Old Herborn University Seminar Monograph 31:
Evolutionary biology of the virome, and impacts in human health and disease.
Editors: Peter J. Heidt, Pearay L. Ogra, Mark S. Riddle and Volker Rusch.
Old Herborn University Foundation, Herborn, Germany: 15-26 (2017).

THE ORIGIN, NATURE AND DEFINITION OF VIRUSES (AND LIFE): NEW CONCEPTS AND CONTROVERSIES

PATRICK FORTERRE

Institute Pasteur, Department of Microbiology, and
Institute for Integrated Cell Biology, Department of Microbiology,
University of Paris-Saclay, Paris, France

SUMMARY

During the last two decades, our vision of the viral world has been profoundly modified by several discoveries in different fields of biology. Many of these discoveries conflict with the traditional view of viruses as inert biological objects that played a minor role in evolution and mainly evolve by picking genes from their hosts. It has been realized that viruses are very ancient, predating the Last Universal Cellular Ancestor (LUCA), extremely diverse, and have played a major role in life evolution. New definitions and concepts of viruses have been proposed to take these new discoveries into account. In particular, the virocell concept states that viruses are cellular organisms and emphasizes their ability to produce their own genes. Although the virocell concept remove arguments against the non-living status of viruses, the definition of life and living organisms remains challenging. Here, viruses are defined as capsid encoding organisms and life as the mode of existence of biological entities (individuals in philosophical term).

INTRODUCTION

For years, most biologists considered viruses as by-products of biological evolution that could have only play a minor role in the history of the living world. This has gradually changed recently as a result of several advances in different fields of biology. The molecular ecologists have highlighted the extraordinary abundance of viral particles and viral genes in the environment (*Kristensen* et al., 2010, *Suttle*, 2013), the structuralists biologists have shown unexpected kinship between viruses infecting organisms belonging to different cellular domains (archaea, bacteria or eukaryotic) determining the structure of the proteins forming the viral capsid (*Abrescia* et al., 2012). At the same time, the study of archaeal virus revealed a fascinating world of different viruses previously unknown in bacteria and eukaryotes (*Prangishvili*, 2013). To top it all, the discovery of giant virus has caught the imagination of the scientific community by revealing the existence of viruses whose genomes are greater than those of many bacteria and archaea (*Raoult* et al., 2013). All these findings have revived interest in viruses and rested the issue of their nature - living or not - and the definition of life itself. Here, I review the definition of viruses and life that I proposed recently (*Forterre*, 2016, and references therein) and I discuss the virocell concept, that challenges the traditional view that assimilate viruses to their virions.

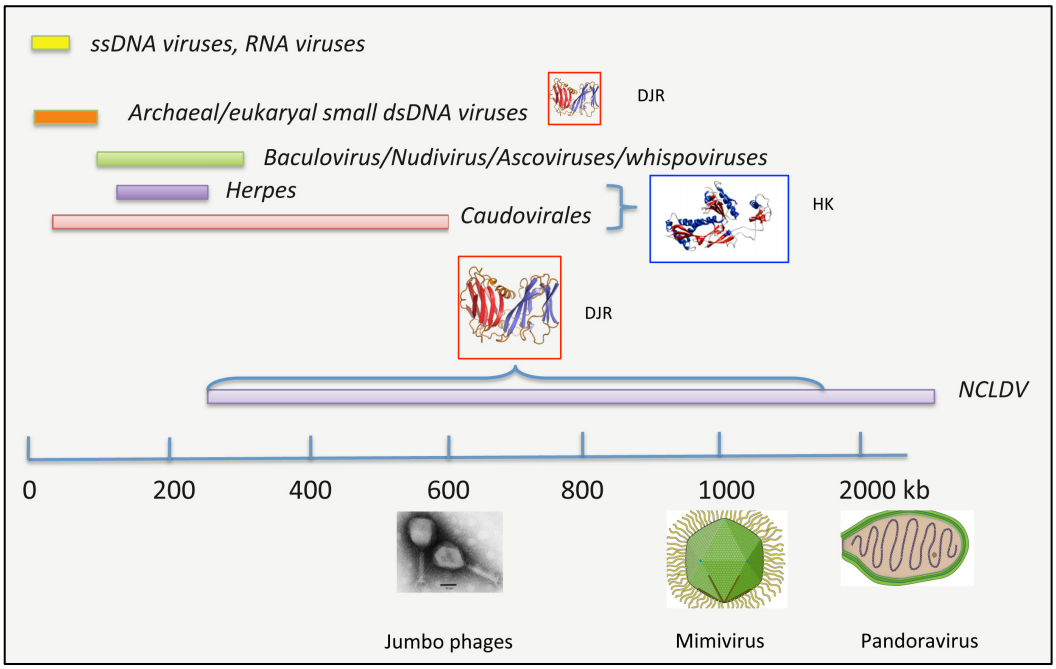

Figure 1: Genome size scale of viruses. DJH: double-jelly roll fold, HK: Hong Kong fold. Drawings and pictures are from ViralZone (*Hulo de Castro* et al., 2012).

WHAT IS A VIRUS?

The discovery of giant viruses, such as Mimivirus and Pandoravirus, has raised new interest in the problem of the nature, origin and role of viruses in life evolution (*Raoult* et al., 2004, *Forterre*, 2010, *Philippe* et al., 2013, *Forterre*, 2017). These viruses produce virions that are visible under the light microscope and have genomes larger than the genomes of some free-living microbes. However, dividing viruses between 'giants' and 'small' viruses is artificial since there is a continuous gradient in genome size between the smallest virus (encoding two genes) and the Pandoravirus, encoding more than 2000 genes (*Forterre* et al., 2014) (Figure 1). The challenge is to find a definition of viruses that takes into account this diversity. Ten years ago, Didier Raoult and myself suggested classifying the living world in two major realms: "*capsid-encoding organisms*" (viruses) and "*ribosome-encoding organisms*" (Archaea, Bacteria and Eukarya) (*Raoult* and *Forterre*, 2008). We proposed the term "*orphan replicons*" for mobile elements such as plasmids, transposons, etc. that are evolutionary related to viruses (capsidless viruses according to *Koonin* and *Dolja*, 2013). Notably, considering the capsid to be the hallmark of the virus allows distinguishing between viruses and orphan replicons. This can be illustrated by comparing the smallest known plasmid encoding one protein (a replication protein) to the smallest known virus that encode two proteins, a replication protein and a capsid protein (*Krupovič* and *Bamford*, 2010) (Figure 2). Importantly capsids should themselves be defined as a set of proteins (at least one) associated to the viral nucleic acid to form a virion.

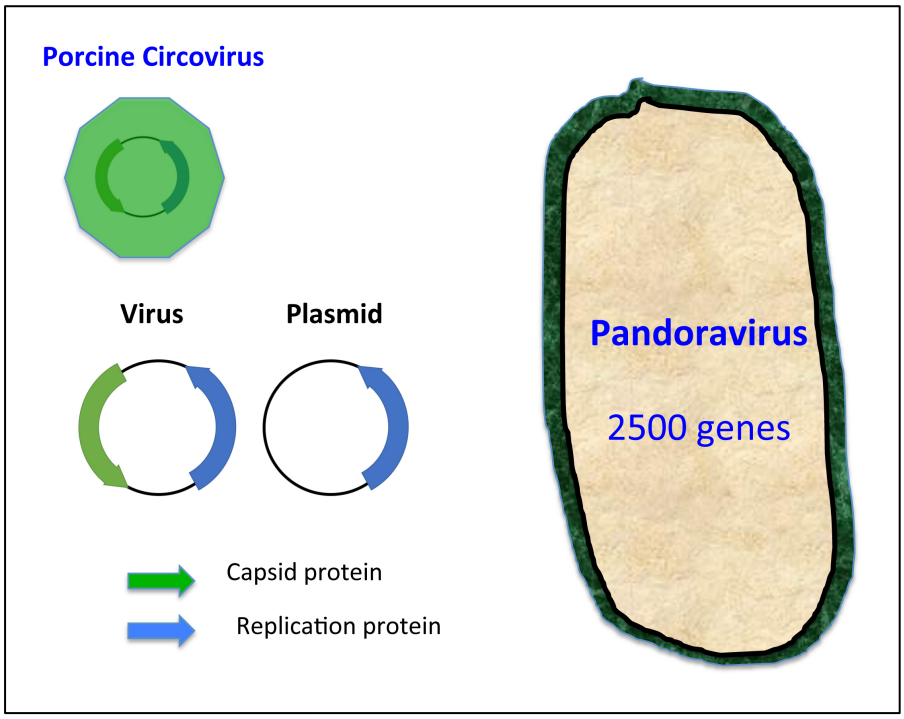

Figure 2: Definition of viruses as capsid encoding organism allows distinguishing viruses and plasmids and is valid for both the smallest and the largest viruses (drawing partly inspired by *Krupovič* and *Bamford*, 2010).

THE ORIGIN OF VIRUSES

The definition of viruses as organisms encoding protein-based capsids implies that viruses originated after the emergence of the ribosome, i.e. after the emergence of rather sophisticated cells (Figure 3). This definition thus clearly refutes all "virus first" theory for the origin of life. On the other hand, comparative analyses of some key viral proteins, have shown that viruses were most probably already present in the biosphere at the time of the last universal common ancestor (LUCA) (*Abrescia* et al., 2012) (Figure 4). Indeed, at least two major lineages of viruses characterized by their specific major capsid proteins (MCP) and packaging ATPases have members in the three ensembles of viruses infecting each of the three domains of life. In addition to the MCP characteristics of these two lineages, many other types of non-homologous MCP have been identified by structural biologists (*Krupovic* and *Koonin*, 2017), confirming that viruses are polyphyletic, and preventing the definition of a viral "LUCA". This indicates that viruses originated several times independently, some of them before LUCA, other possibly later on, by recombination between various replicons and cassettes encoding sets of genes required to make a virion (see for instance *Krupovic*, 2013).

The first viruses were most likely RNA viruses infecting RNA/protein cells that originated from the association of parasitic RNA replicons and simple capsid proteins (*Forterre*, 2006a). Ancient structures present in

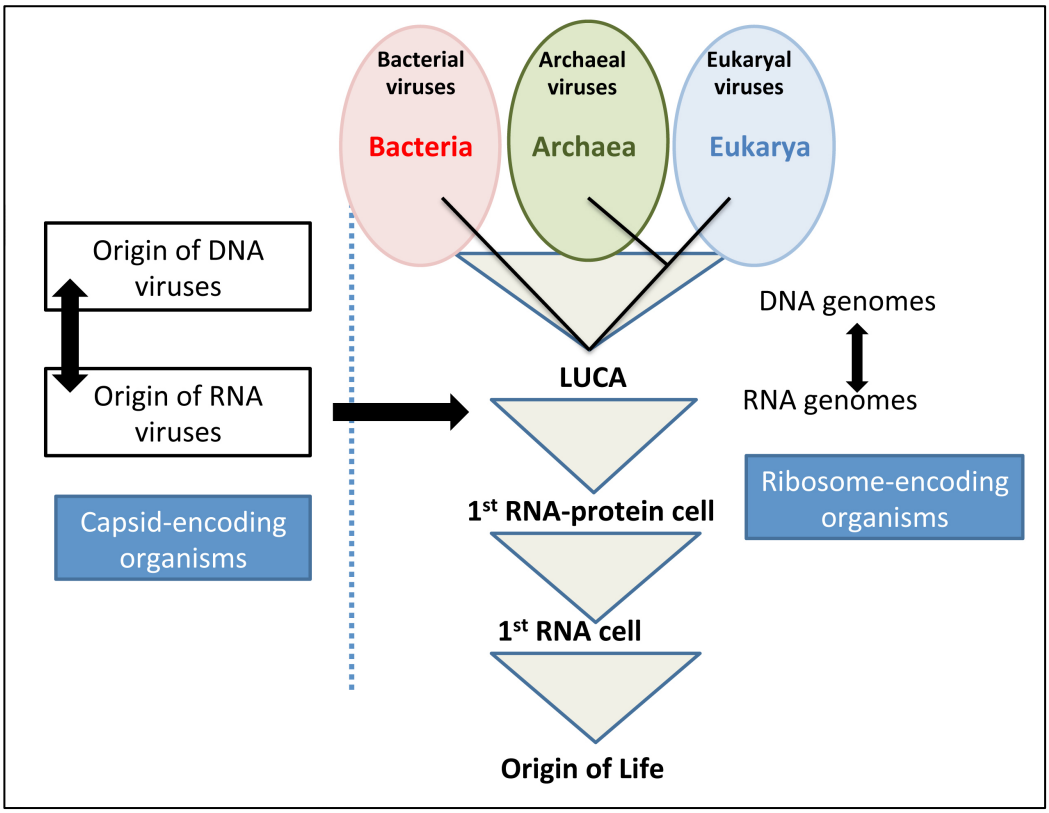

Figure 3: Schematic evolutionary pathway from the origin of life to the modern world of ribosome and capsid encoding organisms. LUCA: the Last Universal Common Ancestor.

these RNA/protein cells, such as membrane vesicles, intracellular compartments, or primitive chromosome scaffolds, may have provided the basis for the emergence of different types of simple virions (pleomorphic vesicle-like virions, icosahedral capsids and nucleocapsids) (*Forterre* and *Krupovic*, 2012). Later on, DNA viruses possibly originated from RNA viruses (Figure 3) and/or from the association of DNA replicons with capsids from RNA viruses. I suggested that DNA itself might have appeared in the ancient virosphere, being originally a particular type of modified RNA genome (the out of virus hypothesis for DNA origin) (*Forterre*, 2002, 2006ab). The early emergence of DNA and DNA replication machineries in such ancient virosphere would explain why these mechanisms are much more diverse in the viral world than in the cellular world (*Forterre*, 2013a). Later on, DNA and two non-homologous viral DNA replication mechanisms would have been transferred to cells, one in the bacterial lineage and the other in the "arkaryal lineage" (Arkarya being the name proposed for the clade grouping Archaea and Eukarya) (*Forterre*, 2015).

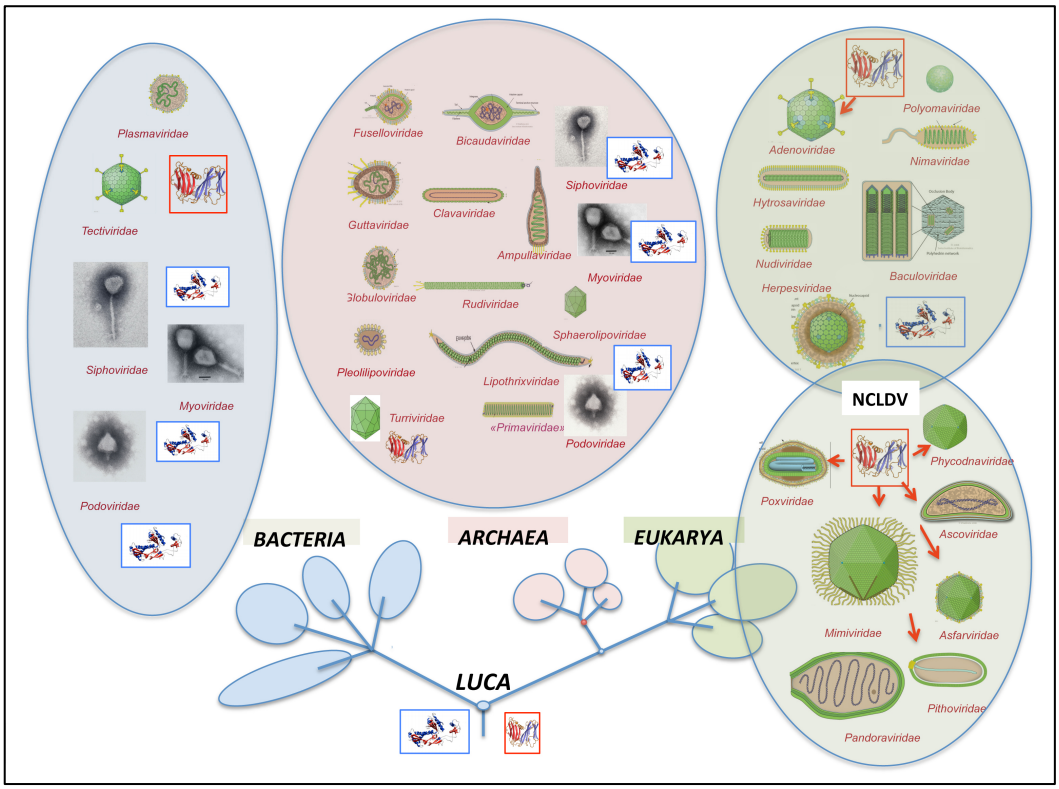

Figure 4: The Universal tree of life and the three ensembles of double-stranded DNA viruses corresponding to each domain (some families) are not indicated in eukaryotes. NCLDV: NucleoCytoplasmic Large DNA Viruses. LUCA: The Last Universal Common Ancestor. The depicted folds correspond to those in Figure 1 (see legend). This picture illustrates the diversity of archaeal dsDNA viruses (*Prangishvili*, 2013) compared to bacterial ones. Caudovirales (Myoviridae, Siphoviridae, Podoviridae) are common to Archaea and Bacteria and represent 95% of known bacterial viruses. Eukaryotes are infected by many different RNA virus families not shown here, whereas Bacteria are also infected by RNA viruses, but much more rarely. Drawings and pictures are from ViralZone (*Hulo de Castro* et al., 2012)

ARE VIRUSES ORGANISMS? NO, ACCORDING TO THE TRADITIONAL VIEW OF VIRUSES ASSIMILATED TO THEIR VIRIONS

Defining virions as "*capsid encoding organisms*" raises the question: are viruses organisms? Usually, viruses are not considered to be organisms because as stated by *Lwoff* (1966) "*an organism is constituted of cells*", whereas viruses are assimilated to macromolecular machines. This is because viruses are usually confused with their virions. The tradition to identify viruses to their virions has both historical and practical reasons **(***Forterre*, 2016). Historically, the assimilation of viruses to their virions can be traced back to the discovery of viruses. The term "virus" was indeed first used to describe the infectious entities able to pass through filters that retain bacteria (*Bos*, 1999) and it turned out that these infectious entities are the viral particles. Later on, images of

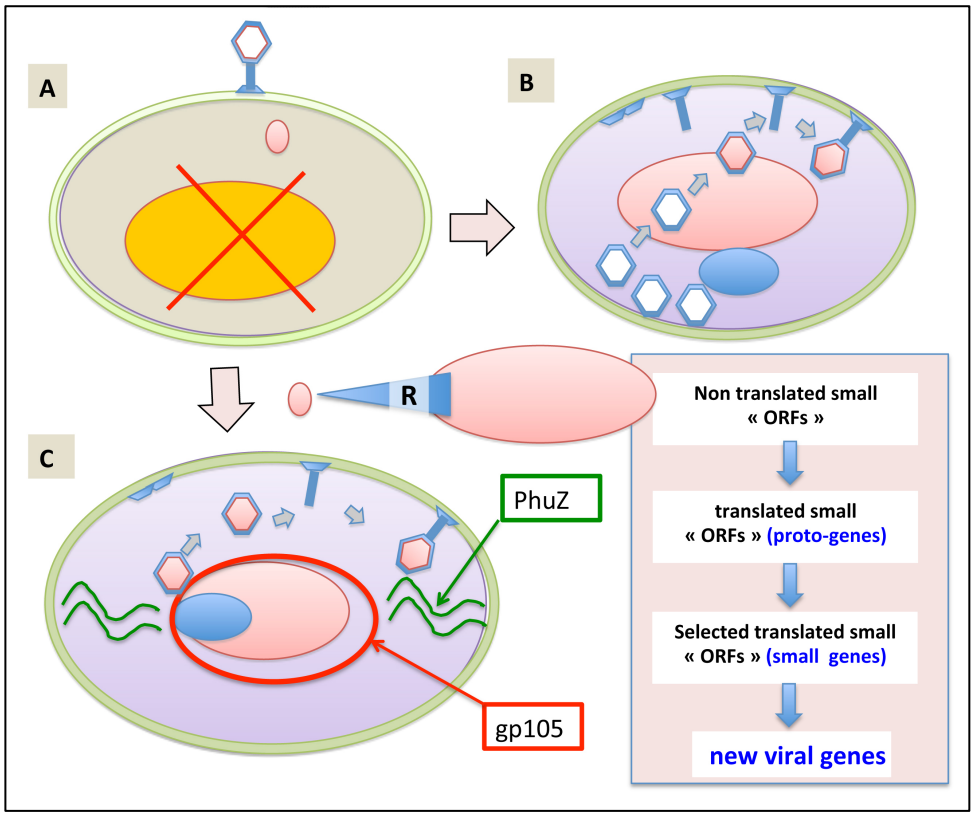

Figure 5: Infection by a virulent bacteriovirus transforms a bacterium (A) into a virocell (B) in which the only present nucleic acid is often the viral DNA (pink) after destruction of the bacterial chromosome (orange). The infection transforms the cellular metabolism and membranes (indicated by the differences in colour). Some bacteriovirus transform the bacterium into a virocell with a nucleus (C). A nuclear membrane is formed by a viral encoded protein (gp105) and the nucleus is positioned at the middle of the cell by a tubuline-like protein (PhuZ) (*Chaikeeratisak* et al. 2017). New genes can originate in the virocell (as in ribocells) during the replication (R) of the viral genome by the mechanism summarized on the lower right panel (*Carvunis* et al., 2012, *Zhao* et al., 2014).

virions have been constantly used to illustrate and popularize the virus concept in publications, textbooks and conferences (as it is the case in Figures 1, 3 and 4 of this paper!).

The assimilation of viruses to their virions had important consequences on the previous definitions of viruses. For instance, *Lwoff* (1957) stated that, in contrast to cells, viruses have only one type of nucleic acid (either RNA or DNA). However, DNA viruses actually possess both DNA and RNA (messenger RNA). Assuming that viruses have only one type of nucleic acid thus clearly means that one identifies the virus and the virion. Another example of this confusion is provided by environmental virologists who always determine the number of viruses in the environment by counting the number of viral-like particles and equal this number to the number of viruses (*Forterre*, 2013b). This is in fact the equivalent to count fish eggs to estimate the number of fishes in the ocean! A striking example can also be found in a seminal review by Jacob and

Wollman in which these authors first described the three possible forms of viruses (including intracellular one) to conclude by defining a virus as "*a genetic element enclosed in a protein coat"* (*Jacob* and *Wollman*, 1961).

VIRUS DEFINITION AND THE ORIGIN OF VIRAL GENES

A damaging consequence of the assimilation of viruses to their virions is that many biologists, especially evolutionists interested in the history of life, underestimate the capacity of viruses to "produce" new genes *de novo*. This is because, once assimilated to their virions, viruses are considered to be passive and inert objects, entirely dependent of their cellular hosts (*Forterre*, 2011). As a consequence, most biologists wrongly assumed that all (or almost all) viral genes are derived from their cellular hosts/victim (*Moreira* and *Lopez-Garcia*, 2009). However, this is probably not correct. One can safely assume that most viral genes originated *de novo* in viral genomes during the intracellular cycle of virus reproduction by the same mechanisms that produce novel genes in cellular genomes (*Forterre*, 2011) (Figure 5). The mechanisms of *de novo* gene emergence have now been revealed by comparative analyses of multiple closely related genomes of *Saccharomyces cerevisiae* and *Drosophila melanogaster* (*Carvunis* et al., 2012, *Zhao* et al., 2014). Most new genes did not originate from gene duplication, as often assumed, but by the selection of short open reading frames (protogenes) arising randomly in intergenic regions. In the case of RNA viruses, such new genes can also originate on the non-coding strand of ancient genes, producing overlapping genes that can be unrelated from one viral strain to another (*Rancurel* et al., 2009). The massive creation of viral genes *de novo* in viral genomes explains well why most of them have no cellular homologues.

The continuous creation of new viral genes *de novo* provides an unlimited resource of new genes for cellular organism, since viral genes can become cellular following the integration of viral genome into cellular chromosomes. Notably, viral genomes can integrate into cellular genomes with few constraints in term of size, whereas the amount of cellular DNA or RNA that a virus can take up is limited by the size of the virion. This explains why the gene flux is overwhelming from viruses to cells than the other way around (*Forterre* and *Prangishvili*, 2013). The integration of viral DNA is a critical mechanism speeding the rate of evolution. There are now many examples of viral proteins whose exaptation has been at the origin of major evolutionary transition in life evolution. Well-known examples are Peg10 and Syncitins which derived from the Gag and Env proteins of an endogenous retrovirus, respectively, and are used by mammals to build the placenta and protect the embryo against the immune system of the mother (*Chen* et al.2015, *Villarreal*, 2016). A less known but also dramatic example is provided by the Arc protein, a master gene of memory acquisition, which is derived from the Gag gene of endogenous retroviruses (*Day* and *Schepherd*, 2015). Many authors now recognize the important role that viruses have played at several critical transitions in life evolution, especially in the emergence of new cellular lineages (*Forterre* and *Prangishvili*, 2009, *Koonin* and *Dolja*, 2013). This is not surprising, considering that viruses are major actors of both variation and selection, the two pillars of Darwinism (*Forterre*, 2012).

ARE VIRUSES ORGANISMS? YES, IF ONE FOCUSES ON THE VIROCELL

The assimilation of viruses to their virions led to an underestimation of the intracellular phase of the virus reproduction cycle, often called the "eclipse phase' (since virions are no more visible) or the vegetative phase. These are unfortunate choices since the intracellular phase is precisely the active phase of the virus reproduction during which the virus indeed behaves as an organism. During this phase, the viral genome is transcribed and replicated and the cellular metabolism is partly or completely reorganized in favour of the virus. The transformation of the host/victim metabolism into a viral metabolism is especially critical for the whole process. For that purpose, the viral genome encodes proteins that redirect the metabolism of its host/victim and/or encodes viral metabolic enzymes complementary to or replacing those of the host cell (*Rosenwasser* et al., 2016).

A few years ago, I proposed the "virocell concept", to focus attention on the intracellular phase of the virus reproduction cycle (*Forterre*, 2011, 2013). In particular, the virocell concept should help taking into account the possible *de novo* origin of most viral genes. However, the virocell concept should not be confused with the definition of viruses because it does not cover the whole process corresponding to the viral organism. The term organism in the definition of viruses describes a biological process and integrates all aspects of the viral reproduction cycle: the virion, the virocells, and the viral genome.

The virocell concept was also proposed to fit with the definition of viruses as capsid encoding ORGANISMS, since the virocell corresponds to a new type of cellular organism (Figure 5B). At the beginning of the infection, two organisms thus co-exist in the same cell, the virus and the ribocell (a bacterium, an archaeon or an eukaryote), fighting each other (in particular via the CRISPR and anti-CRISPR systems). Later on, the two organisms can manage to co-exist pacifically in a form of symbiosis sometimes called carrier state or persistence forming a ribovirocell. However, very often, the virus predominates and the ribocell disappeares, leaving a transient virocell that commit suicide while liberating a wealth of virions. To paraphrase the metaphor from François Jacob: "*the dream of a cell is to produce two cells*", one can say: "*the dream of a virocell is to produce as much virions as possible*". In the ribovirocell, both organisms manage to fulfil their own dream but making a compromise, dividing more slowly for the cell and producing less virions for the virus. The co-existence of different organisms in the same cell is a frequent situation in biology, indicating that one should not confuse the notions of cell and organisms. Many eukaryotic cells harbour a multitude of intracellular bacteria and/or archaea. An amazing example is provided by amoeba infected by a bacterium and a giant virus; the giant virus itself being infected by a virophage (a virus of a virus) (*Moliner* et al., 2010). In that case, four organisms are present in the same cell as in a Russian doll (*Forterre*, 2010, *Mart Krupovic*, personal communication).

By analogy with the eukaryotic nucleus, the viral factories produced by many eukaryotic viruses replicating in the cytoplasm can be considered as the nuclei of these virus virocells. In that case, it is possible that this analogy reflects homology since several authors

have suggested the existence of an evolutionary link between the nucleus of eukaryotic cells and the viral factories (nucleus) of giant DNA viruses infecting eukaryotic cells (*Forterre* and *Gaia*, 2016). Strikingly, it turned out recently that some viruses infecting bacteria can also produce a nucleus in which viral transcription and replication takes place, as well as a simple mitotic-like apparatus to localize this nucleus at the centre of the infected bacterium (*Chaikeeratisak* et al., 2017) (Figure 5C). This again emphasizes the power of viral creativity and increases the appeal of the so-called viral eukaryogenesis hypothesis for the origin of eukaryotes (*Forterre* and *Raoult*, 2017).

ARE VIRUSES ALIVE?

The answer to this question changed frequently depending of the period and the authors. Originally, viruses were often considered to be living because the "infectious fluid" detected by the pioneers of virology displayed all the classical properties of life: reproduction, multiplication and evolution (*Bos*, 2000). Later on, most biologists conclude that viruses are not living when they realized that virions are "simply" inert nucleoprotein particles devoid of metabolism (*Bos*, 2000, *Van Regenmortel*, 2003; *Moreira* and *Lopez-Garcia*, 2009). This was of course another consequence of the assimilation of viruses to their virions. For instance, *Van Regenmortel* (2003), former president of the ICTV, wrote that: "*viruses do not possess many of the essential attributes of living organisms, such as the ability to capture and store free energy and they lack the characteristic autonomy arising from the presence of integrated, metabolic activities*". However, this is not true for the virocell, since the latter is characterized by a specific metabolism working for the benefit of the virus (*Rosenwasser* et al., 2016). It is thus tempting to conclude that viruses are actually living after all. This conclusion then raises further questions. If viruses are living, one can go one step further and ask: are plasmids living? As previously mentioned, the only difference between the smallest plasmid and the smallest virus is the presence in the latter of a gene encoding a capsid protein (Figure 2). Does this mean that addition of a single gene is sufficient to transform a non-living biological object (the plasmid) into a living organism (the virus)?

I recently noticed another example that dramatically illustrates the difficulty to define a living organism, the transition between an intracellular bacterium and an organelle (*Forterre*, 2016). All biologists would agree that intracellular bacteria are living and most of them would also assume that mitochondria are not because "*they lack autonomy and a life cycle*" (*Van Regenmortel*, 2010). However, it is impossible to determine when the transition from living to non-living occurred in the evolution leading from the alphaproteobacterium ancestor of mitochondria to *bona fide* mitochondria. This is because the autonomy of an endosymbiont towards its host decreases in a continuous manner in the course of reductive evolution **(***Forterre*, 2016**).** It is thus hopeless to search for the gene or the gene set that would define life and/or determine the degree of autonomy of an organism. One cannot define a clear-cut border between living and non-living organisms/organelles based on quantitative or qualitative features.

This conclusion raises a challenging question: should we exclude the terms "life" and "living" from the biological literature, since they cannot be rigorously defined scientifically? This seems difficult considering that biology is the science of "life". To solve this conundrum, I suggested considering as living all biological entities as long as they are operational in the process of "life" (*Forterre*, 2016). To discriminate between biological entities that can be living (i.e. a protein or a chromosome) and biological entities that cannot, such as a protein domain or a gene, I have proposed using the philosophical distinction between "individual" and "particular" (*Chauvier*, 2008, *Pradeu*, 2010). Individuals should be "*separable, countable and have acceptable clear-cut spatial boundaries*" (a protein, a chromosome) whereas a particular is "*everything that can be designed through a demonstrative reference*" (a protein domain or a gene) (*Chauvier*, 2008, *Pradeu*, 2010). In these proposals, life can be defined as *"the mode of existence of biological individuals*" (*Forterre*, 2016) to paraphrase the definition proposed by Friedrich Engels in the 19th century "*life is the mode of existence of an albuminoid body"* (*Engels*, 2006 [1883]).

CONCLUSION

Viruses and evolutionary related mobile elements are a major component of the biosphere beside the descendants of LUCA (archaea, bacteria and eukarya). Deciphering the history of the co-evolution between viruses, mobile elements and cells will be a major task of this century. Recently, most efforts in the fields have been made in analysing viromes from various environments. The limitation of this approach is that most sequences in viromes correspond to unknown viruses and data analyses end up focusing on the small set of sequences retrieved from already known viruses. A major effort should be now to isolate new virus-host systems, especially for viruses infecting some understudied organisms that represent most of the biodiversity on earth, such as the various phyla of protists in the eukaryotic domain, or the recently described new archaeal lineages that seem widespread in all types of environments.

ACKNOWLEDGMENT

Patrick Forterre research groups are supported by an ERC grant from the European Union's Seventh Framework Program [340440 Project EVOMOBIL].

LITERATURE

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

Bos, L.: Beijerinck's work on tobacco mosaic virus: historical context and legacy. Philos. Trans. R. Soc. Lond., B Biol. Sci. 354, 675-685 (1999).

Bos L.: 100 years of virology: from vitalism via molecular biology to genetic engineering. Trends Microbiol. 8, 82-87 (2000).

Carvunis A-R., Rolland, T., Wapinski, I., Calderwood, M.A., Yildirim, M.A., Simonis, N., Charloteaux, B., Hidalgo, C.A., Barbette, J., Santhanam, B., Brar, G.A., Weissman, J.S., Regev, A., Thierry-Mieg,

N., Cusick, M.E., and Vidal, M.: Protogenes and de novo gene birth. Nature 487, 370-374 (2012).

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

Chauvier, S.: Particuliers, individus et individuation. In: L'individu: perspectives contemporaines (Eds.: Ludwig P. and Pradeu T.). Librairie Philosophique J. Vrin, Paris: 11-35 (2008).

Chen, H., Sun, M., Liu, J., Tong, C., and Meng, T.: Silencing of Paternally Expressed Gene 10 Inhibits Trophoblast Proliferation and Invasion. PLoS One.10, e0144845 (2015).

Cortez, D., Forterre, P., and Gribaldo, S.: Hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial genomes. Genome Biol. 10, R65 (2009).

Daubin, V. and Ochman, H.: Start-up entities in the origin of new genes. Curr. Opin. Genet. Dev. 14, 616-619 (2004).

Day, C. and Shepherd, J.D.: Arc: building a bridge from viruses to memory. Biochem. J. 469, e1-3 (2015).

Engels, F.: Dialectics of Nature. (Translation of Dialektik der Nature, 1883). Wellred publications, London (2006).

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

Forterre, P.: The origin of viruses and their possible roles in major evolutionary transitions. Virus Res. 117, 5-16 (2006a).

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, 3669-3374 (2006b).

Forterre, P.: Giant viruses: conflicts in revisiting the virus concept. Intervirology 53, 362-378 (2010).

Forterre, P.: Manipulation of cellular syntheses and the nature of viruses: the virocell concept. C.R. Chim. 14, 392-399 (2011).

Forterre, P.: Darwin's goldmine is still open: variation and selection run the world. Front. Cell. Infect. Microbiol. 2, 106 (2012).

Forterre P.: Why are there so many diverse replication machineries? J. Mol. Biol. 425, 4714-4726 (2013a)

Forterre, P.: The virocell concept and environmental microbiology. ISME J. 7, 233-236 (2013b).

Forterre, P.: The universal tree of life: an update. Front. Microbiol. 6, 717 (2015).

Forterre, P.: To be or not to be alive: How recent discoveries challenge the traditional definitions of viruses and life. Stud. Hist. Philos. Biol. Biomed. Sci. 59, 100-108 (2016).

Forterre, P. and Gaïa, M.: Giant viruses and the origin of modern eukaryotes. Curr. Opin. Microbiol. 31, 44-49 (2016).

Forterre, P. and Krupovic, M.: The origin of virions and virocells: the escape hypothesis revisited. In: Viruses: Essential Agents of Life (Ed.: Witzany, G.). Springer Science + Business Media, Dordrecht: 43-60 (2012).

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

Forterre, P. and Prangishvili, D.: The major role of viruses in cellular evolution: facts and hypotheses. Curr. Opin. Virol. 3, 558-565 (2013).

Forterre, P. and Raoult, D.: eLetter: (http://science.sciencemag.org/content/355/6321/194/tab-e-letters) (2017).

Forterre, P., Krupovic, M., and Prangishvili, D.: Cellular domains and viral lineages. Trends Microbiol. 22, 554-558 (2014).

Hulo de Castro, E., Masson, P., Bougueleret, L., Bairoch, A., Xenarios, I., and Le Mercier, P.: ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Res. 39, D576-D582 (2011).

Jacob, F. and Wollman, E.L.: Viruses and genes. Sci. Am. 204, 93-107 (1961).

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

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

Kristensen, D.M., Mushegian, A.R., Dolja, V.V., and Koonin, E.V.: New dimensions of the virus world discovered through metagenomics. Trends Microbiol. 18, 11-19 (2010).

Krupovic, M: Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. Curr. Opin. Virol. 3, 578-586 (2013).

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

Krupovic M, and Koonin E.V.: Multiple origins of viral capsid proteins from cellular ancestors. Proc. Natl. Acad. Sci. USA 114, E2401-E2410 (2017).

Lwoff, A.: The concept of virus. J. Gen. Microbiol.17, 239-253 (1957).

Lwoff A.: Interaction among virus, cell, and organism. Science 152, 1216-1220 (1966

Moliner, C., Fournier, P.E., and Raoult, D.: Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol. Rev. 34, 281-294 (2010).

Moreira, D. and López-García, P.: Ten reasons to exclude viruses from the tree of life. Nat. Rev. Microbiol. 7, 306-311 (2009).

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

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

Pradeu, T.: What is an organism? An immunological answer. Hist. Phil. Life Sci. 32, 247-268 (2010).

Rancurel, C., Khosravi, M., Dunker, A.K., Romero, P.R., and Karlin, D.: Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. J. Virol. 83, 10719-10736 (2009).

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

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

Rosenwasser, S., Ziv, C., Creveld, S.G., and Vardi, A.: Virocell Metabolism: Metabolic Innovations During Host-Virus Interactions in the Ocean. Trends Microbiol. 24, 821-832 (2016).

Suttle, C.A.: Viruses: unlocking the greatest biodiversity on Earth. Genome 56, 542-544 (2013).

Van Regenmortel, M.H.: Viruses are real, virus species are man-made, taxonomic constructions. Arch. Virol. 148, 2481-2488 (2003).

Van Regenmortel, M.H.: Logical puzzles and scientific controversies: the nature of species, viruses and living organisms. Syst. Appl. Microbiol. 33, 1-6 (2010).

Villarreal, L.P.: Viruses and the placenta: the essential virus first view. APMIS 124, 20-30 (2016).

Zhao, L., Saelao, P., Jones, C.D., and Begun, D.J.: Origin and spread of de novo genes in Drosophila melanogaster populations. Science 343, 769-772 (2014).