Studies in History and Philosophy of Biological and Biomedical Sciences 59 (2016) 89–99

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Giant viruses: The difficult breaking of multiple epistemological barriers☆

Jean-Michel Claverie [a,b,*], Chantal Abergel [a]

[a] Structural and Genomic Information Laboratory, UMR 7256 (IMM FR 3479) CNRS Aix-Marseille Université, Luminy Campus, Marseille, France
[b] Assistance Publique des Hôpitaux de Marseille, La Timone, 13005, Marseille, France

ARTICLE INFO

Article history:
Received 21 January 2016
Accepted 27 February 2016
Available online 10 March 2016

Keywords:
Bachelard
Epistemological obstacles
Reductive evolution
Virus definition
Virus origin

ABSTRACT

The discovery of the first "giant virus", Mimivirus, in 2003 could solely have been that of an exceptional freak, a blind alley of evolution as occasionally encountered in biology, albeit without conceptual significance. On the contrary, once broken this epistemological barrier, additional unrelated families of giant viruses such as the Pandoraviruses, the Pithoviruses and most recently Mollivirus, were quickly unraveled, suggesting that an entire chapter of microbiology had been ignored since Pasteur and Ivanovski. In this article, we examine to what extent the giant viruses challenge previous definitions of viruses, the diversity of forms they could take, and how they might have evolved from extinct ancestral cellular lineages. Inspired by the epistemology of Gaston Bachelard, we will also suggest the reasons for which giant viruses laid hidden in plain sight for more than a century. Finally, we propose a new definition for "viruses" that paradoxically emphasize the fact that they do not encode a single universally shared macromolecule or biochemical function.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Biological and Biomedical Sciences*

1. Introduction: Giant viruses were not meant to be discovered by virologists

The recent discovery of a whole diversity of giant viruses in quick succession (Arslan, Legendre, Seltzer, Abergel, & Claverie, 2011; Legendre et al., 2014, 2015; Philippe et al., 2013; Raoult et al., 2004) demonstrated that biology is still a scientific area where some of the most established concepts might be proven wrong, or seriously misleading, even though there were not challenged for more than a century. It also reminded us of the danger of solely envisioning (and funding) biological research in the context of biomedical, economical or societal challenges. The giant viruses that we know today do not cause any harm to humans or animals, and do not destroy crops, the three main incentives that guided the development of virology (Helvoort, 1996) since its very beginning with the isolation of the Tobacco mosaic disease virus (Ivanovski, 1892). Such utilitarian attitude was actually reinforced by a basic technical reason: studying viral diseases provided the researchers both with the virus and its host at once, a *sine qua non* condition to study and propagate such obligatory parasites unable to multiply outside specific cells. Ironically, soon after the serendipitous discovery of the first - totally innocuous- giant virus, it has become clear that their marine relatives played an essential role in regulating the populations of unicellular plankton the equilibrium of which depends on half of the oxygen production and carbon

☆ This article challenges the traditional notion of "virus". In order to describe how this notion should be changed in the light of recent discoveries, we needed to start from the traditional sense of the word, as accepted by most contemporary biologists. For the majority of them, "virus" designates the inert infectious particle which uses the biosynthetic capacity of the cell it infects to multiply. Instead, we will propose that the essence of a virus is the intracellular process akin to the development of a transient microorganism. There is therefore an unavoidable ambiguity each time the word "virus" is used in this article, depending on whether we are meaning the "particle", or the whole process. For instance, given the central role that the small size of viruses played in the history (and the methods) of virology, "giant virus" is definitely an oxymoron in the traditional meaning of the word. The same expression is devoid of sense when using our own new definition.

* Corresponding author. Structural and Genomic Information Laboratory, UMR 7256 (IMM FR 3479) CNRS Aix-Marseille Université, Luminy Campus, Marseille, France.
E-mail address: Jean-Michel.Claverie@univ-amu.fr (J.-M. Claverie).

http://dx.doi.org/10.1016/j.shpsc.2016.02.015

sequestration on our planet (Fischer, Allen, Wilson, & Suttle, 2010; Weitz et al., 2015), one of the most pressing societal challenge of our time.

Following a quick historical account of the discovery of giant viruses and a description of their unusual properties, we will examine the well-accepted views that they appear to challenge. We will more specifically develop three main topics: the new notion that viruses exhibit a gradation in their "absolute" parasitism, the increasingly blurred frontier between the viral and cellular worlds, and how viruses might have evolved to their present diversity. Along the text, we will point out the various kinds of epistemological obstacles (*sensu* Bachelard) that precluded the discovery and the recognition of the viral nature of giant viruses. Finally, we will propose a new definition for "viruses" that we hope (but do not really expect) could stand the test of time and remain applicable to the increasingly exotic types of microorganisms that remain to be discovered.

2. The traditional concept of virus

2.1. "Virus" as a failed microbe

The germ theory of diseases often hailed as the most important work of Louis Pasteur (although initiated by Lister and refined by Koch) paradoxically set up the stage for the discovery of viruses. In front of the French Academy of Medicine, Pasteur proposed in 1878 that infectious diseases were caused by the proliferation of specific – living – microorganisms, visible under the light microscope and cultivable on a nutritious broth (Pasteur, 1878a,b; Pasteur et al. 1878). Few years later Charles Chamberland designed a porcelain filter capable of retaining these microbes, thus providing the first straightforward experimental protocol to rapidly demonstrate the microbial nature of any infectious agent (Chamberland, 1884). Ironically, the year of Pasteur's Jubilee (1892) celebrating his lifelong accomplishments, Dimitry Ivanovski, a young Russian botanist at the beginning of his career, poked the first hole in the newly established paradigm by showing that the agent transmitting the highly contagious Tobacco mosaic disease was not retained by the Chamberland filter, neither could be seen under the microscope, nor could it be cultivated in traditional growth media (Ivanovski, 1892).

Retrospectively, it was very fortunate that this unambiguous falsification (*sensu* Karl Popper) of the barely established theory of Louis Pasteur did not resurrect the fallacious miasma theory which states that contagious diseases are communicated by corrupted air. Instead, following the confirmation of Ivanovski's experiment by Martinus Beijerinck (Beijerinck, 1898), the unexpected filterability of the tobacco mosaic disease agent triggered the emergence of the concept of "virus" as qualitatively different from the usual microbes (i.e. bacteria). Yet, Beijerinck's definition of the new "virus" as a non-corpuscular living fluid ("contagium vivum fluidum") was more of a regression than a progress, uncomfortably close to the antique acceptance of the word "virus" designating anything from stench, poison, or a viscous secretion. Following this nebulous start, the notion of "filterable virus" remained enigmatic until the first electron microscope images of Tobacco mosaic viruses (TMV) were produced in 1939 (Kausche, Pfankuch, & Ruska, 1939).

2.2. Awaiting for the "modern" definition of viruses

Beijerinck's views were so opposed to the prevalent ideas of the time that they did not receive much attention. Already in 1903, Roux challenged the "fluid contagiosum" hypothesis by dubbing it "very original", and considered these filterable agents as not different from the tiny mycoplasma cells he just discovered (Roux, 1903). However, Chamberland's filtering protocol led to the rapid discovery of many other "filterable" viruses. By 1931, nearly two dozen diseases had already been associated with viruses, including yellow fever, rabies, fowl pox, and foot-and-mouth disease in cattle (reviewed in Helvoort, 1996). Yet, the nature of these "filterable viruses" remained elusive, with competing hypotheses ranging from replicating molecules (proteins) to small intracellular parasitic bacteria such as Rickettsia. Until 1950, viruses continued to be defined by three negative properties: they were invisible under the light microscope, they were uncultivable in absence of living cells, and they were not retained by Chamberland's filter (on the use of filtration as a criterion for being a virus and on the related "negative" definition of viruses, see Méthot, 2016). Later in that period, it was realized that viruses did not multiply by binary fission, and that their multiplication within the infected cell was preceded by an "eclipse" phase, during which traces of them were no longer visible. This apparent lack of "organismal" continuity, as well as the - epistemologically unfortunate - crystallization of TMV by Wendell Stanley in 1935 (whom received the 1946 Nobel Prize in Chemistry – not Physiology/Medicine- for his work) weighted a lot in relegating the viruses outside of mainstream microbiology as far as considering them outside of the living world, an opinion still shared by many modern biologists and the general public (about the status of viruses as "alive or not", see Forterre, 2016).

2.3. Lwoff's criteria to discriminate viruses from cells

The study of bacteriophages (i.e. viruses infecting bacteria) and his special taste and talent for rigorous conceptual thinking, led André Lwoff to provide the first formal definition of viruses or, more exactly, a list of properties to be used to ***discriminate*** them from cellular microorganisms (Lwoff, 1957), as follows:

1) typical microorganisms contain both DNA and RNA, viruses contain only one type;
2) all microorganisms are reproduced from the integrated sum of their constituents while viruses are produced from their nucleic acid only;
3) during the growth of a microorganism the individuality of the whole is maintained and culminates in binary fission. There is no binary fission in viruses;
4) viruses lack the system of enzymes which convert the potential energy of foodstuffs into the energy needed for biochemical syntheses (at that time called the "Lipmann system") that is present in cellular microorganisms.

Following the discovery of the ribosome, one more discriminative criterion was added (Lwoff & Tournier, 1966):

5) viruses make use of the translation machinery of their host cells.

These last two criteria (#4 and #5) make the virus an absolute parasite of its cellular host. Note that criterion #1, proposed before mRNA had been discovered, simply reflected the absence of ribosomal RNA (>80% of the cellular RNA), hence is nowadays redundant with criterion #5.

With his list of well-thought and carefully designed binary criteria, André Lwoff not only provided a rigorous and operational way to discriminate viruses from cells, while forcibly affirming his view that an infectious agent could not be intermediate between viruses and nonviruses, a possibility entertained by few microbiologists of his time, to his great irritation (page 46, Lwoff & Tournier, 1966). After fifty years of holding tight, the broadly accepted dichotomy between the viral and the cellular world appeared to be

challenged by the discovery of giant viruses. The next sections will examine to what extent this is true.

3. Size and shape as the first epistemological barriers

3.1. Mimivirus, the first virus easily seen under a light microscope

Although the size criterion was at the very origin of the discovery of viruses, it was not part of Lwoff's classification scheme. Within the context of his binary thinking, that could only be done by imposing a threshold (e.g. smaller than 0.3 μm) that he rightly thought would be arbitrary. Lwoff was also aware of the existence of tiny bacteria such as mycoplasma, the small size and elasticity of which allows them to pass through the usual "sterilizing" Chamberland "F" filter (Roux, 1903). Yet, he noted that the (small) size of viruses "was correlated to some of their essential properties". Although absent from Lwoff's formal criteria, size kept an operational value for the isolation of viruses until today: infectious agents visible under a regular light microscope and retained by the above filter, could not belong to the viral fraction of a sample. This unwarranted generalization constituted a physical barrier (giant viruses are retained by the filter) that delayed the discovery of giant viruses that are quite abundant in aquatic environments (Monier, Claverie, & Ogata, 2008). But it also turned into an unconscious epistemological barrier (*sensu* Bachelard)[1] as shown by the circumstances of the discovery of Mimivirus, the first recognized "giant" virus. Mimivirus was initially spotted in 1992 by Dr. Tim Rowbotham, interpreted as an intracellular parasitic bacterium (a *Legionella*-like amoebal pathogens) and accordingly called "Bradfordcoccus" (reported in Raoult, La Scola, & Birtles, 2007). Despite numerous unsuccessful cultivation and characterization attempts, the deeply rooted size criteria delayed the recognition of its viral nature for 12 years (La Scola et al., 2003),[2] as was also the case for two of the three other giant viruses more recently discovered (see below).

Following into the breach, many viruses from the same family (the Megaviridae), have since then been identified in various environments (Arslan et al., 2011). All of them are propagated as icosahedral particles, approximately 0.7 μm in diameter, making them easily visible as small spherical "cocci-like" microbes under a regular light microscope. Nobody nowadays dispute the fact that they are *bona fide* viruses, even though their unusually large DNA genome and gene content at first appeared to weaken Lwoff's non-negotiable principle of non-continuity between viruses and cells. We will address to what extent in later sections.

3.2. Pandoravirus and pithovirus: when shape adds to the confusion

The geometrically regular shape (an icosahedron) of the Mimivirus particle was the main hint that it could be a virus, prompting the follow-up studies that confirmed it. Yet, nothing limits the shape of virus particles to regular, symmetrical appearances. The dreadful Ebola virus filamentous shape, the rod-like helicoïdal TMV, or the oval, rounded brick form of the smallpox virus are good illustrations of the fact that not all viral particles (also called "virions") are shaped like regular polyhedrons. Yet, the large predominance of icosahedral particles among the smallest pathogenic viruses as well as bacteriophages polluted the mind of many microbiologists (again victims of Bachelard's misleading "connaissance générale"). The regular geometrical shape of virus particles was rationalized as a consequence of their "simplicity" (i.e. small gene content), allowing only a few proteins to be devoted to the making of a symmetrical virus "box" through a spontaneous self-assembly process. This is actually true for many "classical" viruses the particles of which are made of multiple copies of a major "capsid" protein assembled into an icosahedron, the regular polyhedron that requires the minimal surface (or minimal number of surface components) for a given enclosed volume. As a consequence of the rigid symmetry governing their self-assembly, the resulting virus particles are identical to each other, allowing some of them to form crystals. This property, usually associated to minerals, did not weight in favor of classifying viruses among the living world.

Following the well-publicized discovery of Mimivirus in 2003, new giant viruses would have been expected to be recognized quickly among already known intra-cellular parasitic microbes resisting traditional cultivation attempts. But here again, unwarranted generalizations[3] about the symmetrical and reproducible shape of known viruses seem to have delayed the research process. For instance, a literature search performed after the discovery of the amphora-shaped, 1 μm-long, Pandoravirus (Philippe et al., 2013), revealed that it had been spotted as early as 2008 (Scheid, Zöller, Pressmar, Richard, & Michel, 2008). Despite noticing that this "obligate intracellular microorganism, only proliferating within amoebic hosts … didn't grow on five different nutrient media plates (suitable to grow bacteria and fungi)",[4] these authors could not jump over the epistemological barrier of Pandoravirus' unusual look to postulate that it was a new giant virus (Claverie & Abergel, 2015; Scheid, Hauröder, & Michel, 2010).

Amazingly, the same scenario was reenacted for Pithovirus, the first representative of a third family of giant virus with an amphora-shape particle of even larger dimension (1.2 μm in length, 0.5 micron wide) that was described in 2014 (Legendre et al., 2014). While scanning the literature, we discovered that an "endocytobiont" with extremely similar characteristics had been described more than 15 years before, with the authors proposing all kinds of hypotheses except that it could be a virus (Hoffmann, Michel, Muller, & Schmid, 1998)!

The circumstances of the discovery of giant viruses nicely illustrate how the history of a discipline and the building of a fertile - albeit faulty- paradigm, erected an unconscious epistemological barrier. The first viruses were discovered because they were filterable. Soon after however, it was shown that the bacterial agent of Bovine peripneumonia (nowadays known as *Mycoplasma mycoides*) could pass through the Chamberland "F" filter and remain invisible under the light microscope (Nocard & Roux, 1898; Roux, 1903). This immediately demonstrated that some bacteria could be as small as viruses. Yet, nobody could apparently conceive the symmetrical case where some viruses may be as big as bacteria

[1] According to Bachelard, this illustrates the misleading influence of a combination of two different types of epistemological obstacles: "l'expérience première" (the initial experience) (Bachelard, 1993; chapter 2) of Ivanovski discovering a totally invisible infectious agent, and the "connaissance générale" (premature generalization) (Bachelard, op. cit., Chapter 3) erecting as a rule that all subsequent "viruses" had to be similarly tiny and unstoppable by sterilizing filters. Understandably, the first fifty successful years of virology only dealt with "viruses" defined as such by this property.

[2] This delay was clearly not due to a lack of available techniques, as the genome of many large DNA viruses had been sequenced at that time (since Baer et al., 1984) and Rowbotham's laboratory already mastered the genomic identification of other intracellular "microbes" (Fry, Rowbotham, Saunders, & Embley, 1991).

[3] The hurdle of the "connaissance générale" (Bachelard, op. cit.; chapter 3). This is again the barrier of "undue generalization": not only all previously known virus particles where invisible under a light microscope, but most of them (including that of Mimivirus) exhibited regular, symmetrical (often icosahedral) particles.

[4] Words evocating Pasteur's definition of germs and, as a negation, that of a virus.

and be retained by these filters. Giant viruses have been in front of our eyes for more than 20 years before being recognized as such. This probably would not have happened if the first virus ever discovered had been a large poxvirus rather than the tiny one causing a Tobacco disease.

4. Conceptual problems raised by the diversity of viral genomes

As of today, giant viruses (e.g. those visible under the light microscope) have been shown to belong to four distinct families. By order of discovery they are the Megaviridae (Arslan et al., 2011; Raoult et al., 2004), the Pandoraviridae (Antwerpen et al., 2015; Philippe et al., 2013), Pithovirus (Legendre et al., 2014), and Mollivirus (Legendre et al. 2015). The Megaviridae propagate their DNA genome of up to 1.2 million nucleotides (encoding about 1000 proteins) in pseudo-icosahedral particles approximately 0.7 micron in diameter. The Pandoraviridae propagate their DNA genome of up to 2.8 million nucleotides (encoding up to 2,550 proteins) in amphora-shaped particles approximately one micron in length and 0.5 micron in diameter. Pithovirus also exhibits amphora-shaped particles, although differing from those of Pandoraviridae and slightly elongated (1.5 micron in length). However, the Pithovirus smaller DNA genome is made of 600,000 bp and only encodes 470 proteins. Finally, Mollivirus packs a genome of 610,000 bp (coding for 520 proteins) in a roughly spherical particle 0.6 μm in diameter. These discrepant numbers already indicate that the genome size of giant viruses does not correlate with the volume of their respective particles, in contrast with the broad correlation observed for "regular" viruses with particles in the 50 nm–300 nm diameter range. In that respect, giant viruses are more like cells, the morphology and size of which are not at all linked to the size and complexity of the genome they contain. The variability in genome size and gene content among giant viruses already portend that discriminating viruses from cell on the basis of a genome-derived feature might prove unreliable. Part of the difficulty arises from the discovery of intracellular parasitic bacteria with amazingly reduced genomes (López-Madrigal, Latorre, Porcar, Moya, & Gil, 2011; Nakabachi et al., 2006). This section will discuss some of the issues raised by the extreme diversity exhibited by the gene contents of viral and cellular genomes.

4.1. Too many genes for a small box: the virus is not the virion

Given the key role played by the size of viral particles (the "virion") in the discovery of viruses as a new type of infectious agent, it was natural that no distinction was initially made between "virus" and "virion". This confusion was then perpetuated by the systematic use of ultrafiltration for the isolation of new viruses and latter for measuring their environmental diversity (e.g. using metagenomics). André Lwoff was probably the first to make an explicit distinction between these two concepts, when he recommended that "the definition of a bacteriophage should not be centered on the infectious particle". However, even himself ambiguously used the term "virus" in his discussion of the two fundamental questions: "are virus organisms?" and "are virus alive?", both questions to which he answered "no" (Lwoff, 1957). The confusion between virus and virion is constant in the media but is also rampant in the scientific literature, in particular when debating the position of viruses in the living world (see Forterre, 2016, for a detailed discussion of the virus/virion paradigm). The crystallization of the tobacco mosaic disease virions (i.e. of the inert particles) partly contributed to the view that viruses should not be considered alive, despite their capacity to evolve and self-reproduce (two hallmarks of living entities). Indeed, viruses (as particles) do not exhibit any metabolic activity, a situation *a priori* incompatible with life.

In the mind of the general public and most biologists (see again Forterre, 2016, for numerous examples; on this question, see also Van Regenmortel, 2016), the word "virus" designates the filterable object made of a small number of self-associating proteins and transporting the viral genome. As viruses are – by definition- absolute intracellular parasites of their cellular host, their genome is in principle only required to encode the structural proteins forming the "virion" (also called "capsid" or "particle"), and a few regulatory functions necessary to "highjack" the cell machinery. The rest of the many intracellular functions required to multiply the virus particles can all be provided by the host, and encoded in the host's genome.

Accordingly, many of the viruses we know, including highly pathogenic ones, fit this paradigm and exhibit small virions made of a handful of proteins (such as a major capsid protein and a "core" protein) packaging a genome just large enough to encode the structural components of the particle, plus a few hijacking "weapons" (proteins or non-coding RNAs) interfering with the cell metabolism. This is the case for many virus families such as the Polyomaviridae (5 kb DNA genomes), the Papillomaviridae (7 kb DNA genomes), the Circoviridae (2 kb DNA genome), the Geminiviridae (5 kb DNA genome), the Parvoviridae (5 kb DNA genome), but also the Picornaviridae (7 kb RNA genome, one of which cause poliomyelitis), the Hepadnaviridae (3 kb RNA genome, causing hepatitis) and the Retroviridae (7–10 kb RNA genome, including the dreadful HIV1 causing AIDS). Such minimal viruses are found everywhere and infect the entire spectrum of cellular organisms (unicellular protists, plants, animals, and prokaryotes). As they have been so successful in evolution, the existence of giant viruses with highly complex genomes becomes all the more paradoxical.

Although viruses with intermediate genome sizes and gene contents had been known for sometimes (such as the virus causing smallpox or many bacteriophages which encode from forty to several hundred proteins), it is the shock of the discovery of giant viruses possessing more genes than many cellular microorganisms that made us realize that the concept of virus should not be restricted to that of virion, for two main reasons: 1) the genome of these viruses contains much more than the blueprint of the particle, and 2) the properties exhibited by giant viruses while multiplying within the cell might call for a reappraisal of their nonliving status as well as of the place they should occupy in the "Tree of Life" (Claverie, 2006; Claverie & Abergel, 2009, 2010; Forterre, 2013; Moreira & Lopez-Garcia, 2009; Nasir, Kim, & Caetano-Anolles, 2012; Raoult et al. 2004; Villarreal & Witzany, 2010).

Following the discovery that the Mimivirus genome encoded thousand proteins, it became evident that such a large gene content, disproportionate to the quantity of information required to specify a box (the virion), was in relation to the "virus", seen as a comprehensive intracellular "process" leading to the development of a transient parasitic microorganism within the host cell. A description of the replicative cycle of the Megaviridae will help illustrate this fundamental shift in perspective.

The delivery of the Mimivirus particle inner core within the host cytoplasm is followed by its rapid development into a spectacular intracytoplasmic "virion factory" (Claverie & Abergel, 2009; Mutsafi, Zauberman, Sabanay, & Minsky, 2010). This factory resembles an organelle created *de novo* to become the site of translation, transcription, as well as replication of the viral genome, using the host's pool of metabolic precursors. It is important to notice that this organelle is built from the genetic information provided by the virus and functions independently of the nucleus

that is simply pushed aside at the periphery of the host cell.[5] At this stage, the resemblance with an intracellular bacterial infection is clear. The numerous cellular and biochemical functions involved in rebuilding from scratch this transient microorganism upon each infection cycle is what justifies the complexity of the giant virus genome.

In 2006 we thus proposed that the intracellular virion factory should be recognized as the actual "virus" whereas the virion (the particle) should be reappraised as a mere vehicle by which the virus (and its genome) is multiplied and propagated from cells to cells (Claverie, 2006; Claverie & Abergel, 2010). In other words, the (inert) virion is to the virus what the (inert) seed is to the adult plant. Thus, the genome of giant viruses is not commensurate to the virion, but to the numerous functions that the virus must emulate once temporarily "alive" in the host cell. Making a clear distinction between the virion and the virus in its intra-cellular active stage immediately reinstates all viruses (not only the giant ones) within the living world, solving at once the paradox cited earlier. During the "vegetative phase" (the disconcerting wording used by Lwoff (Lwoff, 1957) to denote the actively replicating bacteriophage in its host) viruses definitely exhibit all the properties of a (parasitic) microorganism and are therefore "alive", while the virion is not. Although such interpretation is relatively common today (e.g. Dupré & O'Malley, 2009; Dupré & Guttinger, 2016), it is rarely acknowledged that it was actually proposed by Arthur Edwin Boycott as early as 1928 (Boycott, 1928), more than ten years before the first sighting of a virus particle under the newly invented electron microscope (Kausche et al., 1939). Making a clear distinction between the virion and the virus as concept will also be central to the general definition we propose in Section 5.3.

Coincidentally, the notion that the Mimivirus virion factory should be considered a transient microorganism received additional support with the discovery that they could be ''infected'' by their own virus, called a "virophage" (La Scola et al., 2008). At odds with previously described defective/satellite viruses, virophages do not have any host–pathogen relationship with the Acanthamoeba host cells but transcribe and replicate their DNA genome within the Mimivirus virion factory once fully deployed (Claverie & Abergel, 2009). This new kind of parasitism might be a common feature among large DNA viruses infecting eukaryotes, as other virophages were recently discovered associated with other less closely related members of the Megaviridae (Fischer & Suttle, 2011; Santini et al., 2013).

In conclusion, the particles are no more representative of a (giant) virus, than a seed is to a plant, albeit they do exhibit the *exact same genome*. In this new conceptual framework, the finding that some viral genomes may be as big and as complex as that of a parasitic cellular organism becomes no longer paradoxical.[6] It is also a warning that a general definition of viruses should not place too much emphasis on their gene contents that exhibit a tremendous variability.

4.2. Different degrees of "absolute parasitism": a quantitative view of life

Making a clear distinction between the "virions" (the metabolically inert particles) and the "virus" (seen as the process encompassing the succession of intracellular stages leading to the production of new infectious particles) opens more abstract ways to think about viruses, disregarding morphological criteria and, to a certain extent, genome sizes, gene contents and phylogenetic relationships. This abstraction may guide us toward a definition of all viruses (extinct, known, and yet to be discovered) more resilient than one built from the features of the sole viruses known today. We feel that an important step in that abstraction process is to recognize that viruses exhibit vast differences in their dependency *vis-à-vis* the host cell they infect although they all equally qualify as *bona fide* absolute cellular parasites (*sensu* Lwoff). At first, the above sentence may seem paradoxical as the notion of "absolute parasitism" is an all or none notion incompatible with a quantitative scale. In this section, we hope to convince the reader that the paradox is only apparent and that its resolution leads to a plausible scenario for the evolution of viruses and the extent of their diversity.

According to the current paradigm, "life" is a property of the cell and requires the coordinated operation of five well individualized subsystems: the cell division apparatus, the genome replication apparatus, the genome transcription (from DNA to mRNA) apparatus, the protein translation apparatus, and a metabolism both in charge of providing the above subsystems with biochemical building blocks (amino-acids, nucleotides) and the energy required to perform biosynthesis (e.g. ATP).

However, comprehensive metabolisms are rarely found encoded in the genome of a unique microorganism but often requires the collaboration of several of them each bringing its own specific savoir-faire.[7] Photosynthetic microorganisms and plants constitute a fundamental exception as they literally can live of sun, thin air, water, and a few minerals. They constitute the basis of life on our planet ("the bottom of the food chain") and are called "autotrophs", by opposition to "heterotrophs" that cannot live without some help from others. More exotic autotrophs also consist of bacteria capable of metabolizing mineral carbon sources such as carbon dioxide and methane. The metabolic collaboration required by heterotrophs is reached through various modes of association ranging from the acquisition of organic compounds from "dead" matter (food) or through symbiosis or parasitism. In the extreme cases, the parasitic (endosymbiotic) microorganism can only live within its cellular host, most often in its cytoplasm, but examples are known of exotic bacteria confined to the nucleus (Schulz et al., 2014; Zielinski et al., 2009).

The residual metabolic pathways encoded by these obligate intracellular parasites exhibit various combinations of defects in the biosynthesis of amino-acids, nucleotides, or membranes, as well as in energy production (e.g. ATP biosynthesis). Some of the most defective endosymbionts/parasites hardly have any recognizable metabolism left (López-Madrigal et al., 2011; Nakabachi et al., 2006; McCutcheon & Moran, 2011). At this point it becomes tempting to rank cellular microorganisms according to their level of (remaining) autonomy on a scale along which some would be considered more "alive" than others. Life, usually a binary concept (an organism is dead or alive), is then shifted from its qualitative meaning to a quantitative one (see also Koonin & Starokadomskyy, 2016; Kostyrka, 2016). An objective gradation could be made from the "lesser alive" to the "fully living"

[5] Interestingly, the "virion factory" was not recognized as such in the original publication where it was confused with (and mislabeled as) the cell "nucleus" (see Fig. S1 in La Scola et al., 2003).

[6] But still illustrates the capacity of evolution to design amazingly different solutions to the same problem: here propagating a DNA genome in a (small) macromolecular container using an absolute intracellular parasitic "lifestyle". Pandoraviruses and Polyomaviruses (although belonging to the same dsDNA virus class in the Baltimore classification) exhibit a ratio of 8,000 between their virion volumes and a ratio of 510 between their gene numbers (approximately as Human *vs. Escherichia coli*!).

[7] (Benomar et al., 2015) and (Wrighton et al. 2014) are two good entry points to the rapidly developing research area on metabolic interdependency.

Table 1
Various degrees of autonomy exhibited by dsDNA viruses

virus family	Ribosomes	tRNA aminoacyl- ligase	DNA polymerase	RNA polymerase encoded	RNA polymerase (in virion)	Genome size
Megaviridae	−	±	+	+	+	1 Mb
Pandoraviridae	−	±	+	+	−	2.5 Mb
Pithoviridae	−	−	+	+	+	0.6 Mb
Mollivirus	−	−	+	+	−	0.6 Mb
Poxviridae	−	−	+	+	+	0.3 Mb
Herpesviridae	−	−	+	+	−	0.3 Mb
Chloroviruses	−	−	+	−	−	0.4 Mb
Adenoviridae	−	−	+	−	−	25 kb
Polyomaviridae	−	−	−	−	−	5 kb

organisms by simply counting the number of genes that need to be reintroduced into the genome of any parasite to reinstate its defective pathways until it could recover an independent lifestyle. Notice that such a "Gedanken experiment" could become a real one soon (Gibson et al., 2010). According to such a scale, cyanobacteria would be considered more alive than heterotrophic bacteria such as E. coli, themselves more alive than parasitic/endosymbiotic bacteria, the lowest point in the scale being reached (as of today) by the endosymbiotic bacterium *Tremblaya princeps* the genome of which only encode 121 proteins (López-Madrigal et al., 2011). Interestingly enough, the question of the limit at which these bacteria should still be considered "alive" is rarely raised, although their anabolic and energy producing capabilities are almost entirely absent and that key-components of their division apparatus, DNA replication apparatus, transcription apparatus, and translation apparatus are not fully encoded by their genomes.

In continuity with what we just did for parasitic cellular organisms, we can now analyze the gene contents of the various DNA viruses to establish their level of dependency with respect to their host cell, hence their degree of "absolute parasitism", again a property normally regarded as qualitative.[8] As shown in Table 1, DNA viruses naturally fall on a 4-degree scale, based on the presence/absence of encoded components of the translation apparatus, the DNA replication apparatus, and the transcription apparatus. Since we focus our discussion on the notion of ***absolute*** parasitism (i.e. a dependency on a cell machinery not mere biochemical compounds) we disregarded the diverse partial biosynthetic pathways (mostly for nucleotides and carbohydrates) eventually encountered in these viruses. Four levels of increasing absolute parasitism can be defined as follows:

- Viruses entirely replicating inside the host cytoplasm, such as the Megaviridae. These viruses encode their own transcription apparatus and, in addition, load their particle with the corresponding enzymes (RNA polymerase and transcription factors) allowing them to initiate their replication cycle without the help of the cell's nucleus. Interestingly, the most "autonomous" viruses (known as of today) also encode several components of a translation apparatus, albeit no ribosomes (Claverie & Abergel, 2010).
- Viruses encoding their own transcription apparatus, but not loading it into their particle, such as Pandoraviruses (Philippe et al., 2013). These viruses must initiate their replication cycle with an initial help from the cell nucleus, where a first round of transcription has to be performed by the cellular RNA polymerase. Thus, not packaging the viral RNA polymerase in the particle, far from being insignificant, has important consequences on the mode of replication.
- Viruses not encoding their own (or an incomplete) transcription apparatus. These viruses are condemned to use the cell transcription machinery and thus must go through a fully intranuclear stage. Paradoxically, many of these viruses (such as the ubiquitous Adenoviruses) still encode their own DNA replication apparatus, while the cellular one is readily available inside the nucleus where the viral genome has to reside to be transcribed.
- Viruses not even encoding their own DNA replication apparatus. As before, these viruses must replicate in the nucleus.

Interestingly, there is no known example of DNA virus encoding a RNA polymerase but no DNA polymerase. This suggests to us that there is a strict evolutionary hierarchy between the two apparatus. However, neither of them is required to make highly successful and ubiquitous viruses such as the Polyomaviruses or the Papillomaviruses.

Similar to the cellular microorganisms exhibiting various levels of autonomy with respect to their environments (from autotrophy to obligate intracellular parasitism), DNA viruses exhibit a gradation in their autonomy *vis-à-vis* the host cell within which they replicate (on the notion of autonomy and its application to viruses, see also Dupré & Guttinger, 2016, as well as Pradeu, 2016). In a Gedanken experiment, it should be possible to multiply a Megaviridae using a cell-free system (providing a functional translation apparatus, amino-acids and ATP) making it more "alive" and cell-like than simpler viruses. This definitely suggests that the notion of continuity between "life" forms, ranging from cellular autotrophs to the simplest viruses, initially rejected by Lwoff with its famous aphorism "viruses are viruses" (Lwoff, 1957, p. 240), might be worth revisiting on the combined light of the recently discovered ultra-parasitic bacteria and quasi-autonomous giant viruses. Breaking this potential epistemological barrier[9] is the only way to justify the search for missing links that are simply impossible to conceive within the current paradigm.

What could these missing links look like? For instance, some viruses could be found to encode a minimal ATP-producing pathways (such as glycolysis) to transiently boost the energy available in his host, hence enhancing their own fitness. Other even more cell-like viruses could possess a full complement of tRNA-aminoacyl ligases, or their own ribosomes (that would need to be both encoded and packaged). Although drastically violating two of the most central Lwoff's criteria, these entities could still be classified

[8] By reappraising the meaning of "absolute parasitism" in the light of contemporary microbiology and our new knowledge of microbial genomes we are now tackling an epistemological barrier of the "verbal" type, according to Bachelard (op. cit.; chapter IV). A concept until now considered obvious and self-explanatory (i.e. derived from a familiar image inherited from our everyday experience) suddenly requires a semantic inspection, taking into account previously unsuspected details or degrees brought about by scientific progresses. This phenomenon, akin to a "pixelization" of the initial image is highly recurrent in biological research.

[9] e.g. a premature "generalization" of Lwoff's belief, *sensu* Bachelard (op. cit. Chapter 3).

as "viruses" provided we use the new definition presented in a subsequent section. However, before to get to that point we would like to discuss yet another epistemological barrier that is preventing many virologists to correctly interpret the existence of giant viruses.

4.3. One more epistemological barrier[10]: virus do not undergo reductive evolution

The concept of "reductive genome evolution", i.e. the irreversible loss of genes as time goes on, is a characteristic of all parasitic life forms. It has been particularly well documented and studied for endosymbiotic/parasitic bacteria, as the one we discussed in the previous section. After a microorganism enters an endosymbiotic/parasitic association, functions once essential to its survival as a free living organism can be lost if compensated by the host. For instance, the incentive for a parasitic bacterium to synthetize a given metabolite disappears if it can be taken from its surroundings. The genes of the corresponding biosynthetic pathway can thus be lost at no, or minimal, fitness cost (or can even be advantageous if this synthesis was costly in energy). In addition, intracellular microorganisms tend to exist as smaller and more isolated populations than free living ones. These conditions accelerate the fixation of neutral or even slightly detrimental mutations through an irreversible evolutionary mechanism known as "Muller's ratchet" (Moran, 1996).

This slow, cumulative and irreversible loss of functions is the reason why no parasitic[11] (in particular intracellular) microorganism could ever evolve back toward a more independent lifestyle. The "once a parasite, always a parasite" rule is among the few biological ones not suffering exceptions (at least for intracellular parasites) (Poulin, 2007). In the case of symbiosis, gene losses may continue until the parasitic microorganisms become fully integrated in the host cell as an organelle (such as mitochondria), or vanishes altogether, sometimes following the transfer of its genes to the host's genome (Sloan et al., 2014). This evolutionary process (i.e. reductive evolution) is perfectly illustrated by the highly reduced genomes of the parasitic bacteria we described previously. However, it must be noted that for bacteria living inside a eukaryotic cell, the existence of various compartments delimited by membranes (such as the bacterial cytoplasm or the host cell's nucleus), as well as regulatory incompatibilities between the bacterial and eukaryotic subsystems, make some functions more challenging to lose than others. Biosynthetic pathways producing metabolites easy to ferry across membranes (by diffusion or active transporters) will be more readily lost than nuclear-based functions such as DNA replication or transcription. Protein translation, in particular the ribosomes, are also maintained in bacterial parasites despite their redundancy with the host cell's apparatus. This might be due to the significant differences known to exist between the mechanisms of translation initiation of eubacteria *versus* eukaryotes and to the difficulty (maybe an unresolved evolutionary challenge) to import eukaryotic ribosomes across a plasmic membrane. However, analyses of the genome of the most reduced intracellular parasitic bacteria pointed out the absence of numerous central components of their translation apparatus, including aminoacyl-tRNA ligases and ribosomal proteins (López-Madrigal et al., 2011; McCutcheon & Moran, 2011). This suggests that reductive evolution might eventually lead to a violation of a key Lwoff's criterion: "cellular" parasites devoid of protein translation capability.

Amazingly, the phenomenon of genome reduction, so well established for bacterial intracellular parasites, was rarely invoked in the evolutionary context of large DNA viruses although it seems to provide a simple explanation for the wide range of genome sizes and diversity of gene contents they exhibit (Claverie, 2006; Claverie et al., 2006).

To our knowledge, no rationale was ever proposed as to why viruses, the archetypes of obligate intracellular parasites, might be immune to the irreversible genome reduction process. On the contrary, the accepted paradigm was exactly the opposite: large DNA viruses were depicted as efficient "pick-pockets of cellular genes", gaining functions and genes over time, rather than losing them. This traditional way of thinking was boosted after the discovery of the first giant virus (Filée, Siguier, & Chandler, 2007; Iyer, Balaji, Koonin, & Aravind, 2006; Moreira & Brochier-Armanet, 2008; Yutin, Wolf, & Koonin, 2014). In its most recent and extreme version, all large and giant DNA viruses are deemed to derived from the same ancestral mobile element (Krupovic & Koonin, 2015).

Such a persistent refusal to consider viruses as possibly submitted to reductive evolution, as any other intracellular cellular parasites, is an epistemological barrier that we need to cross if we want to take a full advantage of the discovery of giant viruses. We will see that it may even lead us to challenge an even more solidly anchored belief: that of the unique origin of cellular life.

4.4. Viruses as "sort of microbes" lacking essential cellular functions

Within the logic of Lwoff's criteria, viruses are seen as "sort of microbes" missing properties thought to be essential cellular features by the biologists of his time. Such inability to define viruses otherwise than as missing essential cellular properties is eminently compatible with the theory that viruses were derived from the cellular world through the gradual loss of essential functions, forcing them increasingly deeper into "absolute parasitism". If we accept this scenario, the gene contents of various DNA viruses is then expected to be extremely variable, as the phenomenon of "lineage specific gene loss" is a trademark of the stochastic process of genome reduction (Blanc et al., 2007). This is what is actually observed, in particular when comparing the four known families of giant viruses (Abergel, Legendre, & Claverie, 2015).

The existence of a hierarchy of "absolute parasitism" among the DNA viruses depicted in the above section, is also easily explained in the framework of reductive evolution, whereby initially complex viruses trapped into an intracellular life-style will continuously lose genes and functions, although at variable speeds resulting from diverse ecological and physiological constraints. In this context, the slowest evolving lineages might correspond to the most complex viruses while the simplest (i.e. most reduced) viruses are the end products of the fastest evolving lineages.

The discovery of what looked like a vestigial translation apparatus in the Megaviridae (in the form of seven amino-acyl tRNA ligases) (Arslan et al., 2011) brought in an additional support to the reductive evolution scenario. While seven of these enzymes (connecting each amino-acid to its cognate codon) were clearly not enough to constitute a functional tool box for protein translation, it was however too many to be plausibly explained by random horizontal gene acquisitions. Since, by Lwoff's criteria, only cells can perform protein translation, the ancestor of the Megaviridae had to be some sort of cell.

[10] That one does not easily fit within Bachelard's classification of epistemological barriers. We believe it is merely due to virus (particles) not being considered "alive", and thus not obeying the laws of regular (endosymbiotic/parasitic) microorganisms. The dominant theory for the origin of viruses is exactly the opposite, whereby the first viruses had to be small and simple (see its last avatar in Koonin, Krupovic, & Yutin, 2015) and giant viruses originated from smaller ones (Yutin et al., 2014).

[11] We will no longer make the irrelevant distinction between parasites or endosymbionts here.

According to the dominant "gene accretion" paradigm (Filée et al., 2007; Iyer et al., 2006; Moreira & Brochier-Armanet, 2008; Krupovic & Koonin, 2015) we ought to replace, today's giant viruses are hypothesized to have evolved from minimal small genomes and grew in size and complexity by capturing genes from their cellular host, or their environment. While the number of genes estimated to originate from the ancestral virus is not more than 50 (i.e. the "core" genes with homologs in different virus families (Iyer et al., 2006)), this means that more than 90% of the genome of giant viruses should have been acquired from their cellular host (an amoeba), or other cellular organisms. However, only a very small percentage (less than 10% for Pandoravirus) of the genes in giant viruses appear to share a common ancestry with genes from any other known virus or cellular organisms (prokaryote or eukaryotes) and even less so with their amoebal host (Abergel et al., 2015). Thus, if the giant genome of giant viruses was mostly acquired from cellular organisms, these organisms are nowhere to be seen today. Proponents of the gene accretion scenario have tried to explain the absence of recognizable similarity by the supposed fast divergence rate of these viruses. However, 1) there is no evidence that DNA (in contrasts with RNA) viruses evolve much faster than cells (Doutre, Philippe, Abergel, & Claverie, 2014), and 2) this fast divergence should have made impossible the detection of the "core" viral genes supposedly dating from the oldest common ancestor of DNA viruses (Iyer et al., 2006). But the overwhelming argument against the gene accretion scenario is again the huge proportion of giant virus genes (from 2/3 to more than 90%) coding for proteins without homologs in the three cellular domains: eubacteria, archaea, and eukarya (Abergel et al., 2015). If these genes were acquired, where are they coming from? The rationale behind the denial of genome reduction as an alternative scenario for the evolution of giant DNA viruses is not clear. It may be an unconscious legacy from Lwoff's credo that no intermediate could exist between viruses and cells, even including extinct ancestral lineages.

4.5. The unique origin of cellular life: one more epistemological barrier to leapfrog?

Reductive evolution is indeed perfectly consistent with the diversity in size and genome complexity of the many families of eukaryotic DNA viruses such as those listed in Table 1. These diverse families could result from alternative reductive evolutionary pathways, initiated by the loss of functional ribosomes, the most basic system the absence of which is shared by all viruses. Then committed to an absolute parasitic life-style, their evolution was punctuated by the further loss of fundamental functions in some lineages, such as transcription (i.e. a virally-encoded RNA polymerase) or DNA replication (i.e. virally-encoded nucleotide-handling and DNA repair enzymes, and DNA polymerase). These successive losses first led to increasingly host-dependent cytoplasmic viruses, then to viruses replicating within the host nucleus. Further reductions led to the simplest DNA viruses we know today for which most of the replicative functions are performed by the host (Table 1). We know now that such a reduction process can eventually culminate with most of the functions required for the virus replication becoming encoded in the host genome (Herniou et al., 2013).

However, as attractive as it is, this simple scenario is not compatible with the very high percentage of giant virus genes without traceable ancestry to one of today's cellular domains (prokaryote or eukaryote). It is also not compatible with the fact that the four known families of giant viruses are as different from each other as they are different from extent cellular organisms. To rescue our reductive evolution scenario, we have to hypothesize that these various virus lineages have different origins, possibly in multiple ancestral protocell-types that were once in competition with the one that gave rise to LUCA, the last common universal ancestor to the Eubacteria, Archaea and Eukarya. At the end of a fierce evolutionary battle, the "loser" proto-cells would have partially survived as parasites (or endosymbionts) of the "winner" cellular lineages, giving rise to the diversity of DNA viruses we know today. DNA viruses would thus be the descendant of these vanished cell-types following a billion year of co-evolution in a variety of extant cellular organisms derived from LUCA (Abergel et al., 2015). Paradoxically, virus-bearing ancestral cells might also have enjoyed a selective advantage through the accelerated evolution of their genomes promoted by increased virus-induced genes exchanges and nucleic acid shuffling. The suggestion that (giant) DNA viruses might predate LUCA has been made, in various forms, by different authors (Forterre, 1992; Nasir et al., 2012; Abrescia, Bamford, Grimes, & Stuart, 2012 and references herein). Most recently, a detailed phylogenomic analysis of all virus types (including ARN and DNA viruses) concluded to an evolutionary scenario very similar to ours (Nasir & Caetano-Anollés, 2015). As speculative as these ideas might seem, the recent discovery of 2.1 billion-year old fossils possibly corresponding to an aborted pre-metazoan lineage, reminded us of the fragility of our current knowledge about the early days of life on our planet (El Albani et al., 2010; El Albani et al., 2014). Our inborn fascination for unique causes to which attributing the emergence of complex phenomena, such as life, might constitute the most serious epistemological barrier that remains to be broken to understand the true significance of giant viruses.

5. The concept of virus in the post giant virus era

The discovery of giant viruses was delayed for up to 10 years, and probably much more, due to the blind confidence that microbiologists put into an initial size-based paradigm that was too hastily built and cast in stone. As a result, three of the four giant viruses known today were spotted years before their viral nature was recognized. This paradigm was based on a number of known viruses too small to encompass their diversity in size, shape, structure, biological complexity and habitat. However, it was also extremely fruitful in leading to the rapid discovery of the great majority of "regular viruses", most of them pathogenic to animals or plants. Initially useful generalizations, later on turning out to be scientifically detrimental are among the epistemological barriers recognized by Bachelard. Amazingly, the validity of Lwoff's carefully drafted criteria suffered little damage in the process. His discrimination protocol suffered more from the discovery of unanticipated highly defective cellular microorganisms than from that of unexpectedly complex viruses.

5.1. The status of Lwoff criteria today

Taking into account what we know today about viruses and cellular organisms (including endosymbionts and intracellular parasites), the situation is as follows:

1) Typical cellular organisms contain both DNA and RNA, viruses only contain one type;

This statement is ambiguous as it mixes the concept of cell, virus and virion. As a replicative process inside a cellular host, there is obviously no difference between virus and cell (see Forterre, 2013, 2016 for the "virocell concept"). This criterion is thus useless. If we understand the word "virus" as meaning "the particle", the criteria is also dismissed as mRNA have been shown to be packaged in the particle of giant viruses (Raoult et al., 2004). However, taking into

account the historical context, this sentence should be understood as "there is no ribosome (i.e. ribosomal RNAs) in viral particles". We believe the modern criterion truly translating Lwoff's thoughts would be: viruses do not encode ribosomes, cells do. Although still formally valid – no virus has yet been found encoding a fully functional translation apparatus, this criterion is being weakened by the discovery of parasitic bacteria that do not encode either a full complement of ribosomal proteins or aminoacyl-tRNA ligases, pending the eventual discovery that ultimately reduced bacteria may entirely rely on the ribosomes of the cellular host they infect.

2) All microorganisms are reproduced from the integrated sum of their constituents; viruses are produced from their nucleic-acid only;

This is now obviously wrong, as the DNA genome of the large or giant viruses infecting eukaryotes is not infectious in absence of a large array of well-organized proteins. This criterion is also ambiguous as it compares the "cell" to the "virion", and not to the "virus", defined as the whole replication process that takes place within the host cell.

3) During the growth of a microorganism, the individuality of the whole is maintained, and culminates in binary fission. There is no binary fission in viruses;

Again, this statement is ambiguous. If the word "virus" is taken as meaning "the particle", this is obviously true, but the comparison is meaningless: a dormant seed or a bacterial spore do not divide either, but initiate a whole developmental program once put into suitable conditions, just as a viral particle generates a complex intracellular "virion factory" after penetrating the host cytoplasm. If we take the word "virus" as meaning the "intracellular replicative process", the criterion remains true, as "virion factories" do not multiply by division, but by propagating from cell to cell using infectious virions. However, certain intracellular (or intra-nuclear) bacteria either do not appear to encode a division apparatus, or grow into unseptated filaments that eventually separate in multiple cells. Thus if the multiplication by some sort of (not always binary) division remains the privilege of cellular organisms, it does not correspond to a conserved set of telltale genes. This implies that genomic information alone is insufficient to establish the existence of a division process without the direct observation (hence cultivation) of the most reduced cellular microorganisms.

4) Viruses lack the system of enzymes which convert the potential energy of foodstuffs into the energy necessary to biochemical syntheses (i.e. an ATP producing machinery).

This criterion is not valid anymore, as many intracellular parasitic bacteria are unable to produce ATP and use the one available in the surrounding host cytoplasm. On the opposite, some viruses (e.g. cyanophages) have been found to encode parts of a photosynthetic apparatus boosting the energy balance of their host upon infection (Sharon et al., 2009). Furthermore, without violating any basic biological rule, a large virus genome might be one day found to encode the few enzymes required to produce ATP by glycolysis.

5) Viruses make use of the ribosomes of their host cells;

As discussed above, this statement is still valid, but some extremely reduced parasitic bacteria may also do likewise, making this criteria less discriminant than originally thought.

5.2. On the formal impossibility to discriminate viruses from cells on the basis of their gene contents

The present status of Lwoff's criteria points out protein translation and binary division as the sole remaining key biological functions allowing a formal discrimination of most "regular" cells from most "regular" viruses. Yet, the continuous discovery of increasingly host-dependent intracellular parasitic bacteria concurrently to that of increasingly autonomous viruses is making their discrimination more and more challenging, in particular from the sole analysis of their respective gene contents (as often done in the context of metagenomics).

As we already mentioned, the translation apparatus encoded by some intracellular bacteria is far from being complete, with up to 11 amino-acyl tRNA ligase missing, as well as 15 ribosomal proteins (McCutcheon & Moran, 2011). The mechanisms by which these defects are compensated are unknown. In the meantime, *Megavirus chilensis* encodes 7 aminoacyl tRNA ligases, as well as many other translation factors, but no ribosomal proteins (Arslan et al., 2011).

On the other hand, many parasitic bacteria have been found to lack any recognizable genes related to known mechanisms of cell division and/or membrane synthesis pathways. A direct observation of the microorganism caught in the process of dividing becomes thus required to establish its cellular nature.

Thus, without the need to further wait for the future discovery of even more atypical microbes, we can already infer that it is impossible to propose a robust definition of virus or cells based on their sole gene contents (such as the proposed dichotomy between capsid-encoding and ribosome-encoding microorganisms (Raoult & Forterre, 2008; Forterre, 2016). We must renounce to a classification scheme by which viruses are defined by their lack of a common subset of cellular functions, because that subset steadily tends to zero. Virology is also longing for a definition of viruses more positive than a list of the biological functions they do not possess. On the other hand, we must also depart from a classification of "cellular" microbes based on a common subset of functions they might all possess, as this subset also clearly tends toward zero as more endosymbiotic/parasitic bacteria are discovered. It appears thus impossible to propose a lasting definition of viruses and cells that will forever encompass the tremendous diversity of their gene contents.

5.3. Our new definition of viruses

If we conclude that viruses and cells cannot be formally and rigorously discriminated from each other on the basis of their respective gene contents, is it possible to propose an alternative classification scheme that will be consistent with all our present knowledge and may remain valid as long as Lwoff's criteria? We believe the answer is yes. Irrespective of their metabolic capacities, genome types, particle structures, sizes, or morphologies, the key distinctive feature by which viruses and cells can still be recognized unambiguously is by the way they ***propagate their genome***.

Thus, in the most general sense, we propose to define as a "virus" any biological entity the genome (nucleic acid molecule) of which is:

1) Replicated by a system of macromolecules that it does not entirely encode (absolute parasitism)
2) Disseminated using a metabolically inert structure the maintenance of which does not require energy.

At variance with previous conceptions, such a definition encompasses genomes encapsulated within any type of particles (without constraints of size, morphology, or biochemical composition), as well as "naked" infectious nucleic acid molecules such as plasmids. We are looking forward with great anticipation to the debate this new proposed definition will no doubt trigger among our colleagues.

5.4. Closing remarks

Despite their relative abundance in the environment, and the ease with which they could be multiplied and visualized, giant viruses failed to be discovered and recognized as such for many years. Rapidly following the initial discovery of TMV by Dimitri Ivanovski, an unwarranted generalization of what viruses should look like put the mere concept of giant viruses outside of the sight of microbiologists. Entities that are not conceivable within a given paradigm are simply not accessible *via* a rational experimental approach and can only be stumbled across by scientists uninhibited by a previous knowledge of the field. The serendipitous discovery of the first giant viruses allowed original ways to reflect on the mere concept of viruses and speculate on their evolution and the origin of cellular life. New debates have now been initiated that will lead virologists to explore avenues that they did not know existed ten years ago. This exciting episode of totally unplanned basic research should serve as a precious reminder that, even in experimental biology, "discovery is to see what everybody else has seen, and to think what nobody else has thought" (Szent-Gyorgyi, 1957).

Acknowledgments

The initial version of this article was written during a summer visit of the St. Giles laboratory of Human Genetics of Infectious diseases, Rockefeller University, New York, at the invitation of Prof. Jean-Laurent Casanova. We wish to thank him and his collaborators for providing us with ideal working conditions. We also thank Thomas Pradeu, Gladys Kostyrka, and John Dupré for their useful comments on earlier versions of this article. The IGS laboratory is partially supported by the Centre National de la Recherche Scientifique and Aix-Marseille University.

References

Abergel, C., Legendre, M., & Claverie, J.-M. (2015). The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiology Reviews, 39*(6), 779–796.

Abrescia, N. G., Bamford, D. H., Grimes, J. M., & Stuart, D. I. (2012). Structure unifies the viral universe. *Annual Reviews Biochemistry, 81*, 795–822.

Antwerpen, M. H., Georgi, E., Zoeller, L., Woelfel, R., Stoecker, K., & Scheid, P. (2015). Whole-genome sequencing of a pandoravirus isolated from keratitis-inducing acanthamoeba. *Genome Announcement, 3*(2). http://dx.doi.org/10.1128/genomeA.00136-15.

Arslan, D., Legendre, M., Seltzer, V., Abergel, C., & Claverie, J.-M. (2011). Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proceedings National Academy Sciences U.S.A., 108*(42), 17486–17491.

Bachelard, G. (1993). *[1934]. La formation de l'esprit scientifique*. Paris: Librairie Philosophique J. Vrin (pocket edition).

Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., et al. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature, 310*(5974), 207–211.

Beijerinck, M. W. (1898). Concerning a contagium vivum fluidum as cause of the spot disease of tobacco leaves. Akademie van Wettenschappen te Amsterdam, 65(2), 3-21. (English translation in: Johnson, J. (ed) Phytopathological classics no. 7. American Phytopathological Society, Saint Paul, 1942, pp 33–52).

Benomar, S., Ranava, D., Cárdenas, M. L., Trably, E., Rafrafi, Y., Ducret, A., et al. (2015). Nutritional stress induces exchange of cell material and energetic coupling between bacterial species. *Nature Communications, 6*, 6283.

Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., Vestris, G., et al. (2007). Reductive genome evolution from the mother of Rickettsia. *PLoS Genetics, 3*(1), e14.

Boycott, A. E. (1928). The transition from live to dead: The nature of filterable viruses. *Proceedings Royal Society Medicine, 22*(1), 55–69.

Chamberland, C. (1884). A filter permitting to obtain physiologically pure water. *Comptes Rendus Académie des Sciences, 99*, 247–248.

Claverie, J.-M. (2006). Viruses take center stage in cellular evolution. *Genome Biology, 7*(6), 110.

Claverie, J.-M., & Abergel, C. (2009). Mimivirus and its virophage. *Annual Review Genetics, 43*, 49–66.

Claverie, J.-M., & Abergel, C. (2010). Mimivirus: The emerging paradox of quasi-autonomous viruses. *Trends Genetics, 26*(10), 431–437.

Claverie, J.-M., & Abergel, C. (2015). From extraordinary endocytobionts to pandoraviruses. Comment on scheid et al.: Some secrets are revealed: Parasitic keratitis amoebae as vectors of the scarcely described pandoraviruses to humans. *Parasitology Research, 114*(4), 1625–1627.

Claverie, J.-M., Ogata, H., Audic, S., Abergel, C., Suhre, K., & Fournier, P. E. (2006). Mimivirus and the emerging concept of "giant" virus. *Virus Research, 117*(1), 133–144.

Doutre, G., Philippe, N., Abergel, C., & Claverie, J.-M. (2014). Genome analysis of the first Marseilleviridae representative from Australia indicates that most of its genes contribute to virus fitness. *Journal of Virology, 88*(24), 14340–14349.

Dupré, J. A., & Guttinger, S. (2016). Viruses as living processes. *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 109–116.

Dupré, J., & O'Malley, M. A. (2009). Varieties of living things: Life at the intersection of lineage and metabolism. *Philosophy & Theory in Biology, 1*. http://dx.doi.org/10.3998/ptb.6959004.0001.003. Michigan Publishing.

El Albani, A., Bengtson, S., Canfield, D. E., Bekker, A., Macchiarelli, R., Mazurier, A., et al. (2010). Large colonial organisms with coordinated growth in oxygenated environments 2.1 Gyr ago. *Nature, 466*(7302), 100–104.

El Albani, A., Bengtson, S., Canfield, D. E., Riboulleau, A., Rollion Bard, C., Macchiarelli, R., et al. (2014). The 2.1 Ga old Francevillian biota: Biogenicity, taphonomy and biodiversity. *PLoS One, 9*, e99438.

Filée, J., Siguier, P., & Chandler, M. (2007). I am what I eat and I eat what I am: Acquisition of bacterial genes by giant viruses. *Trends Genetics, 23*(1), 10–15.

Fischer, M. G., Allen, M. J., Wilson, W. H., & Suttle, C. A. (2010). Giant virus with a remarkable complement of genes infects marine zooplankton. *Proceedings National Academy Sciences U.S.A, 107*(45), 19508–19513.

Fischer, M. G., & Suttle, C. A. (2011). A virophage at the origin of large DNA transposons. *Science, 332*, 231–234.

Forterre, P. (1992). New hypotheses about the origins of viruses, prokaryotes and eukaryotes. In J. K. Trân Thanh Vân, J. C. Mounolou, J. Shneider, & C. Mc Kay (Eds.), *Frontiers of life* (pp. 221–234). Gif-sur-Yvette, France: Editions Frontières.

Forterre, P. (2013). The virocell concept and environmental microbiology. *International Society Microbial Ecology Journal, 7*(2), 233–236.

Forterre, P. (2016). To be or not to be alive: how recent discoveries challenge the traditional definitions of viruses and life. *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 100–108.

Fry, N. K., Rowbotham, T. J., Saunders, N. A., & Embley, T. M. (1991). Direct amplification and sequencing of the 16S ribosomal DNA of an intracellular Legionella species recovered by amoebal enrichment from the sputum of a patient with pneumonia. *FEMS Microbiology Letters, 67*(2), 165–168.

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., et al. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science, 329*(5987), 52–56.

Helvoort, T. V. (1996). When did virology start? *American Society Microbiology News, 62*, 142–145.

Herniou, E. A., Huguet, E., Thézé, J., Bézier, A., Periquet, G., & Drezen, J.-M. (2013). When parasitic wasps hijacked viruses: Genomic and functional evolution of polydnaviruses. *Philosophical Transactions Royal Society London B Biological Sciences, 368*(1626), 20130051.

Hoffmann, R., Michel, R., Muller, K.-D., & Schmid, E. N. (1998). Archaealike endocytobiotic organisms isolated from Acanthamoeba sp. (Gr II). *Endocytobiosis Cell Research, 12*, 185–188.

Ivanovski, D. (1892). *Über die Mosaikkrankheit der Tabakspflanze*. Bulletin Scientifique publié par l'Académie Impériale des Sciences de Saint-Pétersbourg/Nouvelle Serie III (in German & Russian) (St. Petersburg) 35, 67-70. Translated into English in Johnson, J., Ed. (1942) Phytopathological classics (St. Paul, Minnesota: American Phytopathological Society) No. 7, pp. 27–-30.

Iyer, L. M., Balaji, S., Koonin, E. V., & Aravind, L. (2006). Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Research, 117*(1), 156–184.

Kausche, G. A., Pfankuch, E., & Ruska, H. (1939). Die Sichtbarmachung von pflanzlichem virus im Ubermikroskop. *Naturwissenschaften, 27*, 292–299.

Koonin, E. V., Krupovic, M., & Yutin, N. (2015). Evolution of double-stranded DNA viruses of eukaryotes: From bacteriophages to transposons to giant viruses. *Annals New York Academy Science, 1341*, 10–24.

Koonin, E. V., & Starokadomskyy, P. (2016). Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question,. *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 125–134.

Krupovic, M., & Koonin, E. V. (2015). Polintons: A hotbed of eukaryotic virus, transposons and plasmid evolution. *Nature Reviews Microbiology, 13*, 105–115.

Kostyrka, G. (2016). What roles for viruses in origin of life scenarios? *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 135–144.

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., et al. (2003). A giant virus in amoebae. *Science, 299*, 2033.

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature, 455*(7209), 100–104.
Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proceedings National Academy Sciences U.S.A, 111*(11), 4274–4279.
Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., et al. (2015). In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proceedings National Academy Sciences U.S.A, 112*(38), E5327–E5335.
López-Madrigal, S., Latorre, A., Porcar, M., Moya, A., & Gil, R. (2011). Complete genome sequence of "Candidatus Tremblaya princeps" strain PCVAL, an intriguing translational machine below the living-cell status. *Journal Bacteriology, 193*(19), 5587–5588.
Lwoff, A. (1957). The concept of virus. *Journal General Microbiology, 17*, 239–253.
Lwoff, A., & Tournier, P. (1966). The classification of viruses. *Annual Reviews Microbiology, 20*, 45–74.
McCutcheon, J. P., & Moran, N. A. (2011). Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology, 10*(1), 13–26.
Méthot, P.-O. (2016). Writing the history of virology in the twentieth century: discovery, disciplines, and conceptual change. *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 145–153.
Monier, A., Claverie, J.-M., & Ogata, H. (2008). Taxonomic distribution of large DNA viruses in the sea. *Genome Biology, 9*, R106.
Moran, N. A. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proceedings National Academy Sciences U.S.A, 93*(7), 2873–2878.
Moreira, D., & Brochier-Armanet, C. (2008). Giant viruses, giant chimeras: The multiple evolutionary histories of mimivirus genes. *BMC Evolutionary Biology, 8*, 12.
Moreira, D., & López-García, P. (2009). Ten reasons to exclude viruses from the tree of life. *Nature Reviews Microbiology, 7*(4), 306–311 (and following responses)
Mutsafi, Y., Zauberman, N., Sabanay, I., & Minsky, A. (2010). Vaccinia-like cytoplasmic replication of the giant mimivirus. *Proceedings National Academy Sciences U.S.A, 107*, 5978–5982.
Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., et al. (2006). The 160-kilobase genome of the bacterial endosymbiont Carsonella. *Science, 314*(5797), 267.
Nasir, A., & Caetano-Anollés, G. (2015). A phylogenomic data-driven exploration of viral origins and evolution. *Science Advances, 1*, e1500527.
Nasir, A., Kim, K. M., & Caetano-Anolles, G. (2012). Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evolutionary Biology, 12*, 156.
Nocard, E., & Roux, E. (1898). Le microbe de la péripneumonie. *Annales Institut Pasteur, 12*, 240–262.
Pasteur, L. (1878a). La théorie des germes et ses applications à la chirurgie. *Bulletin Académie nationale médecine, 2*(7), 166–167.
Pasteur, L. (1878b). La théorie des germes et ses applications à la chirurgie. *Bulletin Académie nationale médecine, 2*(7), 283–284.
Pasteur, L., Joubert, J., & Chamberland, C. (1878). La théorie des germes et ses applications à la médecine et à la chirurgie. *Comptes rendus Académie des sciences, LXXXVI*, 1037–1043.
Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., et al. (2013). Pandoraviruses: Amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science, 341*, 281–286.
Poulin, R. (2007). *Evolutionary ecology of parasites*. New Jersey, USA: Princeton University Press.
Pradeu, T. (2016). Mutualistic viruses and the heteronomy of life. *Studies in History and Philosophy of Biological and Biomedical Sciences, 59*, 80–88.
Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004). The 1.2-Mb genome sequence of mimivirus. *Science, 306*, 1344–1350.
Raoult, D., & Forterre, P. (2008). Redefining viruses: Lessons from mimivirus. *Nature Reviews Microbiology, 6*(4), 315–319.
Raoult, D., La Scola, B., & Birtles, R. (2007). The discovery and characterization of Mimivirus, the largest known virus and putative pneumonia agent. *Clinical Infectious Diseases, 45*(1), 95–102.
Roux, E. (1903). Sur les microbes dits invisibles. *Bulletin Institut Pasteur Paris, 1*, 7–12, 49-56.
Santini, S., Jeudy, S., Bartoli, J., Poirot, O., Lescot, M., Abergel, C., et al. (2013). Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. *Proceedings National Academy Sciences U.S.A., 110*(26), 10800–10805.
Scheid, P., Hauröder, B., & Michel, R. (2010). Investigations of an extraordinary endocytobiont in acanthamoeba sp.: Development and replication. *Parasitology Research, 106*, 1371–1377.
Scheid, P., Zöller, L., Pressmar, S., Richard, G., & Michel, R. (2008). An extraordinary endocytobiont in Acanthamoeba sp. isolated from a patient with keratitis. *Parasitology Research, 102*, 945–950.
Schulz, F., Lagkouvardos, I., Wascher, F., Aistleitner, K., Kostanjšek, R., & Horn, M. (2014). Life in an unusual intracellular niche: A bacterial symbiont infecting the nucleus of amoebae. *International Society Microbial Ecology Journal, 8*(8), 1634–1644.
Sharon, I., Alperovitch, A., Rohwer, F., Haynes, M., Glaser, F., Atamna-Ismaeel, N., et al. (2009). Photosystem I gene cassettes are present in marine virus genomes. *Nature, 461*(7261), 258–262.
Sloan, D. B., Nakabachi, A., Richards, S., Qu, J., Murali, S. C., Gibbs, R. A., et al. (2014). Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Molecular Biology Evolution, 31*(4), 857–871.
Szent-Gyorgyi, A. (1957). *Bioenergetics*. p. 57. New York: Academic Press.
Villarreal, L. P., & Witzany, G. (2010). Viruses are essential agents within the roots and stem of the tree of life. *Journal Theoretical Biology, 262*(4), 698–710.
Weitz, J. S., Stock, C. A., Wilhelm, S. W., Bourouiba, L., Coleman, M. L., Buchan, A., et al. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *International Society Microbial Ecology Journal, 9*(6), 1352–1364.
Wrighton, K. C., Castelle, C. J., Wilkins, M. J., Hug, L. A., Sharon, I., Thomas, B. C., et al. (2014). Metabolic interdependencies between phylogenetically novel fermenters andrespiratory organisms in an unconfined aquifer. *International Society Microbial Ecology Journal, 8*(7), 1452–1463.
Yutin, N., Wolf, Y. I., & Koonin, E. V. (2014). Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology, 466–467*, 38–52.
Zielinski, F. U., Pernthaler, A., Duperron, S., Raggi, L., Giere, O., Borowski, C., et al. (2009). Widespread occurrence of an intranuclear bacterial parasite in vent and seep bathymodiolin mussels. *Environmental Microbiology, 11*(5), 1150–1167.