JOURNAL OF VIROLOGY, Dec. 2010, p. 12476–12479 Vol. 84, No. 24
0022-538X/10/$12.00 doi:10.1128/JVI.01489-10

COMMENTARY

Order to the Viral Universe[▽]

Mart Krupovič† and Dennis H. Bamford*

Department of Biosciences and Institute of Biotechnology, Viikki Biocenter 2, P.O. Box 56 (Viikinkaari 5), University of Helsinki, FIN-00014 Helsinki, Finland

It is perhaps human nature to seek order in everything that surrounds us. With the number of individual viral particles in the biosphere considerably exceeding that of their hosts (33), the virosphere—or the viral universe, if you wish—shares this quest. In fact, the number of viruses is so high that the entire tree of cellular life, from its roots to the tips of its branches, seems to be immersed in the sea of viruses (5). The attempt to bring order to the virosphere is manifested in the work carried out by the International Committee on Taxonomy of Viruses (ICTV) (9), an official body consisting of numerous experts in the field of virology. The ICTV currently recognizes the hierarchical levels of order, family, subfamily, genus, and species. The dominant (but not exclusive) demarcation criterion employed to delineate viruses into different taxonomic levels is sequence comparison, the main advantage of which is the quantitative reflection of the divergence between taxa. Virus origin and evolution are issues that have never ceased to fascinate biologists and are currently highly debated (see recent studies, e.g., references 11–13, 16, 18–19, 21, 23, 27, and 34). Unfortunately, sequence comparisons do not reach far enough to recognize the relationships between viruses that diversified further back in time while leaving no detectable signal at the sequence level. However, new information, coming mainly from structural studies, might help us (i) to reveal deeper evolutionary connections between viruses that were not previously considered to be related and (ii) to establish biologically relevant higher levels of virus classification, thereby bringing additional biology to taxonomy.

Extensive horizontal gene transfer has had an immense impact on the evolution of viruses, with the result that their genomes are mosaics of genes with distinct evolutionary histories (10, 14). Such genetic mosaicism complicates the reconstruction of the evolutionary relationships for distantly related viruses to the point that it becomes impossible to infer a single phylogeny that would represent the evolution of all or even most of the genes in these viral genomes. The two most obvious solutions to this problem are (i) to represent the evolution of viruses using network-like diagrams that capture the dynamics of the reticulate relationships between viral genomes (26) and (ii) to select a set of viral genes that reflect properties identifying a biological object as a virus and to consider all horizontal gene transfer events to be relative to this selected gene set, which is inherited vertically (5). The disadvantage of the first approach is that it relies on sequence comparisons and therefore fails to recognize relationships between viruses that diverged further back in time. Furthermore, such representation masks the signal of vertical inheritance in viral genomes. In this commentary, we advocate the second of the two approaches and argue that deep evolutionary connections between viruses can be traced by a careful analysis of the architectural principles of the virions. We provide two examples of potential higher-level viral taxons. What the nomenclature for such a taxon would be is a matter of debate, but we have used the term (viral) lineage.

Inclusion in a lineage would be dependent less on sequence similarity or common gene content and more on the structural characteristics of the virion. It is therefore important to discuss why virion structure is considered a more relevant criterion for higher-level virus classification than, e.g., genome replication or transcription apparatuses. (i) The capacity to build a viral capsid is the sole characteristic which distinguishes viruses from other genetic elements, such as plasmids. A good illustration of this point is the comparison between porcine circovirus 1 (PCV1; GenBank accession number AY660574) and the *Helicobacter pylori* plasmid pAL236-5 (GenBank accession number HM125989), which both possess a minimal number of genes required for their propagation (two and one, respectively) (Fig. 1). Both genetic elements encode rolling-circle replication initiation proteins, but only PCV1 encodes a capsid protein and forms a virion. Even skeptics will perhaps admit that it is the capsid that makes PCV1 a virus (Fig. 1). The same logic is certainly also applicable to more complex viruses, a prerequisite condition for a more universal higher-level classification scheme. (ii) There is a vast amount of evidence indicating that viruses can exchange genes for genome replication proteins horizontally, not only with other mobile elements but also with cellular organisms (14, 22, 25, 35). For example, tailed double-stranded DNA (dsDNA) viruses of bacteria utilize a plethora of nonhomologous proteins for their genome replication (Fig. 2) (35), while archaeal viruses on multiple occasions captured the genes for replicative minichromosome maintenance (MCM) helicases from their hosts, as reported recently (25). Tracing the evolution and descent of these viral functions is nevertheless of great value and interest, since it sheds light on the evolutionary dynamics in the virosphere and

* Corresponding author. Mailing address: Department of Biosciences and Institute of Biotechnology, Viikki Biocenter 2, P.O. Box 56 (Viikinkaari 5), University of Helsinki, FIN-00014 Helsinki, Finland. Phone: 358 919159100. Fax: 358 919159098. E-mail: dennis.bamford@helsinki.fi.

† Present address: Institut Pasteur, Unité de Biologie Moléculaire du Gène chez les Extrêmophiles, Department of Microbiology, 28 rue du Dr Roux, 75015 Paris, France.

▽ Published ahead of print on 6 October 2010.

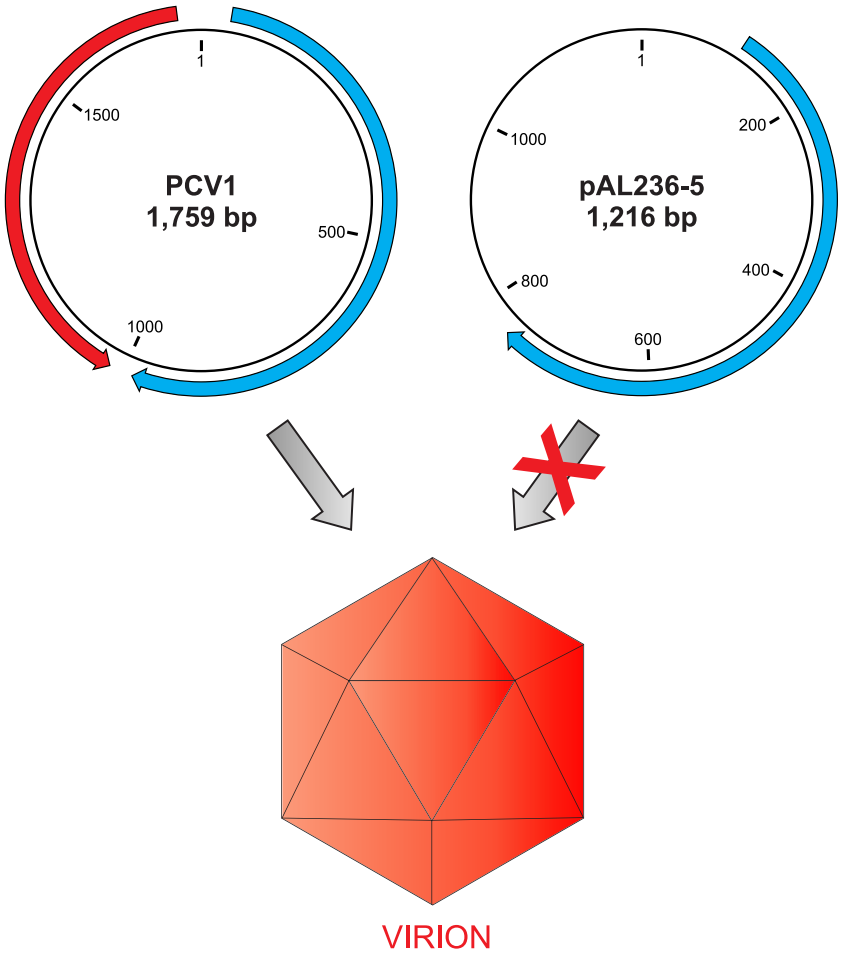

FIG. 1. What makes a virus a virus? Genome comparison of two mobile genetic elements, porcine circovirus 1 (PCV1; GenBank accession number AY660574) on the left, and *Helicobacter pylori* plasmid pAL236-5 (GenBank accession number HM125989) on the right. Genes for the rolling-circle replication initiation proteins in both mobile elements are shown by blue arrows, while the capsid protein-coding gene of PCV1 is depicted by a red arrow. Note that even though 50% of the PCV1 genes (1 out of 2) have counterparts in pAL236-5 plasmid, only the former can be considered a genuine virus, capable of forming a virion.

allows us to comprehend the magnitude of genetic exchange between viruses and other genetic elements, be they cellular chromosomes or plasmids (19). However, if we prefer a more biologically relevant higher-level virus classification, we should probably consider putting more weight on the viral functions responsible for virion construction than on those involved in genetic processes and virus-host interactions (21).

One of the examples of viral lineages is the one uniting large dsDNA viruses that utilize double β-barrel major capsid proteins (MCPs) to build their virions (Fig. 3). The architectural principles common to all these viruses have been recently reviewed elsewhere (23). This lineage includes viruses infecting hosts from all three domains of cellular life. The monophyletic origin for the majority of the eukaryote-infecting members of the lineage has been firmly established using sequence-based approaches (7, 15, 20). This assemblage of eukaryotic viruses is referred to as *n*ucleo*c*ytoplasmic *l*arge *D*NA *v*iruses (NCLDVs) (15). A number of structural studies of individual NCLDVs have complemented the sequence-based analyses, further supporting the common origin for these viruses (7, 28, 37–38). Structural and biochemical analyses revealed that other viruses, infecting not only eukaryotes but also bacteria and archaea, utilize architectural principles (from the same structural fold of the MCPs to common capsid assembly and genome packaging) very similar to those for NCLDVs (1, 3, 17, 31–32). The extent of similarity in virion organization of these viruses (Fig. 3) leaves no room for the possibility of convergence (23), suggesting that the common features were inherited from a common viral ancestor. This lineage currently is composed of nine officially recognized virus families and three additional viruses that have not been yet assigned to a family (Fig. 3).

Tailed dsDNA viruses of bacteria and archaea and eukaryotic herpesviruses represent yet another example of a viral lineage that unifies viruses from several different families, infecting hosts from all three domains of life. The molecular principles of virion assembly and maturation, genome packaging, and virion structure are very similar for these viruses (in some cases not only at the structure level but also at the sequence level), suggesting a common ancestry (4, 6, 8, 24, 30, 36). A combination of sequence-based (for more closely related viruses) and structural (for more distantly related ones)

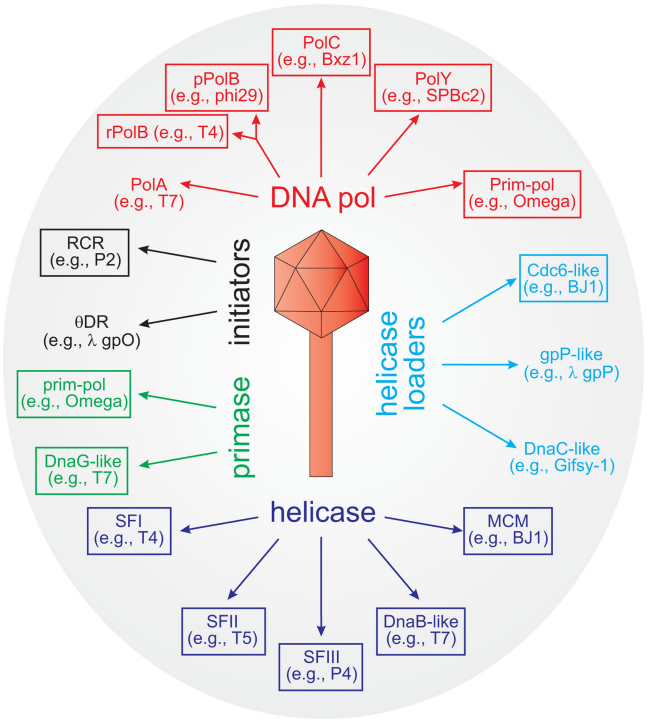

FIG. 2. Diversity of genome replication proteins utilized by tailed dsDNA viruses infecting bacteria and archaea. Replication proteins that are also encoded by nontailed viruses and/or plasmids are boxed. Examples of viruses utilizing the indicated genome replication proteins are depicted. Abbreviations: PolA, -B, -C, and -Y, types A, B, C, and Y DNA polymerases, respectively; rPolB, RNA/DNA-primed type B DNA polymerase; pPolB, protein-primed type B DNA polymerase; prim-pol, primase-polymerase protein; MCM, minichromosome maintenance helicase; SFI-III, superfamily I to III helicases; θDR, theta-type DNA replication; RCR, rolling-circle DNA replication.

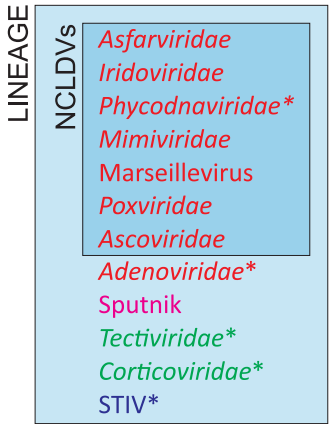

FIG. 3. An example of a higher-level taxonomic grouping, a lineage, for dsDNA viruses encoding the double β-barrel capsid proteins. Viral family names are written in italics; if a virus has not been assigned to a family, its name is provided. Virus (family) names are color coded according to the host the corresponding viruses infect: red, eukaryotic viruses; green, bacterial viruses; blue, archaeal *Sulfolobus* turreted icosahedral virus (STIV); magenta, a satellite virus, Sputnik, of mamavirus that is associated with a eukaryotic host. The group of nucleocytoplasmic large DNA viruses (NCLDVs) is indicated. The asterisk indicates that an X-ray structure is available for the major capsid protein of a representative virus.

analyses thus can be utilized to reveal deep evolutionary connections between these viruses. This information can then be used to establish a higher-level taxonomy for viruses that are considered to share an ancestor.

Why would it be advantageous to have an additional higher level in virus taxonomy, and how would it move forward our understanding of virus origin and evolution? The grouping of evolutionarily related viruses with hosts from different domains of life indicates that the viral universe is ancient, dating back at least to the last universal common ancestor (LUCA) of the cells. Furthermore, the presence of several distinct viral lineages leads to the conclusion that the LUCA was already infected by several types of viruses (progenitors of current viral lineages), some of which were already rather sophisticated and complex (capable of forming empty procapsids that were subsequently filled with nucleic acids with the aid of virus-encoded packaging NTPases). If so, the contemporary virosphere, which has evolved from a set of viruses that were grazing on the LUCA, is polyphyletic in its origins as opposed to the monophyletic origin of cellular organisms.

The assignment of virus families to lineages would bring a new dimension to viral classification by allowing evolutionarily related viruses, which have diversified beyond recognition by standard sequence-targeted approaches, to be rightfully grouped together. This in turn would bring more order to the virosphere and would more accurately reflect the evolution of viruses rather than the coevolution of viruses with their hosts. Even though the antiquity of viruses is generally appreciated, the current virus classification does not reflect this point. The only exception is the order *Caudovirales* (tailed dsDNA viruses infecting bacteria and archaea). However, even in this case, the horizontal interdomain spread of tailed dsDNA viruses from bacteria to archaea has been favored to explain the presence of these evolutionarily related viruses in the two cellular domains (29). The assignment of viruses infecting evolutionarily distant hosts into lineages would therefore unify viruses into a more continuous virosphere and would abrogate the tendency to perceive the viral universe as a diffuse assemblage composed of three distinct groups of noncellular entities, each associated with its own domain of cellular organisms.

It is obvious that not all virus families are equally well covered by structural and functional analyses, nor have we sampled the virosphere extensively enough to be able to place all currently recognized virus families into distinct lineages. In addition, it might prove difficult to find a criterion which could be universally used to assign pleomorphic viruses (e.g., influenza virus) to structure-based higher taxonomic levels (2). These caveats notwithstanding, we are confident that it is time to acknowledge the antiquity of viruses by implementing a higher-level classification scheme that would gradually bring order to the viral universe.

ACKNOWLEDGMENT

This work was supported by the Finnish Center of Excellence Program (2006–2011) of the Academy of Finland (grant 1129684 to D.H.B.).

REFERENCES

1. **Abrescia, N. G., J. J. Cockburn, J. M. Grimes, G. C. Sutton, J. M. Diprose, S. J. Butcher, S. D. Fuller, C. San Martin, R. M. Burnett, D. I. Stuart, D. H. Bamford, and J. K. Bamford.** 2004. Insights into assembly from structural analysis of bacteriophage PRD1. Nature **432:**68–74.
2. **Abrescia, N. G., J. M. Grimes, E. E. Fry, J. J. Ravantti, D. H. Bamford, and D. I. Stuart.** 2010. What does it take to make a virus: the concept of the viral "self," p. 35–58. *In* P. G. Stockley and R. Twarock (ed.), Emerging topics in physical virology. Imperial College Press, London, United Kingdom.
3. **Abrescia, N. G., J. M. Grimes, H. M. Kivelä, R. Assenberg, G. C. Sutton, S. J. Butcher, J. K. Bamford, D. H. Bamford, and D. I. Stuart.** 2008. Insights into virus evolution and membrane biogenesis from the structure of the marine lipid-containing bacteriophage PM2. Mol. Cell **31:**749–761.
4. **Baker, M. L., W. Jiang, F. J. Rixon, and W. Chiu.** 2005. Common ancestry of herpesviruses and tailed DNA bacteriophages. J. Virol. **79:**14967–14970.
5. **Bamford, D. H.** 2003. Do viruses form lineages across different domains of life? Res. Microbiol. **154:**231–236.
6. **Bamford, D. H., J. M. Grimes, and D. I. Stuart.** 2005. What does structure tell us about virus evolution? Curr. Opin. Struct. Biol. **15:**655–663.
7. **Boyer, M., N. Yutin, I. Pagnier, L. Barrassi, G. Fournous, L. Espinosa, C. Robert, S. Azza, S. Sun, M. G. Rossmann, M. Suzan-Monti, B. La Scola, E. V. Koonin, and D. Raoult.** 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc. Natl. Acad. Sci. U. S. A. **106:**21848–21853.
8. **Duda, R. L., R. W. Hendrix, W. M. Huang, and J. F. Conway.** 2006. Shared architecture of bacteriophage SPO1 and herpesvirus capsids. Curr. Biol. **16:**R11–13.
9. **Fauquet, C. M., M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (ed.).** 2005. Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, United Kingdom.
10. **Filée, J., and M. Chandler.** 2010. Gene exchange and the origin of giant viruses. Intervirology **53:**354–361.
11. **Forterre, P.** 2010. Defining life: the virus viewpoint. Orig. Life Evol. Biosph. **40:**151–160.
12. **Forterre, P.** 21 August 2010, posting date. Manipulation of cellular syntheses and the nature of viruses: the virocell concept. Comptes Rendus Acad. Sci. doi:10.1016/j.crci.2010.06.007.
13. **Forterre, P., and D. Prangishvili.** 2009. The origin of viruses. Res. Microbiol. **160:**466–472.
14. **Hendrix, R. W.** 2002. Bacteriophages: evolution of the majority. Theor. Popul. Biol. **61:**471–480.
15. **Iyer, L. M., L. Aravind, and E. V. Koonin.** 2001. Common origin of four diverse families of large eukaryotic DNA viruses. J. Virol. **75:**11720–11734.
16. **Jalasvuori, M., and J. K. Bamford.** 2008. Structural co-evolution of viruses and cells in the primordial world. Orig. Life Evol. Biosph. **38:**165–181.
17. **Khayat, R., L. Tang, E. T. Larson, C. M. Lawrence, M. Young, and J. E. Johnson.** 2005. Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. Proc. Natl. Acad. Sci. U. S. A. **102:**18944–18949.
18. **Koonin, E. V.** 2009. On the origin of cells and viruses: primordial virus world scenario. Ann. N. Y. Acad. Sci. **1178:**47–64.
19. **Koonin, E. V., Y. I. Wolf, K. Nagasaki, and V. V. Dolja.** 9 February 2009. The complexity of the virus world. Nat. Rev. Microbiol. [Epub ahead of print]. doi:10.1038/nrmicro2030-c2.
20. **Koonin, E. V., and N. Yutin.** 2010. Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology **53:**284–292.
21. **Krupovič, M., and D. H. Bamford.** 2009. Does the evolution of viral polymerases reflect the origin and evolution of viruses? Nat. Rev. Microbiol. **7:**250.
22. **Krupovič, M., and D. H. Bamford.** 2007. Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria. BMC Genomics **8:**236.
23. **Krupovič, M., and D. H. Bamford.** 2008. Virus evolution: how far does the double beta-barrel viral lineage extend? Nat. Rev. Microbiol. **6:**941–948.
24. **Krupovič, M., P. Forterre, and D. H. Bamford.** 2010. Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. J. Mol. Biol. **397:**144–160.
25. **Krupovič, M., S. Gribaldo, D. H. Bamford, and P. Forterre.** 25 June 2010, posting date. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitch-hiking of mobile genetic elements. Mol. Biol. Evol. doi:10.1093/molbev/msq161.
26. **Lima-Mendez, G., J. Van Helden, A. Toussaint, and R. Leplae.** 2008. Reticulate representation of evolutionary and functional relationships between phage genomes. Mol. Biol. Evol. **25:**762–777.
27. **Ludmir, E. B., and L. W. Enquist.** 2009. Viral genomes are part of the phylogenetic tree of life. Nat. Rev. Microbiol. **7:**615.
28. **Nandhagopal, N., A. A. Simpson, J. R. Gurnon, X. Yan, T. S. Baker, M. V. Graves, J. L. Van Etten, and M. G. Rossmann.** 2002. The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. Proc. Natl. Acad. Sci. U. S. A. **99:**14758–14763.
29. **Prangishvili, D., P. Forterre, and R. A. Garrett.** 2006. Viruses of the Archaea: a unifying view. Nat. Rev. Microbiol. **4:**837–848.
30. **Steven, A. C., J. B. Heymann, N. Cheng, B. L. Trus, and J. F. Conway.** 2005. Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. Curr. Opin. Struct. Biol. **15:**227–236.
31. **Strömsten, N. J., D. H. Bamford, and J. K. Bamford.** 2005. In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses. J. Mol. Biol. **348:**617–629.
32. **Sun, S., B. La Scola, V. D. Bowman, C. M. Ryan, J. P. Whitelegge, D. Raoult, and M. G. Rossmann.** 2010. Structural studies of the Sputnik virophage. J. Virol. **84:**894–897.
33. **Suttle, C. A.** 2007. Marine viruses—major players in the global ecosystem. Nat. Rev. Microbiol. **5:**801–812.
34. **Villarreal, L. P., and G. Witzany.** 2010. Viruses are essential agents within the roots and stem of the tree of life. J. Theor. Biol. **262:**698–710.
35. **Weigel, C., and H. Seitz.** 2006. Bacteriophage replication modules. FEMS Microbiol. Rev. **30:**321–381.
36. **Wikoff, W. R., L. Liljas, R. L. Duda, H. Tsuruta, R. W. Hendrix, and J. E. Johnson.** 2000. Topologically linked protein rings in the bacteriophage HK97 capsid. Science **289:**2129–2133.
37. **Xiao, C., Y. G. Kuznetsov, S. Sun, S. L. Hafenstein, V. A. Kostyuchenko, P. R. Chipman, M. Suzan-Monti, D. Raoult, A. McPherson, and M. G. Rossmann.** 2009. Structural studies of the giant mimivirus. PLoS Biol. **7:**e92.
38. **Yan, X., N. H. Olson, J. L. Van Etten, M. Bergoin, M. G. Rossmann, and T. S. Baker.** 2000. Structure and assembly of large lipid-containing dsDNA viruses. Nat. Struct. Biol. **7:**101–103.

The views expressed in this Commentary do not necessarily reflect the views of the journal or of ASM.