including viruses in the TOL, or not, concerns primarily the evolutionary relationship between the ancestral viruses (i.e. the AVs) and the cellular domains. If the AVs originated before cells as predicted by the Pre-cellular Theory, then their diversification leading to the extant viral lineages would generate a TOL that is independent, or parallel, to that of cellular species (Fig. 4, panel I). In this case, the history of life would be represented by two separate line-of-descent TOLs, a cellular TOL and a viral TOL. If we assume that the first cellular organisms descended from viruses, then the cellular branch of the TOL would root at the base of the viral branch of the TOL (Fig. 4, panel II). In contrast, the fusion model for the origin of viruses predicts that thousands of AVs originated from parasitic cellular species, both before and after the origin of three cellular domains, Bacteria, Archaea and Eukarya (23), (Fig. 4, panel III).

The living or non-living status of viruses notwithstanding, the rationale for not including viruses in the TOL is based on the interpretation of viral molecular phylogenetic data in context of the conventional views on the origin of viruses, which leads to profound conceptual disagreements and methodological problems [see (15) and associated correspondence (111-117)].

Interestingly, by using the fusion model as working hypothesis for interpreting the current molecular phylogenetic data, many of these problems and methodological constraints are resolved. For example, the fusion model predicts that viruses are polyphyletic (Fig. 4, panel III), that they did not originate from pre-cellular ancestors, and that by originating from parasitic cellular species by reductive evolution, many of the extant complex viruses are expected to contain cellular remnants and genes that are closely related to those of their hosts. This last aspect is highly relevant for the interpretation of the current data in the field of viral molecular phylogenetics, because many of the viral genes that currently are regarded as cellular genes acquired by viruses from their host by LGT [reviewed in (15)] might represent genuine viral genes that descended from their parasitic cellular ancestors.

There is no doubt that owing to rapid evolution and intense LGT, much of the vertical sequence-based phylogenetic signal of the early viral lineages has probably eroded. Considering also that, as predicted by the fusion model, thousands of new viral lineages originated from cellular species throughout the history of life, it would be very difficult, if not impossible, to root many of the extant viral lineages to the TOL in a conventional way.

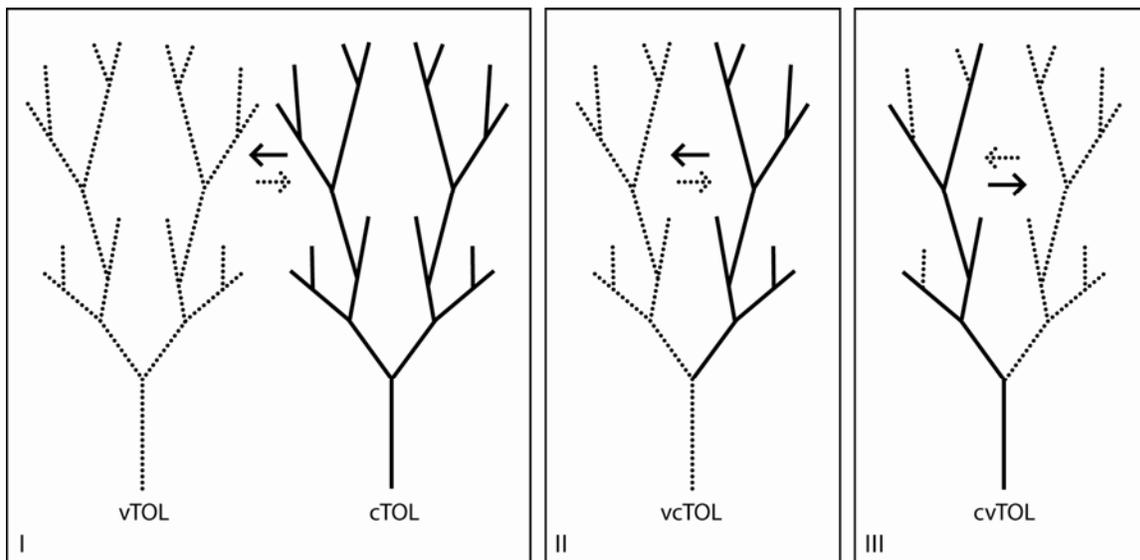


Figure 4. Schematic representation of TOL based on the current views on the origin and evolution of viruses (see text for discussion). For simplicity, the trees are drawn with a bifurcating topology and with just a few illustrative branches. The cellular lineages are shown as continuous lines and the viral lineages as dotted lines. The arrows indicate exchanges of genetic material between viral and cellular lineages. Panel I represents two independent TOLs, a viral TOL (vTOL) and a cellular TOL (cTOL). Panel II represents a viral-first/cellular TOL (vcTOL), in which the ancestral cellular lineages are rooted at the base of an ancestral viral branch. Panel III represents a cellular-first/viral TOL (cvTOL) in which multiple viral lineages originated from cellular lineages.

However, these methodological difficulties should not nullify viruses the right to be included in the TOL. A possible solution to this methodological problem is to envision that the TOL is embedded in viral shell (23).

Summary and perspective

The dogma of viruses as viral particles misrepresents the nature of viruses and has set viruses apart from mainstream biological and evolutionary paradigms. Moreover, as discussed here and in the other articles in this series, this dogma has constrained full potential for progress in many bio-medical fields (22;23). Clearly, the dogma of viruses as virus particles can no longer productively integrate the rapid advances in our knowledge about viruses and cells. It is not surprising, therefore, that alternative views are emerging (4;21).

Similar to other parasitic species, viruses pass through multiple developmental stages in their life cycle and, therefore, an integrated sum of all these stages defines them. It is within their host cell environment, however, that viruses express the essence of all organisms – synthesis of their specific molecules, growth, and reproduction. To identify viruses during the intracellular stage when their molecules are dispersed within the host cell, this paper proposed the concept of molecular structure and labeled viruses as molecular organisms. These concepts set the foundation for including diverse biological entities, such as plasmids, endogenous viruses, transposable elements, viroids, phages, and viruses, within the same domain of biological organization - the viral, or molecular, domain.

The fusion model proposes that viruses originated from parasitic cellular organisms that fused with their host cell. Among modern viruses, the life cycle of poxviruses and other complex viruses provides compelling evidence for the fusion model. Some of these viruses have indisputable cellular remnants that are best explained by an evolutionary origin from cellular ancestors.

One of the most remarkable implications of the fusion model is that new viral lineages originated from parasitic cellular species throughout life's history and that this process might still be active. Surprisingly, as predicted by the fusion model, several extant parasitic fungal and algal species fuse with their host cells and develop as molecular organisms. In addition to these parasitic eukaryal species, there are also several bacterial and archaeal parasites that might be evolving into molecular organisms, albeit by special fusion-like mechanisms.

The fusion model for the origin of viruses supports the notion that they are alive, and that they should be included in the TOL. However, because thousands of new viral lineages originated from parasitic cellular species throughout life's history, and because of their rapid evolution and intense LGT, it is difficult to place the viral lineages on TOL. Although, some of the lineages that originated more recently might be rooted to TOL in a conventional way most probably cannot. A potential

solution would be to envision that many of the branches of the TOL are embedded in a viral shell (23).

In perspective, the present model on origin and evolution of viruses as molecular organisms has immediate research applications (22) and it could radically change our view about the history of life (23).

Acknowledgments:

I thank Francisco Ayala, William Bains, Rebecca Bandea, Philip Bell, Carolyn Black, Allan Campbell, Jean-Michel Claverie, Cristina Crisco, Esteban Domingo, Lynn Enquist, Patrick Forterre, Mark Gibbs, Johann Gogarten, Eugene Koonin, Norman Pace, David Penny, Anthony Poole, David Prangishvili, Frank Ryan, and Luis Villarreal for reviewing this or previous versions of this paper. I also thank Jennifer Hulsey for help with the illustrations.

The findings and conclusions in this report are those of the author and do not necessarily represent the views or the opinions of U.S. Department of Health and Human Services and Centers for Disease Control and Prevention.

References

1. Edwards,R.A., and Rohwer,F. 2005. Viral metagenomics. *Nat. Rev. Microbiol.* **3**:504-510.
2. Suttle,C.A. 2007. Marine viruses--major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**:801-812.
3. Angly,F.E., Felts,B., Breitbart,M., Salamon,P., Edwards,R.A., Carlson,C., Chan,A.M., Haynes,M., Kelley,S., Liu,H. et al 2006. The marine viromes of four oceanic regions. *PLoS. Biol* **4**:e368.
4. Claverie,J.M. 2006. Viruses take center stage in cellular evolution. *Genome Biol.* **7**:110.
5. Filee,J., Forterre,P., and Laurent,J. 2003. The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies. *Res. Microbiol.* **154**:237-243.
6. Forterre,P. 2006. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* **117**:5-16.
7. Koonin,E.V., Senkevich,T.G., and Dolja,V.V. 2006. The ancient Virus World and evolution of cells. *Biol. Direct.* **1**:29.
8. Domingo,E., C.Parrish, and J.J.Holland 2008. *Origin and Evolution of Viruses, 2nd edition.* Elsevier. Oxford.
9. Villarreal,L.P. 2004. *Viruses and the Evolution of Life.* ASM Press. Washington.
10. Gibbs,A.J., Calisher,C.H., and García-Arenal,F. 1995. *Molecular Basis of Virus Evolution.* Cambridge University Press. Cambridge.
11. Morse,S.S. 1994. *The Evolutionary Biology of Viruses.* Raven Press. New York.
12. Ryan,F.P. 2007. Viruses as symbionts. *Symbiosis* **44**:11-12.
13. Woese,C.R. 2000. Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. U. S. A* **97**:8392-8396.
14. Doolittle,W.F. 1999. Phylogenetic classification and the universal tree. *Science* **284**:2124-2129.
15. Moreira,D., and Lopez-Garcia,P. 2009. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* **7**:306-311.

16. Knipe,D.M., P.M.Howley, D.E.Griffin, R.A.Lamb, M.A.Martin, B.Roizman, and S.Straus 2007. *Fields Virology*. Lippincott Williams & Wilkins.
17. Ghabrial,S.A. 1998. Origin, adaptation and evolutionary pathways of fungal viruses. *Virus Genes* **16**:119-131.
18. Lucas,W.J. 2006. Plant viral movement proteins: agents for cell-to-cell trafficking of viral genomes. *Virology* **344**:169-184.
19. Gifford,R., and Tristem,M. 2003. The evolution, distribution and diversity of endogenous retroviruses. *Virus Genes* **26**:291-315.
20. Bandea,C.I. 1983. A new theory on the origin and the nature of viruses. *J. Theor. Biol.* **105**:591-602.
21. Raoult,D., and Forterre,P. 2008. Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* **6**:315-319.
22. Bandea,C.I. 2009. Endogenous viral etiology of prion diseases. *Nature Precedings* <http://precedings.nature.com/>.
23. Bandea,C.I. 2009. A unifying scenario on the origin of and evolution of cellular and viral domains. *Nature Precedings* <http://precedings.nature.com/>.
24. Hughes,S.S. 1977. *The virus. A History of the Concept*. Heinemann Educational. London.
25. Waterson,A.P., and L.Wilkinson. 1978. *An Introduction to the History of Virology*. Cambridge University Press. London.
26. Norrby,E. 2008. Nobel Prizes and the emerging virus concept. *Arch. Virol.* **153**:1109-1123.
27. Stanley,W.M. 1935. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus. *Science* **81**:644-645.
28. Bowden,F.C., and N.W.Pirie 1937. The isolation and some properties of liquid crystalline substances from solanaceous plants infected with three strains of tobacco mosaic virus. *Proc. R. Soc.* **123**:274-320.
29. Watson,J.D. 1976. *Molecular Biology of the Gene*. Benjamin-Cummings. Menlo Park.
30. Rowson,K.E.K., T.A.L.Rees, and B.W.J.Mahy 1981. *A Dictionary of Virology*. Blackwell Scientific. Oxford.
31. Joklik,W.K., H.P.Willett, D.B.Amos, and C.M.Wifert 1992. *Zinsser Microbiology*. Appleton and Lange. Norwalk.
32. Alberts,B., D.Bray, K.Hopkin, A.Johnson, J.Lewis, M.K.Roberts, and P.Walter 2004. *Essential Cell Biology*. Garland Science. New York and London.
33. Lwoff,A. 1957. The concept of virus. *J. Gen. Microbiol.* **17**:239-253.
34. Woese,C.R., Kandler,O., and Wheelis,M.L. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U. S. A* **87**:4576-4579.
35. Woese,C. 1998. The universal ancestor. *Proc. Natl. Acad. Sci. U. S. A* **95**:6854-6859.
36. Penny,D., and Poole,A. 1999. The nature of the last universal common ancestor. *Curr. Opin. Genet. Dev.* **9**:672-677.
37. Doolittle,W.F. 2000. The nature of the universal ancestor and the evolution of the proteome. *Curr. Opin. Struct. Biol* **10**:355-358.
38. Forterre,P., and Philippe,H. 1999. Where is the root of the universal tree of life? *Bioessays* **21**:871-879.
39. Luria,S.E., and J.E.Darnell 1967. *General Virology*. Wiley. New-York.
40. Burnet,F.M. 1955. *Principles of Animal Virology*. Academic Press. New-York.
41. Burnet,F.M. 1945. *Virus as Organism*. Harvard University Press. Cambridge, Massachusetts.
42. Matthews,R.E.F. 1983. The origin of viruses from cells. *Int. Rev. Cyto. (Suppl.)* **15**:245-280.
43. Moss,B. 2006. Poxvirus entry and membrane fusion. *Virology* **344**:48-54.
44. Novoa,R.R., Calderita,G., Arranz,R., Fontana,J., Granzow,H., and Risco,C. 2005. Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol. Cell* **97**:147-172.
45. Netherton,C., Moffat,K., Brooks,E., and Wileman,T. 2007. A guide to viral inclusions, membrane rearrangements, factories, and viroplasm produced during virus replication. *Adv. Virus Res.* **70**:101-182.
46. Neiman,A.M. 2005. Ascospore formation in the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **69**:565-584.
47. Dawes,I.W. 1981. Sporulation in evolution. *Symp. Soc. Gen. Microbiol.* **30**:85-130.
48. Granados,R.R., and Roberts,D.W. 1970. Electron microscopy of a poxlike virus infecting an invertebrate host. *Virology* **40**:230-243.
49. Dale,S., and Pogo,B.G.T. 1981. *Biology of Poxviruses*. (Virol.Monogr.18), Springer-Verlag. Wien.
50. Koonin,E.V., and Dolja,V.V. 2006. Evolution of complexity in the viral world: the dawn of a new vision. *Virus Res.* **117**:1-4.
51. Wernegreen,J.J. 2005. For better or worse: genomic consequences of intracellular mutualism and parasitism. *Curr. Opin. Genet. Dev.* **15**:572-583.
52. Cavalier-Smith,T. 2003. Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philos. Trans. R. Soc. Lond B Biol. Sci.* **358**:109-133.
53. Andersson,S.G., and Kurland,C.G. 1998. Reductive evolution of resident genomes. *Trends Microbiol.* **6**:263-268.
54. Keeling,P.J., and Slamovits,C.H. 2005. Causes and effects of nuclear genome reduction. *Curr. Opin. Genet. Dev.* **15**:601-608.
55. Slamovits,C.H., Fast,N.M., Law,J.S., and Keeling,P.J. 2004. Genome compaction and stability in microsporidian intracellular parasites. *Curr. Biol.* **14**:891-896.
56. Gilbert,W. 1986. Origin of life, the RNA world. *Nature* **319**:618.
57. Holmes,E.C. 2003. Molecular clocks and the puzzle of RNA virus origins. *J. Virol.* **77**:3893-3897.
58. Belshaw,R., Pybus,O.G., and Rambaut,A. 2007. The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res.* **17**:1496-1504.
59. Duffy,S., Shackelton,L.A., and Holmes,E.C. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**:267-276.
60. Keeling,P.J., and Fast,N.M. 2002. Microsporidia: biology and evolution of highly reduced intracellular parasites. *Annu. Rev. Microbiol.* **56**:93-116.
61. Temin,H.M. 1980. Origin of retroviruses from cellular moveable genetic elements. *Cell* **21**:599-600.

62. Katzourakis,A., Rambaut,A., and Pybus,O.G. 2005. The evolutionary dynamics of endogenous retroviruses. *Trends Microbiol.* **13**:463-468.
63. Ribet,D., Harper,F., Dupressoir,A., Dewannieux,M., Pierron,G., and Heidmann,T. 2008. An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome Res.* **18**:597-609.
64. Weiss,R.A. 2006. The discovery of endogenous retroviruses. *Retrovirology.* **3**:67.
65. Lower,R., Lower,J., and Kurth,R. 1996. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc. Natl. Acad. Sci. U. S. A* **93**:5177-5184.
66. Suzan-Monti,M., La,S.B., and Raoult,D. 2006. Genomic and evolutionary aspects of Mimivirus. *Virus Res.* **117**:145-155.
67. Raoult,D., Audic,S., Robert,C., Abergel,C., Renesto,P., Ogata,H., La,S.B., Suzan,M., and Claverie,J.M. 2004. The 1.2-megabase genome sequence of Mimivirus. *Science* **306**:1344-1350.
68. Ghedin,E., and Claverie,J.M. 2005. Mimivirus relatives in the Sargasso sea. *Virol. J.* **2**:62.
69. Suhre,K., Audic,S., and Claverie,J.M. 2005. Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proc. Natl. Acad. Sci. U. S. A* **102**:14689-14693.
70. Yamada,T., Onimatsu,H., and Van Etten,J.L. 2006. Chlorella viruses. *Adv. Virus Res.* **66**:293-336.
71. Rosenvinge,L.K. 1888. Sur la formation des pores secondaires chez Polysiphonia. *Bot. Tidskr.* **17**:10-19.
72. Burgeff,H. 1924. Untersuchungen über Sexualität und Parasitismus bei Mucorineen. *Bot. Abh.* **4**:1-135.
73. Wetherbee,R. 1980. Pit Connections and Translocation in Red Algae. *Science* **209**:423.
74. Schultze,K., Schimek,C., Wostemeyer,J., and Burmester,A. 2005. Sexuality and parasitism share common regulatory pathways in the fungus *Parasitella parasitica*. *Gene* **348**:33-44.
75. Kellner,M., Burmester,A., Wostemeyer,A., and Wostemeyer,J. 1993. Transfer of genetic information from the mycoparasite *Parasitella parasitica* to its host *Absidia glauca*. *Curr. Genet.* **23**:334-337.
76. Goff,L.J., and Coleman,A.W. 1984. Transfer of nuclei from a parasite to its host. *Proc. Natl. Acad. Sci. U. S. A* **81**:5420-5424.
77. Goff,L.J., and Coleman,A.W. 1995. Fate of Parasite and Host Organelle DNA during Cellular Transformation of Red Algae by Their Parasites. *Plant Cell* **7**:1899-1911.
78. Goff,L.J., Ashen,J., and Moon,D.A. 1997. The evolution of parasites from their hosts: A case study in the parasitic red algae. *Evolution* **51**:1068-1078.
79. Bauer,R., M.Lutz, and F.Oberwinkler 2004. *Tuberculina*-rusts: a unique basidiomycetous interfungal cellular interaction with horizontal nuclear transfer. *Mycologia* **96**:960-967.
80. Simon,U.K., Bauer,R., Rioux,D., Simard,M., and Oberwinkler,F. 2005. The vegetative life-cycle of the clover pathogen *Cymadothea trifolii* as revealed by transmission electron microscopy. *Mycol. Res.* **109**:764-778.
81. Wetherbee,R., and West,J.A. 1976. Unique Golgi apparatus and vesicle formation in a red alga. *Nature* **259**:566-567.
82. Baumann,P. 2005. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* **59**:155-189.
83. Fenn,K., and Blaxter,M. 2006. Wolbachia genomes: revealing the biology of parasitism and mutualism. *Trends Parasitol.* **22**:60-65.
84. Nakabachi,A., Yamashita,A., Toh,H., Ishikawa,H., Dunbar,H.E., Moran,N.A., and Hattori,M. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**:267.
85. Jahn,U., Gallenberger,M., Paper,W., Junglas,B., Eisenreich,W., Stetter,K.O., Rachel,R., and Huber,H. 2008. Nanoarchaeum equitans and Ignicoccus hospitalis: new insights into a unique, intimate association of two archaea. *J. Bacteriol.* **190**:1743-1750.
86. Waters,E., Hohn,M.J., Ahel,I., Graham,D.E., Adams,M.D., Barnstead,M., Beeson,K.Y., Bibbs,L., Bolanos,R., Keller,M. et al 2003. The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. *Proc. Natl. Acad. Sci. U. S. A* **100**:12984-12988.
87. Guerrero,R., Pedros-Alio,C., Esteve,I., Mas,J., Chase,D., and Margulis,L. 1986. Predatory prokaryotes: predation and primary consumption evolved in bacteria. *Proc. Natl. Acad. Sci. U. S. A* **83**:2138-2142.
88. Rendulic,S., Jagtap,P., Rosinus,A., Eppinger,M., Baar,C., Lanz,C., Keller,H., Lambert,C., Evans,K.J., Goemann,A. et al 2004. A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. *Science* **303**:689-692.
89. Lambert,C., Morehouse,K.A., Chang,C.Y., and Sockett,R.E. 2006. *Bdellovibrio*: growth and development during the predatory cycle. *Curr. Opin. Microbiol.* **9**:639-644.
90. Archibald,J.M. 2007. Nucleomorph genomes: structure, function, origin and evolution. *Bioessays* **29**:392-402.
91. Douglas,S., Zauner,S., Fraunholz,M., Beaton,M., Penny,S., Deng,L.T., Wu,X., Reith,M., Cavalier-Smith,T., and Maier,U.G. 2001. The highly reduced genome of an enslaved algal nucleus. *Nature* **410**:1091-1096.
92. Gilson,P.R., Su,V., Slamovits,C.H., Reith,M.E., Keeling,P.J., and McFadden,G.I. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. *Proc. Natl. Acad. Sci. U. S. A* **103**:9566-9571.
93. Cyrklaff,M., Risco,C., Fernandez,J.J., Jimenez,M.V., Esteban,M., Baumeister,W., and Carrascosa,J.L. 2005. Cryo-electron tomography of vaccinia virus. *Proc. Natl. Acad. Sci. U. S. A* **102**:2772-2777.
94. Krupovic,M., and Bamford,D.H. 2008. Virus evolution: how far does the double beta-barrel viral lineage extend? *Nat. Rev. Microbiol.* **6**:941-948.
95. Bamford,D.H., Grimes,J.M., and Stuart,D.I. 2005. What does structure tell us about virus evolution? *Curr. Opin. Struct. Biol* **15**:655-663.
96. Sara,M., and Sleytr,U.B. 2000. S-Layer proteins. *J. Bacteriol.* **182**:859-868.
97. Sleytr,U.B., Egelseer,E.M., Ilk,N., Pum,D., and Schuster,B. 2007. S-Layers as a basic building block in a molecular construction kit. *FEBS J* **274**:323-334.
98. Tilley,S.J., and Saibil,H.R. 2006. The mechanism of pore formation by bacterial toxins. *Curr. Opin. Struct. Biol* **16**:230-236.

99. Menestrina,G., Dalla,S.M., Comai,M., Coraiola,M., Viero,G., Werner,S., Colin,D.A., Monteil,H., and Prevost,G. 2003. Ion channels and bacterial infection: the case of beta-barrel pore-forming protein toxins of *Staphylococcus aureus*. *FEBS Lett.* **552**:54-60.
100. Yeates,T.O., Tsai,Y., Tanaka,S., Sawaya,M.R., and Kerfeld,C.A. 2007. Self-assembly in the carboxysome: a viral capsid-like protein shell in bacterial cells. *Biochem. Soc. Trans.* **35**:508-511.
101. Reaney,D.C. 1982. The evolution of RNA viruses. *Annu. Rev. Microbiol.* **36**:47-73.
102. Domingo,E., Escarmis,C., Sevilla,N., Moya,A., Elena,S.F., Quer,J., Novella,I.S., and Holland,J.J. 1996. Basic concepts in RNA virus evolution. *FASEB J.* **10**:859-864.
103. Gibbs,M.J., Koga,R., Moriyama,H., Pfeiffer,P., and Fukuhara,T. 2000. Phylogenetic analysis of some large double-stranded RNA replicons from plants suggests they evolved from a defective single-stranded RNA virus. *J. Gen. Virol.* **81**:227-233.
104. Ahlquist,P. 2006. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat. Rev. Microbiol.* **4**:371-382.
105. Sousa,R., and Padilla,R. 1995. A mutant T7 RNA polymerase as a DNA polymerase. *EMBO J* **14**:4609-4621.
106. Ong,J.L., Loakes,D., Jaroslowski,S., Too,K., and Holliger,P. 2006. Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide. *J Mol. Biol* **361**:537-550.
107. Houzet,L., Morichaud,Z., Didierlaurent,L., Muriaux,D., Darlix,J.L., and Mougel,M. 2008. Nucleocapsid mutations turn HIV-1 into a DNA-containing virus. *Nucleic Acids Res.* **36**:2311-2319.
108. Zaccomer,B., Haenni,A.L., and Macaya,G. 1995. The remarkable variety of plant RNA virus genomes. *J Gen. Virol.* **76 (Pt 2)**:231-247.
109. Gogarten,J.P., and Townsend,J.P. 2005. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3**:679-687.
110. Koonin,E.V. 2009. Darwinian evolution in the light of genomics. *Nucleic Acids Res.* **37**:1011-1034.
111. Claverie,J.M., and Ogata,H. 2009. Ten good reasons not to exclude giruses from the evolutionary picture. *Nat. Rev. Microbiol.* **7**:615.
112. Koonin,E.V., Senkevich,T.G., and Dolja,V.V. 2009. Compelling reasons why viruses are relevant for the origin of cells. *Nat. Rev. Microbiol.* **7**:615.
113. Ludmir,E.B., and Enquist,L.W. 2009. Viral genomes are part of the phylogenetic tree of life. *Nat. Rev. Microbiol.* **7**:615.
114. Hegde,N.R., Maddur,M.S., Kaveri,S.V., and Bayry,J. 2009. Reasons to include viruses in the tree of life. *Nat. Rev. Microbiol.* **7**:615.
115. Navas-Castillo,J. 2009. Six comments on the ten reasons for the demotion of viruses. *Nat. Rev. Microbiol.* **7**:615.
116. Raoult,D. 2009. There is no such thing as a tree of life (and of course viruses are out!). *Nat. Rev. Microbiol.* **7**:615.
117. Lopez-Garcia,P., and Moreira,D. 2009. Yet viruses cannot be included in the tree of life. *Nat Rev Micro* **7**:615-617.