OPINION

Ten reasons to exclude viruses from the tree of life

David Moreira and Purificación López-García

Abstract | When viruses were discovered, they were accepted as missing links between the inert world and living organisms. However, this idea was soon abandoned as information about their molecular parasitic nature accumulated. Recently, the notion that viruses are living organisms that have had a role in the evolution of some essential features of cells has experienced a renaissance owing to the discovery of unusually large and complex viruses that possess typical cellular genes. Here, we contend that there is strong evidence against the notion that viruses are alive and represent ancient lineages of the tree of life.

In 1929, when the first hypotheses to explain the origins of life through natural processes were being formulated, Muller proposed that the first living organism was a primitive gene[1]. He thereby positioned himself in the debate between 'nucleocentric' (genetic material first) and 'cytoplasmist' (metabolism first) schools of thought on the origins of life[2]. Advances in the characterization of viruses, which had been discovered only a few decades previously, led to the hypothesis that viruses were autonomous hereditary units that many authors thought could replicate independently. On the basis of their simplicity, Muller proposed that viruses could be present-day functional analogues of primitive genes[1]. At the same time, authors including Simon, hypothesized that "the majority of viruses are living matter"[3], an idea that was rapidly embraced by the emerging field of virology.

The central role of viruses in the origins of life was cemented with Haldane's famous essay, The origin of life, published in 1929, in which he not only considered viruses as living entities, but also asserted that "life may have remained in the virus stage for many millions of years before a suitable assemblage of elementary units was brought together in the first cell."[4] In other words, he said that viruses were the missing link between the non-living and the first cells. In fact, the syllogism 'smallest = virus, smallest = first, so that virus = first' was easy to understand and many scientists enthusiastically adhered to it[5]. Although Haldane later changed his mind and agreed with Oparin[6] that viruses were irrelevant in the origins of life and that the decisive step was the formation of a cell, his initial opinion remained influential.

The discovery in 1944 that DNA encodes genetic information[7] was fatal for the cytoplasmist theory and created a 'DNA-centrist' definition of life that held 'information' and 'evolution' as central tenets. Indeed, even today many scientists accept that if something evolves it is alive. Under this premise, the idea that viruses might be extremely simple ancestral life forms seems logical. The nucleocentric origin-of-life theory became 'virocentric'[2]. As we discuss in this Opinion article, most of the data on viral genetics and biochemistry that have accumulated over the past 50 years strongly oppose this view (TABLE 1) and support the idea that viruses are not alive. Futhermore, that viruses are not alive was officially acknowledged by the International Committee on Taxonomy of Viruses[8] in 2000 and is still held by most virologists[9].

The virocentric origin of life theory has been revived recently for several reasons. First, viruses that infect distantly related organisms may share similar architectures, for instance phages from bacteria and halophilic archaea. In addition, capsid proteins from unrelated viruses can contain similar folds. This has been suggested as evidence of ancient viral lineages that co-evolved with cells and predate the divergence of the three domains of cellular organisms (Bacteria, Archaea and Eucarya)[10]. Second, viruses have been hypothesized as having been instrumental in the evolution of important features in extant cells, including DNA itself, which viruses might have 'invented' in a hypothetical 'RNA world' in which cells had RNA genomes[11]. Third, the discovery of the giant Mimivirus in 2003 (REF. 12) and the sequencing of its 1.2 Mbp-sized genome revealed that an unprecedented number of genes involved in transcription and translation are homologous to those found in cells[13]. Moreover, the phylogenetic analysis of a few of the Mimivirus genes suggested that the Mimivirus might represent a new branch in the tree of life — namely, a fourth domain of life[13], a fundamental claim for the renaissance of the 'virus as living being' hypothesis.

Some of the considerations on which this assumption is based, however, have been convincingly contested in the past, and a whole body of biological evidence exists refuting the contention that viruses are living entities and have a place in the tree of life. Here, we examine this evidence and argue, first, that viruses are not true living entities. In the subsequent nine points, we further argue that, independently of any consideration about their living or non-living nature, attempts to place viruses in the tree of life are artificial, based on the study of viral genes that were acquired from cellular hosts.

Viruses are not alive

Being composed of molecules that are found in cells (nucleic acids, proteins, lipids and complex sugars) and having the capacity to evolve, viruses are often said to be alive. Defining life has never been an easy task; many different definitions, some mutually exclusive, have been proposed[14] (TIMELINE 1), but any valid definition should be able to exclude entities that are not alive. The more reductionist definitions, which are the ones that might apply to viruses, can be divided into two main groups.

The first, in line with cytoplasmic or metabolism-first schools, relies on the self-organization and self-maintenance properties of living beings. Obviously, as viruses lack any form of energy and carbon metabolism, they are not alive according to this type of definition. The second is based on the properties of self-replication and evolution. Only a definition of this sort, entirely and exclusively based on the ability to evolve, could apply to viruses. However, this definition is not robust as it includes, for example, computer viruses, which can be designed to produce copies of themselves with slight changes in their code (that is, mutate). This makes them evolving entities that undergo selection (by their differential capacity to infect new computers and spread, and by their resistance to antivirus software), but few, if any, people would say they are alive.

Even if we accept such an imperfect definition of life, viruses would still be excluded because of a fundamental peculiarity of viral existence that is often overlooked: viruses

Table 1 | **Comparison of cellular and viral traits**

Trait	Cells	Viruses
Information content	Yes	Yes
Self-maintenance	Yes	No
Self-replication	Yes	No
Evolution	Yes	By cells
Common ancestry	Yes	No
Structural historical continuity	Yes	No
Genes involved in carbon metabolism	Yes	Cellular origin
Genes involved in energy metabolism	Yes	Cellular origin
Genes involved in protein synthesis	Yes	Cellular origin

neither replicate nor evolve, they are evolved by cells. Even if some viruses encode their own polymerases, some of them error-prone, their expression and function require the cell machinery so that, in practice, viruses are evolved by cells — no cells, no viral evolution. This applies to other selfish genetic elements and even to cellular genes. Analogously, human technology does not evolve by itself but is evolved by humans. Alexander and Bridges eloquently made this crucial distinction eighty years ago by declaring that viruses are "produced but not self-reproduced"[15]. Along with this line of thought, we can say that viruses are "not living, but lived entities"[16]. In fact, as perfect molecular parasites, viruses depend completely on the metabolic machinery of cells, not only for their reproduction but also for their evolution. Thus, in the absence of cells, viruses are nothing but inanimate complex organic matter.

Imagine a sterile planet with all the physico-chemical requirements that are needed to host life. If we inseminate it with populations of all the viral lineages that are known on Earth, it is evident that nothing will happen except the progressive decay of the molecules composing those viruses. If instead of viruses we inseminate such a planet with populations of, for example, all known bacterial species, part of this bacterial life would most likely self-maintain, reproduce and evolve, colonizing the planet in a stable way. It could be argued that obligate intracellular bacteria are akin to viruses in that they require a host cell to propagate. However, this is not true for two reasons. First, intracellular bacteria or obligate parasitic bacteria maintain some kind of carbon and energy metabolism and, in most cases, given the appropriate complex culture media, they can be grown under laboratory conditions. Second, compelling evidence shows that these bacteria have lost many of their metabolic functions as a result of reductive evolution from more complex, free-living ancestors; there is no evidence to suggest any viruses ever had such functions.

We believe that considering viruses alive or not is not just a matter of opinion, contrary to a commonly held view, but rather is a matter of inference and logic starting from any given definition of life. Of course, one could decide not to define life but, in that case, viruses can neither be regarded as living nor as non-living; otherwise an implicit definition of life is being used. However, independently of the debate about whether or not viruses are alive, there are other distinct and pragmatic reasons that prevent the inclusion of viruses in the tree of life.

Viruses are polyphyletic

A phylogenetic tree is a conceptual representation of evolutionary relationships among taxa. For more than a century, it has been recognized that a phylogenetic tree can only be inferred by studying characteristics that have been inherited from the last common ancestor of the taxa — that is, proper phylogenetic analysis should only be based on homology[17]. This makes it impossible to include viruses in the tree of life: although a few genes are shared between some specific viral lineages and their host cells (see below), viruses as a whole do not share homologous characteristics with cells. Moreover, not a single gene is shared by all viruses or viral lineages. Therefore, from a molecular phylogeny perspective, viruses as a whole can be compared neither among themselves, nor with the cellular organisms that populate the tree of life.

Members of the different viral families are composed of different nucleic acids and capsid constituents and have different gene contents. This strongly suggests that viruses have various evolutionary origins — that is, they are polyphyletic[18]. By contrast, overwhelming evidence shows that all cellular life has a single, common origin[19,20]. Therefore, whereas the inference of a tree for all cellular species is a sensible scientific task, it is an unattainable one in the case of viruses. The absence of common characteristics among viral families, and between viruses and cells, makes any taxonomic scheme that aims to embrace all of these entities artificial and contrary to proper taxonomic practice.

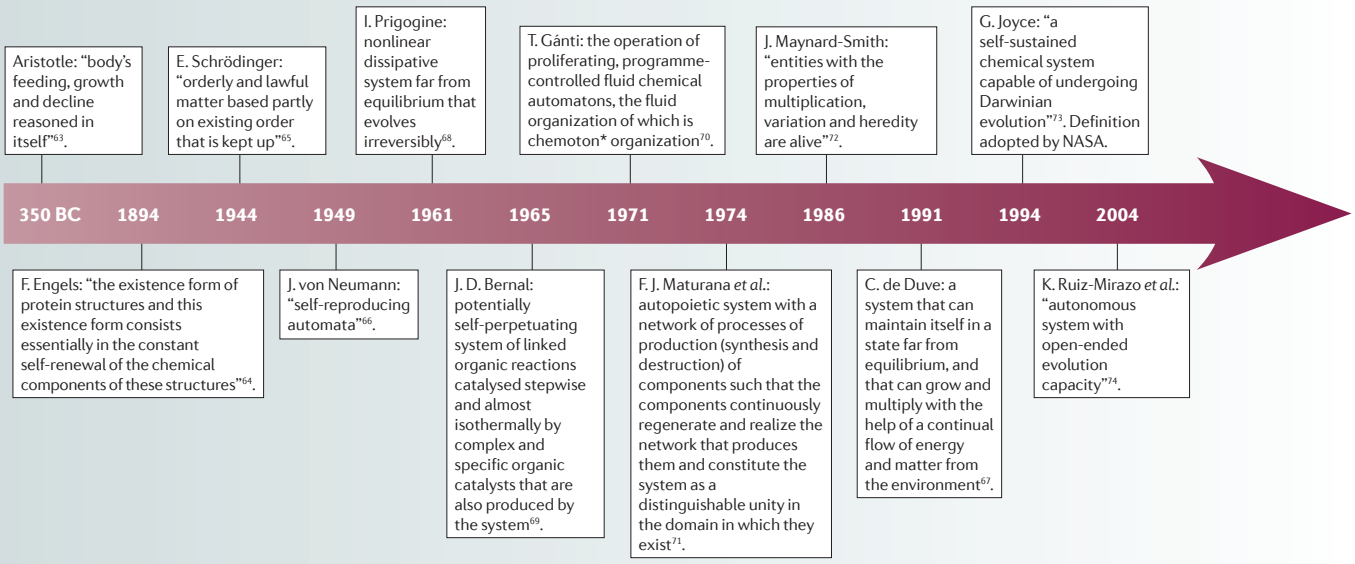
In the early times of viral research, when the nature of viruses was not yet fully understood, such taxonomic schemes were suggested. In 1928 Alexander and Bridges proposed a division of organisms between Ultrabionta (viruses) and Cytobionta (cellular organisms)[15]. However this proposal was abandoned over the next decades when the disconnection between viruses and cells was firmly established. This forgotten scheme has recently been resurrected inadvertently by Raoult and Forterre[21], who proposed a division between 'capsid-encoding organisms' (CEOs, the equivalent of Ultrabionta) and 'ribosome-encoding organisms' (REOs, the equivalent of Cytobionta).

Viruses are not only polyphyletic, but, as an ill-defined group, they are not clearly delineated from other selfish genetic elements, such as plasmids[18]. Many viruses share genes with plasmids (significantly more than with cells), indicating that they have a direct evolutionary connection with these elements. Therefore, if viruses were included in a universal tree, many plasmids, not to mention other genetic elements, should be added too. Raoult and Forterre[21] classify these genetic elements, including plasmids, transposons, viroids, virusoids and RNA satellites, as 'orphan replicons' that do not deserve the title of organisms but that could be included in the tree of life. However, if a tree of life contains elements that are not organisms, is it a tree of life or just a tree of genes from multiple origins? Gene trees may or may not reflect organismal phylogenies but, conceptually, they are clearly different things.

There are no ancestral viral lineages

There is no single gene that is shared by all viruses. Nevertheless, it has been claimed that structural motifs that are shared by capsid proteins from distant viral lineages — for example, enterobacterial phages and eukaryotic adenoviruses — provide evidence, despite their extreme divergence in primary sequence, for a common ancient origin that predates the last common ancestor of cellular organisms: the cenancestor[18,22,23]. Taking into account what is known about viral structure and genome evolution, there are two alternative possibilities to explain the presence of common protein motifs in distinct viral lineages.

Timeline 1 | **Definitions of life or living beings**



*Chemoton is the fundamental unit model of living systems consisting of three functionally dependent autocatalytic subsystems: a metabolic chemical network, a template polymerization and a membrane subsystem enclosing them all.

The first is convergence. Most viral capsids adopt a small number of simple geometrical structures so that their protein tertiary structures are subject to strong constraints. Hence, the probability that proteins converge towards similar folds to adapt to those constraints is far from negligible. Structural convergence occurs in protein motifs under strong selection, such as the active sites of enzymes[24,25]. This could also be the case for viral capsid proteins[26,27] or for viral and bacterial glycoproteins that are involved in cell entry[28]. Bacterial protein-based organelles, such as carboxysomes, have icosahedral shells that have astonishing geometrical similarities to those of viral capsids (FIG. 1), suggesting that this type of molecular architecture is prone to convergence[29].

The second alternative is horizontal gene transfer (HGT), which can move genes between extremely distant species. HGT seems to be rampant in viruses[30–32], which could explain why different viruses share some genes. Because of HGT, speculation about the antiquity of viral lineages just because they harbour one or a few common genes might be misguided. Extensive HGT could scramble the gene content of viral lineages to the point that their identities fade in short time spans. Consequently, high HGT levels are a huge problem in the quest to reconstruct viral evolutionary histories other than those of recent and compact lineages.

Distant hosts do not imply antiquity

It may seem reasonable to think that one or several viral lineages appeared early after, or even simultaneously to, cell evolution. However, this cannot currently be proven. The fact that some viral lineages infect phylogenetically distant hosts is sometimes used as evidence for their ancient origin. This requires a model of co-evolution between viruses and hosts, namely, that viruses speciate after hosts speciate. Accordingly, host and virus phylogenies must be congruent — that is, their topologies must have the same distribution of nodes — as must be their respective ages. For example, since their origin by endosymbiosis, mitochondria have co-evolved with eukaryotic cells so that the bacterial endosymbiont that gave rise to mitochondria is inferred to be as old as the ancestor of all present day eukaryotes[33]. Using this logic, viruses with similar architectures that infect prokaryotes and eukaryotes have been proposed to be at least as old as the cenancestor[21]. However, it is extremely difficult to apply this type of reasoning to parasites, for which host shift (the possibility for a parasite to infect unrelated hosts) is common[34]. For instance, we could deduce that syndinians — a group of parasitic dinoflagellates[35] — are as old as the whole eukaryotic domain because they can infect hosts that belong to distant eukaryotic phyla, including animals and various protists. This conclusion is wrong, however, as the origin of syndinians, a derived lineage of eukaryotes, cannot precede or be simultaneous to that of the ancestor of all eukaryotes; the capacity of syndinians to infect different hosts does not mean that they have co-evolved with them, just that they can shift between distant hosts.

Similarly, many viruses can move between different hosts[36]. At a close taxonomic scale, and taking human hosts as example, we know that viruses can shift hosts (for example, HIV came from primates, avian flu came from birds), which is also true at much larger taxonomic scales. And a strain of the flock house virus, a member of the family *Nodaviridae,* that usually infects insects[37], can infect hosts as distant as plants[38] and fungi[39]. Likewise, head-and-tail viruses that infect hyperhalophilic archaea are probably derived from bacteriophages that have jumped across domains[40]. Such host shifts could lead to false inferences of an ancient origin for widespread viral lineages if it is based only on the diversity of hosts, instead of on a careful phylogenetic analysis of viral and host markers to find the required evidence to prove co-evolution.

Viral lineages lack structural continuity

One universal attribute of cells and, consequently, of living beings, is the possession of membranes. An astonishing characteristic of some cell-membrane systems, such as the cytoplasmic membrane, is that they can only be formed by splitting pre-existing membranes (membrane heredity). These

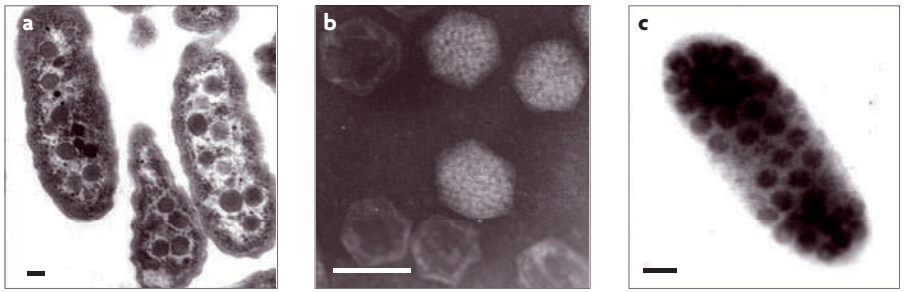


Figure 1 | **Limitations of morphology.** Simple geometric shapes, such as the icosahedral forms that are found in many viruses, can arise by convergence. This is exemplified by non-viral structures that are found in cells, such as the carboxysomes, shown in a bacterial cell (panel **a**) and in a thin section (panel **b**), compared with typical viruses, of comparable size and morphology, infecting a marine bacterium (panel **c**). The scale bars represent 100 nm. Panels **a** and **b** reproduced courtesy of T. O. Yeates, G. C. Cannon and S. Heinhorst (University of California, Los Angeles), panel **c** reproduced courtesy of S. W. Wilhelm and M. Weinbauer (University of Tennessee) (originally published in the Encyclopedia of Earth).

membranes have therefore been called genetic membranes[41]. The concept of membrane heredity implies the persistence of the cytoplasmic membrane from the first cells to contemporary cells. By contrast, there is no evidence for such a structural continuity in viruses: all viral constituents are synthesized *de novo* at each viral infection cycle by the enslaved cellular molecular machineries. This applies also to the lipid membrane envelope that characterizes some viral families; these are present either around the protein capsid (enveloped viruses) or in the protein capsid (as in the case of *Phycodnaviridae*). These membranes cannot be considered as analogues of cell plasma membranes; they do not grow or show structural continuity with previous viral membranes. On the contrary, they are always derived either from the host cytoplasmic endomembranes (endoplasmic reticulum, lysosomes and so on) or from the plasma membrane as the viral particle buds off the cell[42]. Whereas the existence of a genetic membrane provides strong evidence that all modern cells are derived from a single common ancestor[43], its absence in viruses is additional evidence for their polyphyletic origins.

Cellular origin of metabolic genes

As entities that depend entirely on their hosts, the majority of viruses lack genes for energy and carbon metabolism. Recent work on viral metagenomes nonetheless suggests that there are a significant number of genes for energy, carbon and cellular metabolism in viral fractions[44]. So far, however, detailed studies on those genes and the viruses they belong to are missing.

Among the well known cases of metabolism genes in viruses, the most remarkable is the presence of *psbA* and *psbD* genes that encode components of the photosystem II in cyanobacterial phages, which are transcriptionally active during infection[45]. In this and similar cases (such as genes that are involved in central metabolism, including carbonic anhydrases, superoxide dismutases and NlpC/P60 peptidases, as well as in DNA metabolism, including dUTPases, glutathione peroxidases, ribonucleotide reductases, thymidine kinases and uracil DNA glycolases), phylogenetic analyses demonstrate that these genes have been acquired from hosts by HGT[32,45–47]. Moreover, a simple inference of gene content in the ancestors of the different viral families for which complete genome sequences are available shows that they did not contain any of those genes. A lack of metabolic genes in those viral ancestors strongly argues against an ancient origin predating cells, invalidating recent claims that propose this scenario[18].

Cellular origin of translation genes

Some viruses, including the Mimivirus, possess several genes that are involved in protein synthesis[13]. However, these genes have been acquired from the host by HGT[32,48] (FIG. 2), implying that viruses never had the capacity to synthesize their own proteins (an additional reason together with their metabolic deficiency to argue that viruses did not predate cells). This has important phylogenetic implications, as the strongest claims to consider viruses as living beings with a place in the tree of life come from the presence of translation-related genes in certain viruses, which would eventually open the possibility of including them in universal phylogenetic trees based on those genes[13]. However, as those genes are cellular in origin, the corresponding trees do not reflect organismal phylogenies. This is also the case for viral genes that are involved in other informational processes (transcription and replication) that have close homologues in cells. The only known exceptions are genes that encode proteins involved in transcription and replication in mitochondria, which seem to come from a T-odd phage that infected the alphaproteobacterial ancestor of mitochondria[49]. All other cases that have been examined using a phylogenetic approach reveal that viral genes have been acquired from their hosts by HGT[50].

Viruses are gene robbers

Viruses evolve and recombine at much higher rates than cells[51,52]. Moreover, massive sequencing of viral genomes and metagenomes has revealed that viruses possess many genes that have no clear homologues in cells. Consequently, it has been speculated that viruses play a crucial part in the evolution of new protein functions, either by modification of pre-existing genes (as viral genes evolve faster and could soon reach levels of divergence far beyond homology detection) or by creation of completely new genes, and even that viruses may be at the origin of the many orphan genes (ORFans) of cellular genomes[11,18,30,53]. The first systematic survey of such ORFans in a large sample of 277 prokaryotic and 1,456 viral genomes, however, showed that less than 3% of the prokaryotic ORFans have viral homologues[54]. This, and the fact that the analysis did not enable any inference of the direction of the possible HGT events that might account for the presence of that reduced fraction of shared ORFans in prokaryotes and viruses, led the authors to conclude that "the evidence for viral gene transfer as the origin of microbial ORFans in general is currently weak, and even negative."[54]

It could be argued that viral undersampling could partly explain this low percentage. However, most sequenced viruses have been retrieved from prokaryotic or eukaryotic species for which genome sequences are also available and, therefore, it could be counter-argued that at least some of those ORFans should have homologues in the viruses infecting those hosts, which is apparently not the case. Therefore, viruses are unlikely to be donors of massive amounts of new genes to cells, although they may well be a reservoir of cellular genes that can be transferred between different hosts and could thus play a part in cellular adaptation and evolution. All in all, viruses are gene robbers, not gene inventors and massive gene suppliers.

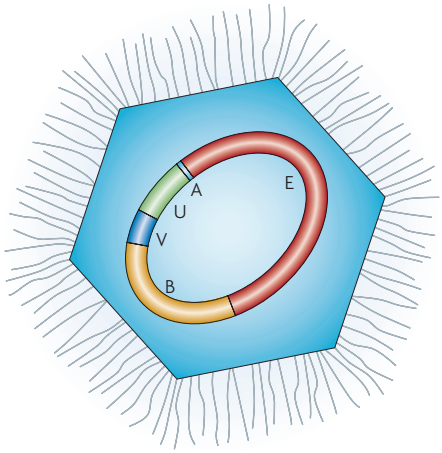


Figure 2 | **Multiple evolutionary origins of Mimivirus genes.** Phylogenetic analysis reveals that among the 126 conserved genes that have cellular homologues, 56% come from eukaryotic donors (E), 29% come from bacteria (B), 1% come from archaea (A), 5% have viral origin (V) and the remaining 9% have unresolved (U) phylogenetic origins[32,48].

Most HGT occurs from cells to viruses

The corollary of the three previous points is not only that most viral genes involved in energy and carbon metabolism, transcription, translation and replication with cellular homologues were acquired by viruses through HGT, but that the cell-to-virus gene flux is quantitatively overwhelming if compared with the opposite event[32,50]. This suggests that viruses have played only a minor part in shaping the gene content of cells (they might have served, however, as vehicles of gene exchange between cells). Paradoxically, although viruses are quantitatively the most abundant organic entities on Earth, to the point that cells live in an 'ocean of viruses'[55], viral genomes appear to live in an ocean of cellular genes. Given such a high frequency of cell-to-virus (as well as of virus-to-virus) HGT and the high recombination rates in viruses[52], the integrity of viral genomes should be lost in short times. In other words, a set of genes that is found together in a viral genome at a given time has little chance to remain linked after a small number of generations. With such a genomic plasticity, trying to reconstruct the evolutionary history of each individual gene of a viral lineage and inferring HGT events is possible, but such histories will not reflect the evolution of the viral lineage as a whole, as lineages cannot have genomic persistence in the presence of high HGT rates[56].

Simplicity does not mean antiquity

Because they are simple, viruses were embraced by many biologists as the missing link between life and non-life. Such a perception had profound historical and psychological grounds[2]. In particular, the anthropocentric view of evolution as a process leading to progress in the form of increased complexity accounted for the promptness to accept the idea that viruses must be extremely ancient because they are extremely simple. An Aristotelian-like 'scale of Nature' view with viruses at the origin of life became firmly anchored in part of the scientific community, despite the compelling evidence that accumulated against the idea that evolution is directional (that is, a progress-linked process) and despite the occurrence of many examples of regressive evolution, including in the microbial world[57]. Thus, structural simplicity implies neither antiquity nor primitiveness.

Regressive evolution is a fundamental process in parasite evolution. Even if certain viral lineages (such as the nucleocytoplasmic large DNA viruses) have increased their genome size and complexity from simpler ancestors[58], the general rule is that viruses are subjected to strong selective pressure to keep a minimal genome size in order to have faster reproduction rates. This is a major force preventing complexification. Consequently, viral simplicity is not evidence of viral antiquity or of primitiveness, but is a consequence of parasitism. A similar confusion applies to even simpler selfish genetic elements, such as the viroids — small single-stranded circular RNA molecules that, on the basis of their simplicity, were proposed to be relicts of a hypothetical pre-cellular RNA world[59]. As Oparin pointed out in 1961, "viruses, like other modern specific proteins and nucleic acids, could only have arisen as products of the biological form of organization."[6]

Concluding remarks

We have discussed multiple reasons that preclude considering viruses as living beings. To overcome most of the problems evoked above, Claverie recently proposed a provocative redefinition of the viral identity wherein the true nature of a virus is not the virion (the infective viral particle). "The virus factory should be considered the actual virus organism when referring to a virus. Incidentally, in this interpretation the living nature of viruses is undisputable, on the same footing as intracellular bacterial parasites"[60]. The "virus factory" comprises the structures that are involved in the replication and assembly of various viruses in the infected cells. They consist of complex assemblages of viral elements combined with recruited cell components, in particular membrane fragments coming from the cytoplasmic membrane, the endoplasmic reticulum, the nuclear membrane and the Golgi apparatus, as wells as from mitochondria and different cytoskeletal constituents[61]. According to Claverie, the virus factory is the virus "soma" and "interpreting the virion particle as 'the virus', is very much like looking at a spermatozoid and calling it a human."[60] The virion would be just a reproductive form, the virus "germline".

We refute this view. First, it is hard to accept that the definition of an organism necessarily requires portions of another organism. This would be akin to defining a tapeworm as the assembly of the parasitic flatworm and the human body that it requires for growth and reproduction. Second, no virus contains all the genes required to build a virus factory, as most of those genes, as well as the machinery to express the virus' own genes, are provided by the host. If the cell-recruited components from the virus factory are removed, the exclusive viral components are completely inert without a host. Third, the use of the apparently appealing analogy of a virion as a spermatozoid and a viral factory as a human is untenable. Virions are part of the viral infectious cycle and spermatozoids and ovules are the haploid components of the *Homo sapiens* life cycle, but both cycles are intrinsically different and not comparable. Viruses do not have sex and do not split any diploid genetic content into haploid gametes that, combined, have all the genes needed to develop the diploid stage of the *H. sapiens* species that we usually call human. Far from that, as just mentioned, viral genomes lack the genes to make any viral factory possible alone. Virions are indeed viruses just as spermatozoids are humans, in the phylogenetic *H. sapiens* species meaning. What else would spermatozoids be but *H. sapiens*? Members from another species? However, and this is an essential distinction that makes the analogy invalid: unlike viral factories, humans, including their gametes, are humans and not 'humans plus something else'.

If viruses are not alive and cannot be included in the tree of life, this does not imply that they have not had, or continue to have, a significant role in the evolution of life on Earth. On the contrary, being abundant (for example, marine waters contain at least one order of magnitude more viruses than prokaryotic cells[55]), and comprising a major selective pressure that exerts a strong control on the populations of many cellular organisms, they are an important source and means of maintaining biodiversity[55,62]. Like other

mobile genetic elements, such as transposons and retroposons that can become part of a host genome, they contribute to the production of genetic variability. Furthermore, as genetic elements with an extracellular phase, they can serve as vehicles to transfer host genes horizontally between cells across species, even from phylogenetically distant taxa. However, none of these points can be used to show that viruses are alive. Taken together, their inability to self-sustain and self-replicate, their polyphyly, the cellular origin of their cell-like genes and the volatility of their genomes through time make it impossible to incorporate viruses into the tree of life.

David Moreira and Purificación López-García are at the Unité d'Ecologie, Systématique et Evolution, UMR CNRS 8,079, Université Paris-Sud, bâtiment 360, 91405 Orsay Cedex, France.
Correspondence to D.M.
e-mail: david.moreira@u-psud.fr

doi:10.1038/nrmicro2108
Published online 9 March 2009

1. Muller, H. J. in *4th International Congress of Plant Science* 917–918 (ed. Duggar, B. M.) 917–918 (Bantha Publishing, Menasha, 1929)
2. Podolsky, S. The role of the virus in origin-of-life theorizing. *J. Hist. Biol.* **29**, 79–126 (1996).
3. Simon, C. E. (ed.) *The Filterable Viruses* (Reinhold, New York, 1928).
4. Haldane, J. B. S. The origin of life. *Rationalist Ann.* 3–10 (1929).
5. Beutner, R. *Life's Beginning on the Earth* (Williams and Wilkins, Baltimore, 1938).
6. Oparin, A. I. *Life: its Nature, Origin and Development* (Academic Press, New York, 1961).
7. Avery, O. T., MacLeod, C. M. & McCarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exper. Med.* **79**, 137–158 (1944).
8. van Regenmortel, M. H. V. in *7th Report of the International Committee on Taxonomy of Viruses* (eds van Regenmortel, M. H. V. *et al.*) 3–16 (Academic Press, San Diego, 2000).
9. van Regenmortel, M. H. V. in *Encyclopedia of Virology* (eds Mahy, B. W. J. & van Regenmortel, M. H. V.) 398–402 (Elsevier/Academic Press, 2008).
10. Bamford, D. H., Grimes, J. M. & Stuart, D. I. What does structure tell us about virus evolution? *Curr. Opin. Struct. Biol.* **15**, 655–663 (2005).
11. Forterre, P. 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 (2006).
12. La Scola, B. *et al.* A giant virus in amoebae. *Science* **299**, 2033 (2003).
13. Raoult, D. *et al.* The 1.2-megabase genome sequence of Mimivirus. *Science* **306**, 1344–1350 (2004).
14. Luisi, P. L. About various definitions of life. *Orig. Life Evol. Biosph.* **28**, 613–622 (1998).
15. Alexander, J. & Bridges, C. B. in *Colloid Chemistry, Theoretical and Applied* (ed. Alexander, J.) 54 (Reinhold, New York, 1928).
16. Guerrero, R., Piqueras, M. & Berlanga, M. Microbial mats and the search for minimal ecosystems. *Int. Microbiol* **5**, 177–188 (2002).
17. Fitch, W. M. Homology a personal view on some of the problems. *Trends Genet.* **16**, 227–231 (2000).
18. Koonin, E. V., Senkevich, T. G. & Dolja, V. V. The ancient virus world and evolution of cells. *Biol. Direct.* **1**, 29 (2006).
19. Doolittle, W. F. The nature of the universal ancestor and the evolution of the proteome. *Curr. Opin. Struct. Biol.* **10**, 355–358 (2000).
20. Ranea, J. A., Sillero, A., Thornton, J. M. & Orengo, C. A. Protein superfamily evolution and the last universal common ancestor (LUCA). *J. Mol. Evol.* **63**, 513–525 (2006).
21. Raoult, D. & Forterre, P. Redefining viruses: lessons from Mimivirus. *Nature Rev. Microbiol.* **6**, 315–319 (2008).
22. Benson, S. D., Bamford, J. K., Bamford, D. H. & Burnett, R. M. Does common architecture reveal a viral lineage spanning all three domains of life? *Mol. Cell* **16**, 673–685 (2004).
23. Rice, G. *et al.* 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 (2004).
24. Terada, T. *et al.* Functional convergence of two lysyl-tRNA synthetases with unrelated topologies. *Nature Struct. Biol.* **9**, 257–262 (2002).
25. Gherardini, P. F., Wass, M. N., Helmer-Citterich, M. & Sternberg, M. J. Convergent evolution of enzyme active sites is not a rare phenomenon. *J. Mol. Biol.* **372**, 817–845 (2007).
26. Wales, D. J. The energy landscape as a unifying theme in molecular science. *Philos. Transact A Math. Phys. Eng. Sci.* **363**, 357–375 (2005).
27. Olson, A. J., Hu, Y. H. & Keinan, E. Chemical mimicry of viral capsid self-assembly. *Proc. Natl Acad. Sci. USA* **104**, 20731–20736 (2007).
28. Barocchi, M. A., Masignani, V. & Rappuoli, R. Cell entry machines: a common theme in nature? *Nature Rev. Microbiol.* **3**, 349–358 (2005).
29. Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C. & Shively, J. M. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reviews Microbiology* **6**, 681–691 (2008).
30. Koonin, E. V. & Dolja, V. V. Evolution of complexity in the viral world: the dawn of a new vision. *Virus Res.* **117**, 1–4 (2006).
31. Koonin, E. V., Makarova, K. S. & Aravind, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709–742 (2001).
32. Moreira, D. & Brochier-Armanet, C. Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC Evol. Biol.* **8**, 12 (2008).
33. Gray, M. W. & Doolittle, W. F. Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **46**, 1–42 (1982).
34. Woolhouse, M. E., Taylor, L. H. & Haydon, D. T. Population biology of multihost pathogens. *Science* **292**, 1109–1112 (2001).
35. Coats, D. W. Parasitic life styles of marine dinoflagellates. *J. Euk. Microbiol.* **46**, 402–409 (2007).
36. Woolhouse, M. E., Haydon, D. T. & Antia, R. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol. Evol.* **20**, 238–244 (2005).
37. Ball, A. & Johnson, K. L. in *The Insect Viruses* (eds. Miller, L. K. & Ball, L. A.) 225–267 (Plenum Publishing, New York, 1998).
38. Selling, B. H., Allison, R. F. & Kaesberg, P. Genomic RNA of an insect virus directs synthesis of infectious virions in plants. *Proc. Natl Acad. Sci. USA* **87**, 434–438 (1990).
39. Price, B. D., Rueckert, R. R. & Ahlquist, P. Complete replication of an animal virus and maintenance of expression vectors derived from it in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **93**, 9465–9470 (1996).
40. Prangishvili, D., Forterre, P. & Garrett, R. A. Viruses of the Archaea: a unifying view. *Nature Rev. Microbiol.* **4**, 837–848 (2006).
41. Cavalier-Smith, T. Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* **5**, 174–182 (2000).
42. Miller, S. & Krijnse-Locker, J. Modification of intracellular membrane structures for virus replication. *Nature Rev. Microbiol.* **6**, 363–374 (2008).
43. Peretó, J., López-García, P. & Moreira, D. Ancestral lipid biosynthesis and early membrane evolution. *Trends Biochem. Sci.* **29**, 469–477 (2004).
44. Dinsdale, E. A. *et al.* Functional metagenomic profiling of nine biomes. *Nature* **452**, 629–632 (2008).
45. Mann, N. H., Cook, A., Millard, A., Bailey, S. & Clokie, M. Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* **424**, 741 (2003).
46. McClure, M. A. Evolution of the DUT gene: horizontal transfer between host and pathogen in all three domains of Life. *Curr. Protein Pept. Sci.* **2**, 313–324 (2001).
47. Bratke, K. A. & McLysaght, A. Identification of multiple independent horizontal gene transfers into poxviruses using a comparative genomics approach. *BMC Evol. Biol.* **8**, 67 (2008).
48. Moreira, D. & López-García, P. Comment on "The 1.2-megabase genome sequence of Mimivirus". *Science* **308**, 1114 (2005).
49. Shutt, T. E. & Gray, M. W. Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet.* **22**, 90–95 (2006).
50. Moreira, D. Multiple independent horizontal transfers of informational genes from bacteria to plasmids and phages: implications for the origin of bacterial replication machinery. *Mol. Microbiol.* **35**, 1–5 (2000).
51. Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. Rates of spontaneous mutation. *Genetics* **148**, 1667–1686 (1998).
52. Awadalla, P. The evolutionary genomics of pathogen recombination. *Nature Rev. Genet.* **4**, 50–60 (2003).
53. Forterre, P. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* **117**, 5–16 (2006).
54. Yin, Y. & Fischer, D. On the origin of microbial ORFans: quantifying the strength of the evidence for viral lateral transfer. *BMC Evol. Biol.* **6**, 63 (2006).
55. Suttle, C. A. Marine viruses—major players in the global ecosystem. *Nature Rev. Microbiol.* **5**, 801–812 (2007).
56. Zhaxybayeva, O. & Gogarten, J. P. Cladogenesis, coalescence and the evolution of the three domains of life. *Trends Genet.* **20**, 182–187 (2004).
57. Lwoff, A. L'évolution physiologique. *Etude des Pertes de Fonctions Chez les Microorganismes* (Hermann et Cie, Paris, 1943).
58. Iyer, L. M., Balaji, S., Koonin, E. V. & Aravind, L. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* **20**, 20 (2006).
59. Diener, T. O. Circular RNAs: relics of precellular evolution? *Proc. Natl Acad. Sci. USA* **86**, 9370–9374 (1989).
60. Claverie, J. M. Viruses take center stage in cellular evolution. *Gen. Biol.* **7**, 110 (2006).
61. Novoa, R. R. *et al.* Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol. Cell* **97**, 147–172 (2005).
62. Futse, J. E., Brayton, K. A., Dark, M. J., Knowles, D. P. Jr & Palmer, G. H. Superinfection as a driver of genomic diversification in antigenically variant pathogens. *Proc. Natl Acad. Sci. USA* **105**, 2123–2127 (2008).
63. Aristotle, D.a. De anima (350 BC) (ed. Hick, R. D.) (George Olms Verlag, Hildesheim, 1990).
64. Engels, F. *Herrn Eugen Dühring's Umwälzung der Wissenschaft* (Dietz Verlag, Stuttgart, 1894).
65. Schrödinger, E. *What is Life?* (Cambridge University Press, Cambridge, 1944).
66. Von Neumann, J. in *Lectures on the Theory and Organisation of Complicated Automata* (ed. Burks, A. W.) (University of Illinois Press, Urbana 1949).
67. de Duve, C. *Blueprint for a Cell* (Patterson, Burlington 1991).
68. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes* (Wiley, New York, 1961).
69. Bernal, J. D. in *Theoretical and Mathematical Biology* (eds. Waterman, T. & Morowitz, H. J.) 96–135 (Blaisdell, New York, 1965).
70. Gánti, T. *The Principles of Life* (Oxford University Press, Oxford, 2003).
71. Varela, F. G., Maturana, H. R. & Uribe, R. Autopoiesis: the organization of living systems, its characterization and a model. *Curr. Mod. Biol.* **5**, 187–196 (1974).
72. Maynard Smith, J. *The Problems of Biology* (Oxford University Press, Oxford, 1986).
73. Joyce, G. F. in *Origins of Life: the Central Concepts* (eds Deamer, D. W. & Fleischaker, G. R.) xi–xii (Jones & Bartlett, Boston, 1994).
74. Ruiz-Mirazo, K., Pereto, J. & Moreno, A. A universal definition of life: autonomy and open-ended evolution. *Orig. Life Evol. Biosph.* **34**, 323–346 (2004).

Acknowledgements
The authors thank three anonymous referees for helpful comments and criticisms, T.O. Yeates and S.W. Wilhelm for permission to use the photographs shown in Figure 1, and the French Agence Nationale de la Recherche (ANR JC05_44,674) and the CNRS for financial support.

FURTHER INFORMATION
Ecologie, Systématique et Evolution: http://www.ese.u-psud.fr/spip.php?rubrique7
Encyclopedia of Earth: http://www.eoearth.org/article/Marine_viruses

ALL LINKS ARE ACTIVE IN THE ONLINE PDF