The Great Billion-year War between Ribosome- and Capsid-encoding Organisms (Cells and Viruses) as the Major Source of Evolutionary Novelties

Patrick Forterre[a,b] and David Prangishvili[a]

[a] *Institut Pasteur, 75015 Paris, France*

[b] *Institut de Génétique et Microbiologie, Univ Paris-Sud, CNRS UMR 8621, 91405 Orsay Cedex, France*

Our conceptions on the origin, nature, and role of viruses have been shaken recently by several independent lines of research. There are many reasons to believe now that viruses are more ancient than modern cells and have always been more abundant and diverse than their cellular targets. Viruses can be defined as capsid-encoding organisms that transform their "host" cell into a viral factory. If capsid-encoding organisms (viruses) and ribosome-encoding organisms (cells) are the major types of living entities on our planet, it seems logical to conclude that their conflict has been a major engine of biological evolution (in the framework of natural selection). In particular, many novelties first selected in the viral world might have been transferred to cells as a consequence of the continuous flow of viral genes into cellular genomes. We discuss recent observations and hypotheses suggesting that viruses have played a major role at different stages of biological evolution, such as the RNA to DNA transition, the origin of the eukaryotic nucleus, or, alternatively, the origin of unique features in multicellular macrobes.

Key words: viruses; cells; natural selection; evolution

Introduction

The existence of the brain, one of the most amazing products of evolution, allows some human beings to make conscientious moral choices. They can decide, for instance, to protect the feeblest and the poor, and to consider all human life to be of identical values. Unfortunately, this holds for a minority, and most humans still obey the rule of natural selection, working hard to get power, wealth, and beautiful women (or men) for the success of their genes. Because the percentage of people with moral values and positive views on humanity is probably greater among scientists than in the whole population (we have no reference for that guess), they tend to be annoyed by some implications of natural selection. Being afraid of possible misinterpretations of social Darwinism, some biologists even tend to restrict the impact of natural selection (especially its life struggle component) in the evolution of the biosphere itself. As a consequence, they can miss some important mechanisms for innovation. It is well known, however, that war is a great source of material progress, as exemplified by all novel devices that were created during World War II (e.g., the invention of radar, the use of nuclear energy), although it was probably the most horrifying event in the history of mankind. We will argue here that the tendency of scientists to be repulsed by the images of war, explain why they have underestimated until now the role played by the war between viruses and cells during the evolution of life on our planet.

Address for correspondence: Patrick Forterre, Institut Pasteur, BMGE, 25 rue du Dr. Roux, 75015 Paris, France. Voice: +33 (0)145-68-87-91; fax: +33 (0)145-68-88-34. forterre@pasteur.fr

Natural Genetic Engineering and Natural Genome Editing: Ann. N.Y. Acad. Sci. 1178: 65–77 (2009).
doi: 10.1111/j.1749-6632.2009.04993.x c 2009 New York Academy of Sciences.

Didier Raoult and one of us have recently proposed to divide the living world into two distinct types of organisms, ribosome-encoding organisms (cellular organisms) and capsid-encoding organisms (viral organisms) in order to make justice to viruses and grant them a proper place in the biosphere.[1] If we agree with this view, it becomes immediately obvious that the war between the two major components of modern life, a war that started billions years ago (and is still raging), would have also been the major determinant for the invention of novelties during the evolution of life on our planet. Of course, a classical objection to this view is that relationships between cells and viruses are not always destructive; ribosome- and capsid-encoding organisms can sometimes cooperate for the benefit of each other. Harmony is of course the desire of most of us and is indeed the new slogan of the Chinese communist party. However, cooperation between viruses and cells is often temporary and/or is only a war alliance against a third party. Hence, many bacteria have enslaved some viral partners in their fight against predatory eukaryotes.[2]

In human war too, different groups can cooperate for a time against a third one, and slavery is a common outcome. Symbioses between cellular and viral organisms usually alternate with dramatic culling episodes in which most cellular targets are devastated.[3] The resulting holobionts itself became a new fighter in the war-game of life. The whole pattern remains that of a struggle to survive and dominate by destroying the enemy or by diverting part of its resources for your own purposes. Weapons are stolen and new ones are continuously created in a red queen race (you should move forward or you will remain behind). To realize that does not transform you into a potential small Hitler or Genghis Khan (once more, the possibility of our brain—for good or worst—remains surprisingly open) but helps to understand the biosphere as it is, so marvellous and so terrible.

For a long time viruses were viewed as (annoying) by-products of evolution, fragments of genetic material that escaped from their host prokaryotic or eukaryotic cells and rebelled against them (for a brief but pertinent historical account, see Ref. 3). Viruses were though as lost (and degenerated) children that have to be controlled and possibly eradicated (as in the case of small pox). Viruses were only interesting as models for molecular biologists and a *raison de vivre* for virologists. They were not so interesting (except for those who knew them well), but we had to study them to fight them (this particular war—human against viruses—being in that case legitimate for everybody, right or left wing oriented). This situation has dramatically changed during the last years when it became clear that modern viruses are not fragments of genetic materials that escaped from mother cells, but descendants of an ancient virosphere that probably even preceded the origin of modern cells (not to be confounded with the origin of cells per se).

We will see below that the war between cells and viruses was most likely already raging at the time of LUCA (the Last Universal Cellular Ancestor), and probably even before, during the period that one of us called "the second age of the RNA world" (after the invention of modern proteins, an obligatory component of viruses as we know them).[4] This period (before the invention of DNA) was certainly dominated by the conflict between RNA cells and RNA viruses (as our modern world is dominated by that between DNA cells and RNA/DNA viruses). In fact, the war between parasites and their hosts probably started even earlier, during the first age of the RNA world (before invention of the ribosome) and may have played a critical role in shaping the evolution of the very first living systems.

The destructive and creative life struggle between viruses and cells (and sometimes Darwinism as well) is usually completely ignored by scientists interested in the origin of life, who often adopt a "Lamarckian view" of biological evolution, in which organisms evolved from simple structures toward complex ones without any obvious reasons (eventually still looking for the magic mathematical formula or

cybernetic concept). However, the continuous struggle between primitive cells and their parasites should have had a major impact (even being the major factor?) in the invention of mechanisms such as the ribosome to produce modern proteins or the spliceosome to produce modern genes. However, we will not discuss these not yet explored possibilities here but focus on recent hypotheses suggesting that the war between ribosome- and capsid-encoding organisms has been at the origin of modern DNA genomes and has shaped the emergence of the three domains of life.

Viruses are Ancient

This conclusion is based on several observations. Firstly, it has been shown that some viruses infecting different members of the three cellular domains of life–archaea, bacteria, eukarya–encode homologous DNA/RNA replication proteins, more closely related between viruses from different domains than to any of their cellular homologues (for an early analysis, see Ref. 5, for recent reviews, see Refs. 6, 7). This suggests that these enzymes first originated in an ancient viral world that predated the divergence of the three domains of life. The same conclusion was independently corroborated by the discovery of homologous features in the virions of viruses infecting members of different domains of life. In particular, two types of protein folds have now been defined that, in both cases, are present in capsids from apparently unrelated viruses infecting either archaea, bacteria, or eukarya (for review, see Refs. 8, 9). In the case of the protein structure known as the "double-jelly roll fold," not only capsid proteins from viruses of the three cellular domain share this common fold, but the topology and organization of these proteins in their capsid surface lattices were found to be similar, excluding the possibility of convergence.[9] Preliminary data on the evolution of one of these proteins already clearly supports the idea that these homologies did not result from virus transfers between domains, but that capsids have coevolved with viral hosts after the formation of the three domains.[10] Indeed, homology between, capsid proteins with the double-jelly roll fold from viruses infecting different domain of life can be only detected by structural analysis, whereas homology between capsid proteins from viruses infecting members from different phyla of the same domain can be detected by sequence similarity. These data strongly suggest that modern viruses originated from ancient viruses that infected members of the community in which the LUCA was living. To explain both the presence of homology between some viruses infecting different domains and of viral families specific for each domain (see below the case of archaeal viruses), it has been suggested that the three ancestral populations of cellular organisms have randomly (and unwillingly!) selected, at the origin of the three modern domains, three different portions of the ancestral virosphere to travel with them in their journey from ancient to modern times.[11] To be in line with this proposal, and to get rid of the old nomenclature (viruses vs. bacteriophages) that arose from the misleading "prokaryote versus eucaryote" dichotomy,[12] we suggest to use the terms ***archeoviruses, bacterioviruses,*** and ***eukaryaviruses*** to name viruses infecting members of these different domains.

Viruses Outnumber Cells: Yesterday and Today

For more than a decade, it is clear from ecological studies and more recently from metagenomic studies that viruses represent the major part of the modern biosphere (for review, see Ref. 13). They outnumber cellular organisms by one log of magnitude in various environments, and each single cellular organism can be infected by several different viral species. There is no reason why the same situation could not already have been prevailing before LUCA, in the RNA world (as soon as viruses actually have originated). If true, this means that the

number of viral genes has always been higher than the number of cellular genes in the genosphere. Even if we use the conservative assumption that lateral gene transfer between cells and viruses occurs (and has always occurred) at the same rate in both directions, this would mean that, all in all, more genes should have been transferred in the course of evolution from viruses to cells than from cells to viruses.

In fact, it is likely that the rate of lateral gene transfer from viruses to cells has always been higher than that from cells to viruses, as indicated by the very low number of cellular homologues found in viral genomes compared to the very high number of viral genes integrated in cellular genomes. It is logical indeed to think that cells are more tolerant than viruses to the integration of new genes, because the size of the virion usually imposes a strict physical limitation to viral genome sizes. We can then conclude with certainty that much more genes were transferred from viruses to cells than in the opposite direction, in the course of evolution both before and after LUCA (until now indeed). Such conclusion is clearly in contrast with the traditional view that presents viruses as "pickpockets" that mainly evolve by recruiting cellular genes.[14] An extreme consequence of this "pickpocket" conception is to deny the existence of real viral genes. In this framework, viruses not being living entities, all viral genes should be ultimately of cellular origin, that is, they should have originated in an ancestor of LUCA or in one of its descendant. For instance, in a recent analysis of the genome of the giant mimivirus, it was conclude that "Mimivirus acquired most of these genes by horizontal gene transfer."[15] However, the authors did not consider in their analysis genes without cellular homologues although they comprise around 80% of the gene content!

Viral genes without cellular homologues probably have originated in very ancient viral lineages before LUCA, whereas others have continuously emerged by duplication and or recombination during the evolution of modern viral lineages, these processes creating an enormous amount of genetic diversity. The diversity and uniqueness of viral proteins is exemplified by the fact that, as in the case of mimivirus, most genes encoded by viruses have no homologues in cellular genomes (at least outside viral elements integrated in these genomes) or have only distantly related homologues (see for example Ref. 11, for the case of viruses infecting archaea). If the number of genes that were transferred from viral to cellular lineages in the course of evolution indeed always outnumbered the reverse flow. One can conclude that the cells are the real pickpockets on a long-term basis. In these conditions, it would be unreasonable to deny that this continuous avalanche of these so diverse and unique viral genes into cellular genomes during the last three or four billions years should have had major consequences in the formation of the cellular domains themselves.

Viruses are Much More Diverse than Previously Thought

The diversity of the viral world is astonishing, and we only start to appreciate correctly its meaning. It has been known for a long time that viruses are more diverse than cellular organisms in terms of genomic features, because their genomes can be made of either RNA, or DNA, single or double stranded, linear or circular, in various combinations. However, it was also widely believed that their morphology (in terms of virion) was quite monotonous, with mainly three families of head and tailed virions for bacterioviruses (order caudavirales), and spherical virions (more or less regular) for viruses infecting eukaryotes. The few exceptions were linear rods (for the virions of some plant viruses), or flexible filaments in the case of the bacteriovirus M13 and a few complex forms of eukaryaviruses. In recent years, this traditional view of viral diversity has been dramatically deepened by the discovery of viruses with unique morphology in the archaeal domain (Fig. 1, for review, see Ref. 16). The

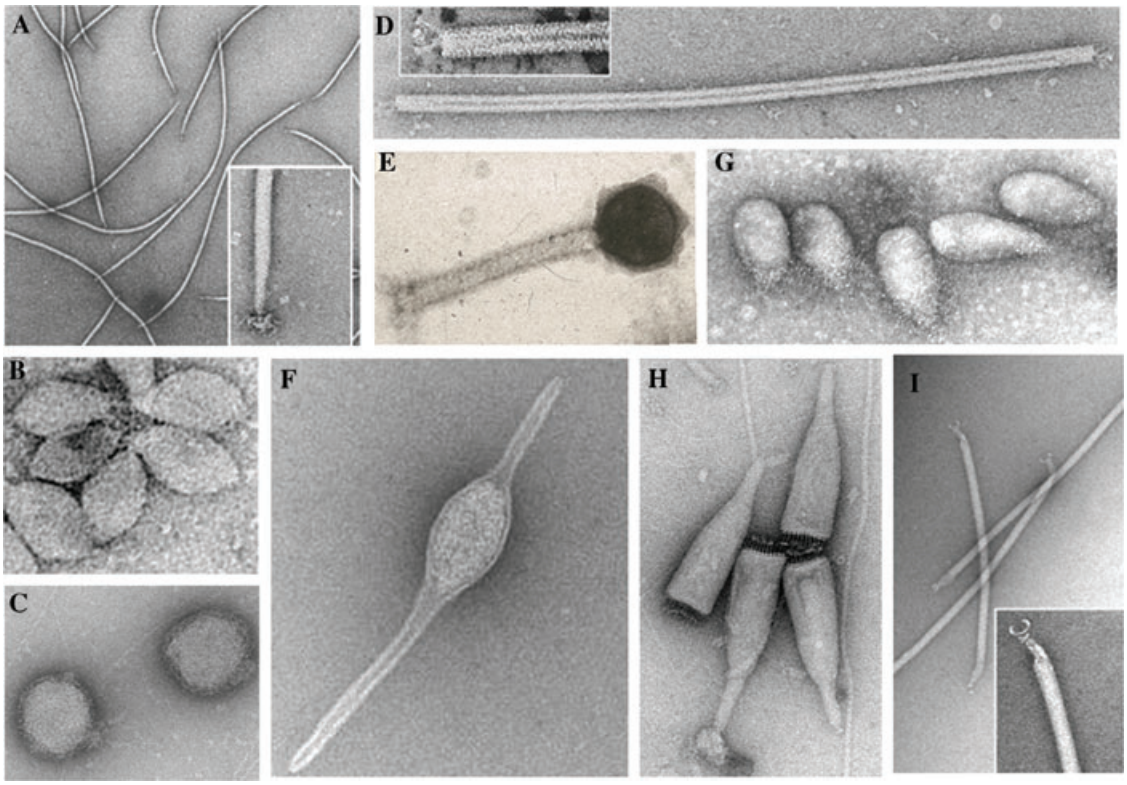

Figure 1. Electron micrographs of virions of double-stranded DNA viruses of archaea: (A) lipothrixvirus SIFV; (B) fusellovirus SSV1; (C) globulovirus PSV; (D) rudivirus SIRV1; (E) halovirus ØH1; (F) bicaudavirus ATV; (G) guttavirus SNFV; (H) ampullavirus ABV; (I) lipothrixvirus AFV1. All negatively stained with uranyl acetate. (A, B, D, E, and G) courtesy of W. Zillig; (C) from Ref. 18 modified with permission from Elsevier; (F) from Ref. 18 modified with permission from Nature; (H) from Ref. 55 modified with permission from American Society for Microbiology; (J) from Ref. 56 modified with permission from Elesevier.

virions of some archaeal viruses exhibit classical viral morphologies, being either spherical or linear (rod-shaped or flexible filaments), but with unique combinations of morphology and genome type, such as linear particles containing double-stranded DNA (vs. single-stranded DNA for linear bacterioviruses or RNA for linear eukaryaviruses). Some of them exhibit astonishing appendages, such as terminal clamps to grasp cellular pili. Even more, the virions of other archaeal viruses exhibit morphologies that were not previously observed, such as lemon-shaped (for the *Fuselloviridae*) or bottle-shaped (for the *Ampulloviridae*) (Fig. 1). In particular, the archaeal virus ATV (*Acidianus*-tailed-virus) encodes a capsid that undergoes the first known case of extracellular virion development.[17] The virions produced by infected cells are lemon-shaped particles that can be stored for months in water at room temperature without modification of their morphology. However, as soon as there are incubated at high temperature (above 70 C) these virions undergo a drastic structural reorganization with the formation of two long tails at opposite ends of the central body.[18] The diversity of morphotypes in archaeal viruses is paralleled by a diversity in their genome structure and content.[11] Viruses from a given family usually only share genes with viruses from the same family with very few exceptions. Considering that seven viral families have already been described by only studying viruses infecting archaea from one order (Sulfolobales), it is already quite clear

that the real diversity of archaeal viruses may be enormous and remains for the most part unexplored.

The Discovery of Giant Viruses: The Virion Factory

The discovery of giant eukaryaviruses (mimivirus and mamaviruses) that infect amoeba, has also shaken previous conceptions on the diversity and nature of viruses.[19–21] These viruses belong to a group of eukaryaviruses, the NCLDV (for nucleocytoplasmic large DNA viruses) that include complex viruses, such as poxviruses, with genomes in the range of 200–300 kb and multilayer envelopes. However, mimivirus breaks the record with a genome of 1200 kb (four times larger than the genome of a mycoplasma), encoding more than 900 proteins. It stains Gram positive and was originally confused with a bacterium (the size of the virion, that incorporates more than 100 proteins, is similar to the size of a mycoplasma).[22] The intracellular viral factory of mimivirus is especially spectacular, with size and shape that are similar to those of the nucleus of the infected protist host.[23] Although the viral factory of mimivirus is exceptional in terms of size, the formation of a complex intracellular factory is not unique to giant viruses but is a common theme among eukaryaviruses.[24,25] These viral factories are usually surrounded by membranes that are recruited from the endoplasmic reticulum. Jean-Michel Claverie correctly pointed out that the viral factory corresponds to the real viral organism, whereas the virion (traditionally confused with the virus) corresponds to the mechanism used by viral factories to spread from one cell to the other.[21] In that sense, the viral factory should probably be called instead the "virion factory" (for a previous paper that focuses on the intracellular stage of the virus cycle instead of focusing on the virion, see Ref. 26).

In the case of archaea and bacteria, we would like to argue that the virus transforms the entire infected cell into a virion factory. After destruction or inactivation of the cellular genome, when the viral genome is the only one to be expressed, the infected "cell" is no more a bacterium or an archaeon, but a virus with a cellular appearance. A nice example of this conversion is provided by cyano-bacterioviruses (cyanophages) that encode their own photosynthetic reaction center protein to replace the decaying cellular one, in order to obtain the proper energy required for its virion factory (Ref. 27 and references therein). The former cyanobacterial cell thus becomes a photosynthetic virus.

From all these examples, it is clear that viruses can themselves be considered as cellular organisms, because they are always living inside a cell. They are of course a particular case of cellular organisms that do not encode their own ribosomes and cell membranes but borrow those from the cells in which they are living. A final touch in this reevaluation of the living status of viruses was the discovery that mimivirus can become ill, after having being infected by another virus, the virophage called sputnik.[20]

Viruses and the Origin of Modern Genomes

The diversity of viral genomes and replication mechanisms is especially striking; whereas all cellular organisms have linear or circular double-stranded DNA genomes and replicate by a symmetric mode of replication (the two strands being replicated concurrently), viruses exhibit both a diversity of genome structures and of replication mechanisms. When double-stranded, they can replicate the two strands of DNA either one after the other (asymmetric replication), or at the same time (concurrently or one being delayed). To initiate DNA replication, they can use an RNA primer, as cellular organisms do, but they can also use a protein molecule as primer, a tRNA or the 3 -OH end of DNA produced by nicking. In terms of

proteins involved in genome replication, viruses encode enzymes that are remarkably different from their cellular counterparts. Some of them clearly have no homologues in the cellular world (for reviews see Refs. 4, 6, 7). For example, this is the case for the monomeric RNA polymerase encoded by the bacteriovirus T7, the DNA helicase of superfamilly III, or the Rep proteins that initiate rolling-circle replication. Others have cellular homologues, but are very distinct from them. This is the case for viral DNA topoisomerases or DNA polymerases, which form specific groups in phylogenetic trees, well separated from their cellular counterparts.[28,29] Phylogenetic analyses indicate that this odd phylogenetic pattern cannot be explained by the faster rate of viral sequences evolution but implies these viral patterns emerged in modern cellular lineages but in an ancient virosphere.[6,7] For these enzymes, it is thus possible to define viral versions, specific to broad viral families, that remind us of the three domain-specific versions of universal cellular proteins.

All these observations clearly exclude trivial explanations suggesting that the diversity of viral genomes, mechanisms of genome replication, and replication proteins are the consequence of the rapid evolutionary rate of viral proteins. The most likely explanation for this diversity is that all these mechanisms and proteins originated in a very ancient virosphere of DNA viruses that already existed at the time of LUCA and possibly even preceded it (in agreement with conclusions drawn from structural analysis of capsid proteins). One of us has suggested that DNA itself and associated mechanisms originated (as a modified form of RNA) in this ancient virosphere, in the framework of the war between ribosome- and capsid-encoding organisms.[30] There is indeed a clear selective advantage for a virus to modify the chemistry of its genome in order to protect it against cellular defense mechanisms. Interestingly, the viral hypothesis for the origin of DNA is in line with the principle of continuity, because the creation of new DNA sequences from RNA sequences via retro-transposition (a process itself a hallmark of the viral world) has been going on continuously, and is still going on in the eukaryal lineage, possibly since the RNA world.[31]

If DNA originated first in an ancient virosphere, many DNA replication, recombination, and repair proteins might have then originated in the viral world, before being transferred from viruses to cells in the course of evolution.[32] It has thus been suggested that modern DNA replication mechanisms originated there too.[30,33,34] At least two transfers have been postulated to explain why archaeal/eukaryal and bacterial DNA replication proteins are not homologous. One of these transfers might have already occurred either before LUCA, requiring the non-orthologous replacement of the ancestral DNA replication mechanism by a new one in archaea/eukarya or in bacteria, or just after LUCA, with the implication of a LUCA with an RNA genome.[30,33] Considering the greater replication fidelity of DNA versus RNA, the viral-induced RNA to DNA transition might have had a critical role in the formation of the cellular domains by slowing down the tempo of gene evolution. Three independent transitions (one at the origin of each domain) could explain the existence of three versions of universal cellular proteins, such as ribosomal proteins.[35] On the other hand, two transitions, one at the origin of bacteria, another in a lineage common to archaea and eukarya, could explain the clear divide observed between the bacterial and the archaeal/eukaryal versions of universal proteins.

Although the above ideas cannot be "experimentally" tested, at least the possibility that viral DNA replication proteins replaced cellular ones in the course of evolution has been demonstrated in the case of mitochondria. Indeed, phylogenetic analyses have shown without any doubt that the RNA polymerase, DNA polymerase, and DNA helicase that transcribe and replicate DNA in modern mitochondria are of viral origin.[28] The viral enzymes, that were originally encoded by a virus integrated in the genome of the bacterium at the origin of

mitochondria, have displaced later on the bacterial RNA polymerase, DNA polymerase, and DNA helicase that originally transcribed and replicated the mitochondrial genome. These three viral proteins are now encoded by the nuclear genome and targeted to mitochondria.

Viruses and the Origin of the Cellular Nucleus

A major role has been suggested for viruses in the origin of the eukaryotic nucleus.[36,37] Indeed, there are striking similarities in the assembling of the eukaryotic nucleus and the intracytoplasmic virion factories of many DNA and RNA eukaryoviruses. In both cases, membranes of the endoplasmic reticulum are recruited to form the nuclear or the viral envelopes and there is a tight coupling between nuclear or viral membrane formation and DNA replication. The similarity is such that the virion factories of poxviruses have been dubbed "mininuclei" by virologists,[38] and the virion factory of the mimivirus (no more a mini but a maxinucleus) has been originally confused by his discoverers with the nucleus of the infected amoeba (Raoult, personal communication).

The authors of the original eukaryogenesis hypothesis have suggested that the nucleus originated from a large DNA virus (resembling a poxvirus) that infected an ancient archaeon, in order to take into account the extensive similarities between archaeal and eukaryal molecular biology. However, such a scenario is unlikely for several reasons, (i) NCLDV appear to be specific of eukaryotic organisms, we don't know presently any giant archaeovirus that resembles poxviruses; (ii) all known archaea lack the system of intracellular membranes that seems to be critical in the formation of a nucleus; (iii) the eukaryotic lipids are strikingly different from the archaeal ones, whereas they should be similar if the host of the ancestral viral proto-nucleus was an archaeon; and (iv) archaea are monophyletic,[39] indicating that eukaryotes did not derive from archaea, but that archaea and eukarya share a common ancestor instead. The last point is especially important because it implies that eukaryotic features, such as the nucleus, are either primitive features that were lost in archaea or derived features that emerged in the stem of the eukaryotic lineage.[40] If the viral eukaryogenesis hypothesis is correct, the host of the ancestral viral proto-nucleus was therefore not an archaeon but either a proto-eukaryotic cell, or a cell predating the divergence between archaea and eukarya, possibly still a member of the RNA world.[35]

As an alternative to the eukaryogenesis hypothesis in which the nucleus directly derived from a large DNA virus, we propose here, as a new hypothesis, that a proto-eukaryotic cell recruited a viral mechanism for membrane formation in order to build the nuclear membrane. For some time this membrane might have protected the cellular genome from the attack of viruses, before some viruses learned how to enter the nucleus. A major question in cell biology is why a nucleus evolved in the first place. The virion factory clearly offers the virus protection against the defense systems of the host. The genome of the virus remains trapped in the virion factory during the viral cycle. In our hypothesis, we postulate that by "inventing" the nucleus some infected cells might have learned how to use this viral strategy for their own purpose. This hypothesis provides a direct and immediate selective advantage for the organism in which the first nucleus arose (as in the case of the viral hypothesis for the "invention" of DNA). Interestingly, in that hypothesis, the cell may have borrowed from the virus at the same time both the nuclear membrane and the transport mechanism (via nuclear pore) allowing the transfer of information and materials between the proto-nucleus and the proto-cytoplasm. Indeed, the mRNA of many eukaryaviruses that replicate in the cytoplasm of their host cell is transcribed inside the virion factory and has to be transferred to the host cytoplasm through nuclear-like pores. Such pores have been recently visualized by electron tomography in the case of the vaccinia virus.[41] The

uncoupling of transcription and translation can thus be viewed as another "viral feature" of eukaryotic cells!

It is remarkable that not only DNA viruses, but RNA viruses too, can recruit membranes of the endoplasmic reticulum to make nucleus-like viral factories.[25] The viral replication apparatus is surrounded by one or two membrane layers in these RNA mininuclei with an opercule for communication with the cytoplasm.[25] This opens the possibility that cellular organisms of the RNA world might have included both types of organisms, with and without nuclei. The emergence of a nucleus has also occurred in some bacteria of the order Planctomycetales that posses an intracellular cytoplasmic membrane (ICM).[42] The ICM is recruited to completely surround the nucleoid in the case of the species *Gemmata obscuriglobus*. Existence of this synkaryotic bacterium (with nucleus) testifies that formation of the eukaryotic nucleus has not been such a unique event, as it is often assumed (although this particular nucleus is indeed especially complex).

Viruses and the Origin of the Eukaryotic mRNA Capping Mechanism

Viral messenger RNA has to be distinguished from cellular RNA, both by the virus that wishes to degrade the cellular messenger RNA and/or favor translation of its own messenger, as well as by the attacked cell that, on the contrary, needs to prevent the translation of viral messenger RNA. The best way to distinguish two messenger RNAs of different origins is to label one of them with a specific tag. Precisely, another specific feature common to Eukarya and some eukaryoviruses is the capping mechanism that modifies the 5 end of messenger RNA to produce methyl-guanosine.[43] To us, it is likely that capping systems of mRNA in eukaryotic cells originated during the conflict between cells and viruses. This hypothesis was already proposed by Shuman who suggested that "it is conceivable that the 5 exoribonucleases present in Eukarya and cap synthesis evolved in tandem in early eukaryotes to provide a primitive immunity to RNA viruses."[43] Most viral capping enzymes are homologous to their cellular counterparts, and it is usually assumed (in the framework of the virus pickpocket paradigm) that they have been recruited from cells. However, in our opinion, if cells are the real pickpocket, one should not exclude the possibility that, the transfer occurred in the other direction.

The Role of Viruses in the Formation of Cellular Domain Phenotypes

The capping system is only one among a number of curious and complex features of eukaryotic molecular biology. One can wonder if this complexity is not, for the most part, a by-product of the billion-years war between cells and viruses. The fact that eukaryotic cells apparently originated from an ancestor that was a predator feeding via phagocytosis, and have evolved toward large sizes, by predation via phagocytosis, make them attractive targets for viruses producing virions that could fuse with their cell membrane and play with their intracellular membrane trafficking systems. On the contrary, the relative simplicity of the molecular mechanisms and cellular structure in archaea and bacteria could reflect the fact that these organisms have been more successful in their fight against at least some forms of viruses (e.g., RNA viruses are scarce in bacteria and for the moment unknown in archaea). Because the first step in the entrance of a virion (or its nucleic acid) into any cellular organism involves an intimate contact between the virion external surface and the cell external envelope, the development of efficient defense mechanisms against viruses might have played a crucial role in the formation of a rigid and thick cell wall at the surface of bacterial cells. The evolution of a complex cell wall in bacteria could

possibly explain the relatively low diversity of bacterioviruses compared to archaeoviruses and eukaryoviruses. The predominance of head and tail viruses (Caudavirales) among bacterioviruses[44] could reflect the evolutionary success of the viral family that invented the powerful tail-entry mechanism to crack the cell wall barrier. In that case, the greater diversity of archaeal viruses could reflect the lack of rigid and thick cell wall in most archaea. In this framework, it would be very interesting to determine the nature of viruses infecting unusual bacteria, such as Planctomycetes, whose cell walls contain no peptidoglycans and have intracellular membrane systems. Finally, considering the importance of the envelopes in defining the "self," one can wonder if the interplay between viruses and cells could have played an important role in "domain speciation" (divergence between domains), and later on in the speciation within domains, because mechanisms invented to prevent viral infection might have also, as a by-product, prevented the extensive genetic exchanges between domains or between domain divisions once they were established.

The fight with viruses probably also influenced the types of genomes found in archaea and bacteria on one side and in eukarya on the other. Archaea and bacteria contain a plethora of plasmids that probably originated from viruses. In fact, the archaeal and bacterial genomes look like megaplasmids and may have originated from ancestral megaplasmids.[35] In general, viruses are probably at the origin of the "mobilome" of archaea and bacteria, that is, the fraction of genomes constituted by mobile elements, not only proviruses, but also transposons, insertion sequences, integrons, and pathogenicity islands.[45] The obvious evolutionary link between plasmids and viruses suggests that mechanisms such as conjugation, restriction-modification, toxin-antitoxin systems, all originated in the framework of the war between cells and viruses.[46] The recent discovery in archaea and bacteria of a conserved defense system against viruses and plasmids, the CRISPR sequences and associated Cas proteins, again reminds us that cells have to make considerable efforts to survive the attack of viruses.[46,47] The eukaryotic chromosomes are strikingly different from the archaeal and bacterial ones, and it might not be a coincidence that the mobilome of eukaryotes (essentially based on retro-elements) is also dramatically distinct from the mobilome of archaea and bacteria.[48] This suggestion fits well with the idea that originally some viruses or derived elements were critical in determining the structure of cellular chromosomes.

To sum up, it is really possible that the various historical turns in the fight between viruses and cells have played the major role in the origin of modern domains and in their different lifestyles. Conversely, the evolution of cellular organisms has also shaped the evolution of the virosphere by selecting the viruses that at each stage were able to successfully bypass the defenses raised by their potential victims.

Viruses and the Explosive Evolution of Macrobes

The pivotal role played by viruses in life evolution has not been limited to the realm of microbes (99% of the biosphere) but has probably been critical for the emergence and evolution of macrobes (as organisms visible with naked eyes could be named).[49] In recent years, the role of viruses in evolution has been especially well documented in multicellular animals.[3] A characteristic feature of animal genomes is the presence of a high proportion of integrated retroviruses and various repetitive elements of retroviral origin. For instance, around 42% of the human genome contains sequences whose origins can still be traced to retroposition.[50] Endogeneous viruses and derived retro-elements can manipulate genomes by changing the pattern of gene expression (via integration and/or recombination) and by creating new genes (in particular via the formation of new sites promoting alternative splicing[31]) (see also Refs. 3, 51 on the creative power of viruses).

Fascinating examples of this creative power have already been identified (see discussion by Villareal of the role of viruses in the origin of the immune system, embryogenesis, or placentation[51–53] and references therein). Considering the extent of viral occupation of their genomes, metazoans can probably be considered as viral symbionts (holobionts), products of ancient and extensive "retroviral symbiogenesis."[51] The principal difference between chimps and us is clearly the number, variety, and, more importantly, the integration loci of elements of viral origin in our genomes. The origin of humanity will certainly be viewed someday as a particular outcome of viral creativity.

Conclusion

The origin of species (life evolution in toto) is the product of variation and selection.[54] For a long time after the initial formulation of this evidence by Charles Darwin, the nature of variations remained unclear. After the discovery of mutations and the cracking of the genetic code, it was usually believed that mutations (originally mostly envisaged as single mutations) were the main sources of variation. More recently, it became evident that more complex processes are at work. Brosius wrote recently, focusing on retroviruses and eukaryotic cells, that "the interaction of hosts with retroviruses, retrotransposons and retroelements is one of the eternal conflicts that drive the evolution of life."[31] We would like to extent this argument here to the whole biosphere by postulating that, once viruses appeared on the stage of life, the major cause of variation became the interplay between viral and cellular genomes, and the major cause of selection became the war between capsid and ribosome-encoding organisms. We thus finally conclude that the billion years war between cells and viruses has been (and still is) the major engine of life evolution.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Raoult, D. & P. Forterre. 2008. Redefining viruses: Lessons from mimivirus. *Nat. Rev. Microbiol.* **6:** 315–319.
2. Brussow, H. 2007. Bacteria between protists and phages: From antipredation strategies to the evolution of pathogenicity. *Mol. Microbiol.* **65:** 583–589.
3. Ryan, R.F. 2006. Genomic creativity. *Biol. J. Linn. Soc.* **88:** 655–672.
4. Forterre, P. 2005. The two ages of the RNA world, and the transition to the DNA world: A story of viruses and cells. *Biochimie* **87:** 793–803.
5. Forterre, P. 1992. New hypotheses about the origins of viruses, prokaryotes and eukaryotes. *In* Frontiers of Life. T. Thanh *et al.*, Eds.: 221–234. Editions Frontieres. Gif-sur-Yvette, France.
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., T.G. Senkevich & V.V. Dolja. 2006. The ancient virus world and evolution of cells. *Biol. Direct* **1:** 29.
8. Bamford, D.H., J.M. Grimes & D.I. Stuart. 2005. What does structure tell us about virus evolution? *Curr. Opin. Struct. Biol.* **15:** 655–663.
9. Krupovic, M. & D.H. Bamford. 2008. Virus evolution: How far does the double beta-barrel viral lineage extend? *Nat. Rev. Microbiol.* **6:** 941–948.
10. Krupovic, M. & D.H. Bamford. 2008. Archaeal proviruses TKV4 and MVV extend the PRD1-adenovirus lineage to the phylum Euryarchaeota. *Virology* **375:** 292–300.
11. Prangishvili, D., R.A. Garrett & E.V. Koonin. 2006. Evolutionary genomics of archaeal viruses: Unique viral genomes in the third domain of life. *Virus Res.* **117:** 52–67.
12. Pace, N.R. 2006. Time for a change. *Nature* **441:** 289.
13. Suttle, C.A. 2007. Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* **5:** 801–812.
14. Moreira, D. & P. Lopez-Garcia. 2005. Comment on "The 1.2-megabase genome sequence of Mimivirus." *Science* **308:** 1114; author reply 1114.
15. Moreira, D. & C. Brochier-Armanet. 2008. Giant viruses, giant chimeras: The multiple evolutionary histories of mimivirus genes. *BMC Evol. Biol.* **8:** 12.
16. Prangishvili, D., P. Forterre & R.A. Garrett. 2006. Viruses of the Archaea: A unifying view. *Nat. Rev. Microbiol.* **4:** 837–848.

17. Haring, M. *et al*. 2004. Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera Pyrobaculum and Thermoproteus: A novel virus family, the Globuloviridae. *Virology* **323:** 233–242.
18. Haring, M., R. Rachel, X. Peng, *et al*. 2005. Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, Acidianus bottle-shaped virus, from a new family, the Ampullaviridae. *J. Virol.* **79:** 9904–9911.
19. Raoult, D. *et al*. 2004. The 1.2-megabase genome sequence of mimivirus. *Science* **306:** 1344–1350.
20. La Scola, B. *et al*. 2008. The virophage as a unique parasite of the giant mimivirus. *Nature* **455:** 100–104.
21. Claverie, J.M. 2006. Viruses take center stage in cellular evolution. *Genome Biol.* **7:** 110.
22. Renesto, P. *et al*. 2006. Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products. *J. Virol.* **80:** 11678–11685.
23. Suzan-Monti, M., B. La Scola, L. Barrassi, *et al*. 2007. Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga mimivirus. *PLoS ONE* **2:** e328.
24. Novoa, R.R. *et al*. 2005. Virus factories: Associations of cell organelles for viral replication and morphogenesis. *Biol. Cell* **97:** 147–172.
25. Miller, S. & J. Krijnse-Locker. 2008. Modification of intracellular membrane structures for virus replication. *Nat. Rev. Microbiol.* **6:** 363–374.
26. Bandea, C.I. 1983. A new theory on the origin and the nature of viruses. *J. Theor. Biol.* **105:** 591–602.
27. Bragg, J.G. & S.W. Chisholm. 2008. Modeling the fitness consequences of a cyanophage-encoded photosynthesis gene. *PLoS ONE* **3:** e3550.
28. Filee, J. & P. Forterre. 2005. Viral proteins functioning in organelles: A cryptic origin? *Trends Microbiol.* **13:** 510–513.
29. Forterre, P., S. Gribaldo, D. Gadelle & M.C. Serre. 2007. Origin and evolution of DNA topoisomerases. *Biochimie* **89:** 427–446.
30. Forterre, P. 2002. The origin of DNA genomes and DNA replication proteins. *Curr. Opin. Microbiol.* **5:** 525–532.
31. Brosius, J. 2003. The contribution of RNAs and retroposition to evolutionary novelties. *Genetica* **118:** 99–116.
32. Forterre, P., J. Filée & H. Myllykallio. 2004. Origin and evolution of DNA and DNA replication machineries. In *The Genetic Code and the Origin of Life*. L. Ribas, Ed.: 145–168. Landes Bioscience. Austin, TX.
33. Forterre, P. 1999. Displacement of cellular proteins by functional analogues from plasmids or viruses could explain puzzling phylogenies of many DNA informational proteins. *Mol. Microbiol.* **33:** 457–465.
34. Villarreal, L.P. & V.R. DeFilippis. 2000. A hypothesis for DNA viruses as the origin of eukaryotic replication proteins. *J. Virol.* **74:** 7079–7084.
35. Forterre, P. 2006. 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.
36. Bell, P.J. 2001. Viral eukaryogenesis: Was the ancestor of the nucleus a complex DNA virus? *J. Mol. Evol.* **53:** 251–256.
37. Takemura, M. 2001. Poxviruses and the origin of the eukaryotic nucleus. *J. Mol. Evol.* **52:** 419–425.
38. Tolonen, N., L. Doglio, S. Schleich, *et al*. 2001. Vaccinia virus DNA replication occurs in endoplasmic reticulum-enclosed cytoplasmic mini-nuclei. *Mol. Biol. Cell* **12:** 2031–2046.
39. Yutin, N., K.S. Makarova, S.L. Mekhedov, *et al*. 2008. The deep archaeal roots of eukaryotes. *Mol. Biol. Evol.* **25:** 1619–1630.
40. Poole, A.M. & D. Penny. 2007. Evaluating hypotheses for the origin of eukaryotes. *Bioessays* **29:** 74–84.
41. Cyrklaff, M. *et al*. 2005. Cryo-electron tomography of vaccinia virus. *Proc. Natl. Acad. Sci. USA* **102:** 2772–2777.
42. Fuerst, J.A. 2005. Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* **59:** 299–328.
43. Shuman, S. 2002. What messenger RNA capping tells us about eukaryotic evolution. *Nat. Rev. Mol. Cell Biol.* **3:** 619–625.
44. Ackermann, H.W. 2007. 5500 phages examined in the electron microscope. *Arch. Virol.* **152:** 227–243.
45. Leplae, R., A. Hebrant, S.J. Wodak, *et al*. 2004. ACLAME: A classification of mobile genetic elements. *Nucleic Acids Res.* **32:** D45–D49.
46. Koonin, E.V. & Wolf, Y.I. 2008. Genomics of Bacteria and Archaea: The emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* **36:** 6688–6719.
47. Brouns, S.J. *et al*. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321:** 960–964.
48. Filee, J., P. Siguier & M. Chandler. 2007. Insertion sequence diversity in Archaea. *Microbiol. Mol. Biol. Rev.* **71:** 121–157.
49. Forterre, P. 2008. In a world of microbes, where should microbiology stand? *Res. Microbiol.* **159:** 74–80.
50. de Parseval, N. & T. Heidmann. 2005. Human endogenous retroviruses: From infectious elements to human genes. *Cytogenet. Genome Res.* **110:** 318–332.
51. Ryan, R.F. 2007. Viruses as symbionts. *Symbiosis* **44:** 11–21.
52. Villarreal, L.P. 2005. *Viruses and the Evolution of Life*. ASM Press. Washington, DC.

53. Villarreal, L.P. 2007. Virus-host symbiosis mediated by persistence. *Symbiosis* **44:** 1–9.
54. Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray. London.
55. Haring, M. *et al*. 2005. Virology: Independent virus development outside a host. *Nature* **436:** 1101–1102.
56. Bettstetter, M., X. Peng, R.A. Garrett, *et al*. 2003. AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. *Virology* **315:** 68–79.