



A brief history of the species concept in virology and an opinion on the proposal to introduce Linnaean binomial virus species names

M. H. V. Van Regenmortel¹

Published online: 6 October 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

The species concept used in virology is based on the logic of the Linnaean hierarchy, which views a species class as the lowest abstract category that is included in all the higher categories in the classification, such as genera and families. As a result of this class inclusion, the members of a species class are always less numerous than the members of higher classes, which become more numerous as one moves up in the hierarchy. Because species classes always have fewer members than any of the higher classes, logic requires that they need more qualifications for establishing membership than any of the higher classes. This invalidates the claim that a species could be defined by a single property present in all its members. Species were only accepted in virus classification in 1991, because virologists assumed that it would lead to the use of Latin species names, which they rejected. Anglicized binomial species names have been used by virologists for more the 40 years and are popular because they consist of a virus name followed by a genus name that most virologists are familiar with. The ICTV has proposed to introduce a new Latinized virus species binomial nomenclature using the genus name followed by a hard-to-remember Latinized species epithet that bears little resemblance to the name of the virus itself. However, the proposal did not clarify what the advantage is of having to learn hundreds of new unfamiliar virus species names. In 2013, the ICTV changed the definition of a virus species as an abstract class and defined it as a group of physical objects, which induced virologists to believe that a virus species could be defined by a few characteristics of the viral genome. In recent years, thousands of viral sequences have been discovered in metagenomic databases, and the ICTV has suggested that it should be possible to incorporate these sequences in the current ICTV virus classification. Unfortunately, the relational properties of these hypothetical viruses that result from their biological interactions with hosts and vectors remain in the vast majority of cases totally unknown. The absence of this information makes it in fact impossible to incorporate these metagenomic sequences in the current classification of virus species.

Introduction

Since viruses are biological objects that commonly infect living organisms, it was natural for virologists to want to classify viruses by using the categories species, genus, family, order, etc. utilized in the hierarchical biological classifications of plants, animals and microorganisms. The root of the word classification is “class”, a term that refers to all the abstract classes and categories to which real individual organisms and viruses are assigned. Class membership is the logical relation that creates a bridge between any of these

conceptual class constructs and the real objects that are its members.

A virus species cannot be defined by a single property present in all its members

The lowest rank is the species class, and its members are automatically members of some of the higher classes that include it. This class inclusion is a logical peculiarity of all hierarchical biological classifications [1], and it obviates the need to repeat the additional properties that a member of a species acquires by virtue of also being a member of a higher taxon such as a genus or a family. This means that these properties of the higher classes above it are also present in members of the species, although they are obviously not species-defining properties that would allow individual species

Handling Editor: Sead Sabanadzovic.

✉ M. H. V. Van Regenmortel
vanregen@unistra.fr

¹ Medical University of Vienna, Vienna, Austria

within a genus or a family to be differentiated. Because of class inclusion, higher taxa such as virus families always have more virus members than lower taxa such as species, and they therefore require fewer properties (for instance, virion structure or replication strategy) to meet the qualification for membership. The logic of the Linnaean hierarchy [1] implies that increasing the number of qualifications decreases membership (in species), whereas decreasing that number increases membership (e.g., in families). A single property such as being a genetic parasite would for instance suffice to define a class that includes all known viruses. This logical principle invalidates the erroneous claim of many virologists [2–4] that it is possible to define a virus species by a single property present in all its members, such as one genetic metric or nucleotide combination motifs, since a species can only be defined by using the variable combination of several properties that characterizes a polythetic class [5].

Species definitions and species names

Although the term species is universally used in all biological classifications, it is remarkable that, 160 years after the publication of Darwin's *On the Origin of Species*, biologists still do not agree on what a plant, animal or microbial species is [6–8]. In 1997, as many as 22 different species concepts were discussed in the proceedings of an international conference [9]. When ICTV members decided to introduce the species concept in virology, they were confronted with a variety of definitions that were not applicable to viruses that replicate as clones, such as the common definition that species are populations of organisms that can only breed among themselves. Several definitions of virus species had been proposed in the 1980s, but none had gained general acceptance [5]. Furthermore, many virologists were opposed to the introduction of species in virus classification because they argued that it would bring about the use of Latin species names, which they strongly opposed [10]. The ICTV had initially advocated a Latinized viral nomenclature [11], but it developed its own rules and a Code that differed from the Biological Code of Nomenclature used in the rest of biology [12] regarding the use of italics and capitals and the formation of binomial names for virus species. Instead of adopting the Linnaean format of Latin binomial species names using the order genus name first/species identifier second, the ICTV in 1976 introduced Anglicized binomial names using the inverse order of species name first and genus name second [13]. A past ICTV president [14] wrote that nothing released adrenalin more readily for many Anglo-Saxon virologists than the suggestion that the names of viruses and virus species should be Latinized.

A definition of virus species as a polythetic class proposed in 1989 was finally adopted by the ICTV [15] only

when the rules regarding the possible use of Latin in viral nomenclature had been abolished [16]. The following definition was endorsed: A virus species is a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche. The main novelty of this definition was that it used the notion of polythetic class that had become established in taxonomy [17]. Whereas monothetic classes are universal classes defined by one or very few properties that are both necessary and sufficient for establishing membership in the class, polythetic classes are defined by a variable combination of numerous properties, none of which is a defining property necessarily present in all the members of the class. This is in line with the logic of the Linnaean hierarchy that a species cannot be defined by a single property that is present in all the members of the class and absent in the members of other classes. This definition has been used by the ICTV for more than 30 years for establishing hundreds of virus species [18]. Some virologists, however, objected to the use of the term polythetic because they thought that it referred to a property present in all the members of the class instead of describing a particular distribution of properties within the class [2]. If it had been a common property in all the members of the class, this would of course have led to the paradox that a polythetic class was actually a monothetic one [5, 19].

In 1998, the ICTV Executive Committee modified the International Code of Virus Classification and Nomenclature and decided that the common English names of viruses would in future become the names of the corresponding virus species by italicizing these names with an initial capital letter in order to provide a visible sign that species are taxonomic classes like italicized genera and families [20]. Measles virus was thus assigned to the species *Measles virus*. This unfortunate decision was soon found to lead to considerable confusion because virologists found it difficult to distinguish between a viral object and a taxonomic species concept only on the basis of typography. As a result, they often wrote that *Measles virus* had been isolated, transmitted or sequenced, as if an abstract taxonomic concept could have a host, a vector or a sequence.

Many plant virologists had in fact been using non-Latinized binomial names (NLBNs) for virus species since the 1970s because these were easier to distinguish from virus names because all biologists are familiar with the fact that binomial names refer to taxonomic species entities. NLBNs had been introduced by the ICTV by Fenner [13] and were commonly used in plant virology papers and books [21–23] as well as in ICTV Reports [24, 25]. Such NLBNs are formed by replacing the terminal word “virus” that occurs in all common English names of viruses with the genus name to which the virus is assigned which also ends in “virus”. Measles virus was thus assigned to the species *Measles morbillivirus*.

Since every virus has a vernacular name that can be used for referring to the infectious agent under study, virologists are able to avoid the common logical problem that is faced by all biologists who have to use a taxonomic Latin species name for referring to a physical member of the millions of species of living organisms that have no common name. The resulting confusion between species as an abstract class and as a concrete object is common in the whole of biology [26] and has even convinced some virologists that there was no need to keep distinguishing between these two different uses of the term species [27, 31].

To make it easier to differentiate between individual viruses and virus species, a proposal was made to generalize the use of NLBNs for all the species that had been created by the ICTV [28]. Hundreds of such names were already in common use and many more were introduced subsequently [29, 30]. These names are popular because they are based on the combination of familiar virus and genus names known by most virologists. Since the major journals and reference books in virology are written in English, which is the predominant communication language used by scientists, most virologists are familiar with English virus names.

Most of the properties useful for defining virus species are the so-called relational properties of the members assigned to a particular species. Such properties arise because of the multiple relations that viruses develop with their biological partners such as hosts and vectors. Relational properties take the form of transmission vectors, host range, disease symptoms, cell and tissue tropism, pathogenicity and immunoreactivity. These properties are the main reasons why a reliable classification and taxonomy of viruses is necessary to allow virologists to communicate with non-virologists such as quarantine officials, regulators, advisors and medical/veterinary professionals who are mainly interested in the relational properties of viruses. Since these properties arise from multiple interactions between the gene products of viruses and of various partners, they can be altered by a few mutations occurring in the viral genome or in the genome of the vector or host. Interactions between modified protein gene products may therefore be altered that may lead to altered relational properties of the virus that cannot be predicted from the viral genome only [31].

In 2012, the ICTV proposed the following new definition of virus species: A virus species is a monophyletic group of viruses whose properties can be distinguished from those of other species by multiple criteria [32]. This definition rejected the traditional view that a biological classification consists of hierarchical constructs of abstract classes that have concrete objects as their members and accepted that viruses do constitute a lineage that must have inherited shared nucleic acid properties from a common ancestor such as nucleotide combination motifs. The viruses were thus linked to the species group by a part-whole relation in

the way that limbs are part of a body. Properties and classes were no longer viewed as related abstract entities, and it was no longer accepted that whatever is said about a thing logically ascribes a property to it [33]. Characters and traits are often used in the sense of both property and part, although nucleotide combination motifs in a viral sequence are parts of a virus and not a property of it [5, 34, 35]. Many objections to the new definition of virus species were posted on the ICTV website, and the ICTV responded with a 4-page, polemical document that tried to defend the new definition [5, 36]. The full ICTV online discussion of 2012 is accessible [36].

Many virologists posted comments on the website that were in favor of the previous definition of virus species as a polythetic class, and a group of 14 senior virologists pointed out many defects of the new species definition, for instance that it applied equally to genera and did not provide any guidelines for establishing new species [37]. The ICTV Executive Committee nevertheless had their new species definition ratified by a fast-track approval process that considerably reduced the time available for posting further objections on the ICTV website [3].

Classification of viral genomes

As more sequences of viral genomes became available, increasing attempts were made to establish new virus species only by considering a part of the genome sequence, and it was overlooked that this produced a phenotypic classification of viral genomes based on molecular sequences rather than a classification of virus species based on the defining relational properties of its members. It is of course possible to classify viral nucleotide sequences on the basis of characteristics such as genome compositional features and organization, gene content, particular nucleotide motifs, and inferred replication strategies, although this would produce a classification of nucleotide sequences or of viral genomes but not a classification of viruses [38].

A reply to the consultation of Siddell et al. [39]

The ICTV EC has recognized that the current situation regarding the names of viruses and of virus species that differ only by typography is no longer acceptable and that a binomial nomenclature for virus species should be established. The current non-Latinized binomial nomenclature consists in giving species a name composed of two parts corresponding to the virus name followed by the genus name, with the species name sometimes containing more than one word. Hundreds of such binomial names (NLBNs)

of virus species had been introduced in this manner since 1976, for instance *Measles morbillivirus* and *Tomato spotted wilt tospovirus*. However, the ICTV in 2016 initiated what they called a thought exercise in which they converted currently existing 175 NLBNs into an inverse Latinized Linnaean binomial format that consists of the genus name to which the species belongs followed by a Latinized species epithet [40]. When the species name contains more than one word, it is necessary to alter the species name significantly in the Latinized epithet with the result that it may then be no longer evident which species is referred to. The two following examples of Latinized Linnaean binomial species names (LLBNs) compared to NLBNs names and their species names illustrate this difficulty:

Adelaide River virus (virus), *Adelaide River ephemero-virus* (NLBN), *Ephemerovirus fumenadelaidense* (LLBN)

Merino Walk virus (virus), *Merino Walk mammarenavirus* (NLBN), *Mammarenavirus viamerinense* (LLBN)

It is immediately apparent that the similarity between the virus name and the NLBN using a known genus name is easy to remember, whereas introducing hundreds of new Latinized LLBNs epithets that are difficult to pronounce and to remember is likely to make it difficult to recognize which is the species that is being referred to. One argument offered in favour of LLBNs is that virus taxonomy is currently excluded from many bioinformatics projects such as the BioCode because the software developed for LLBNs would assume that when a NLBN name like *Lassa mammarenavirus* is used, it would lead the software to assume that “Lassa” is the genus and mammarenavirus is the species epithet. Virologists have always used their own rules and Code of nomenclature, and many of them accept that using NLBNs for virus species is appropriate because they consider viruses to be non-living genetic parasites rather than living organisms. Furthermore, it would also be possible to develop software that would recognize that, in the case of virus species, the species name comes first and the genus name comes second. The thought exercise of Postler et al. [40] certainly demonstrated that it is possible to coin LLBNs, but it did not demonstrate what their advantage is compared to the NLBNs that virologists are familiar with.

Siddell et al. [39] stated that devising LLBNs *would be particularly problematic for viruses “identified” metagenomically because these “may” entirely lack the phenotypic information that assists classification elsewhere in biology*. In fact, the overwhelming majority of viruses that hide in metagenomic databases have not yet been identified at all, and since their hosts and vectors are unknown, it would have been more relevant to admit that in such cases it is actually impossible to know the relational properties of such putative viruses or to incorporate them in the current scheme of ICTV species defined by the phenotypic and relational properties of their members.

A reply to Hull and Rima [41]

In their comments to the consultation paper of Siddell et al. [39], Hull and Rima [41] recommended that the members of the ICTV should continue to develop the very useful taxonomy of viruses that had been developed during the past 50 years. This seemed preferable to becoming embroiled in a counterproductive debate regarding the relative merits of a newly developed Latinized binomial species nomenclature that follows the pattern introduced by Linnaeus several centuries ago (a genus name followed by a Latinized epithet) instead of maintaining the Anglicized binomial species nomenclature that was introduced by the ICTV 40 years ago. Hull and Rima also referred to the possible existence of 10^8 - 10^9 viral species that could possibly hide within the extensive metagenomic databases that have so far been sequenced, and they labelled such sequences “sequence-species”, although such a name is at odds with accepted taxonomic terminology. There is no agreed procedure for incorporating such “sequence-species” within the class of the 5560 species currently characterized by the ICTV on the basis of unknown relational biological properties of such putative members. It has been convincingly argued that, since there is no knowledge of the relational properties of the members of the vast majority of sequence-species found in metagenomic databases, it is in fact impossible to assign them to the ICTV species class [31]. As long as this problem has not been solved, it seems premature to want to create a Latinized species nomenclature for thousands of such unknown entities.

Although Simmonds and Aiewakun [42] claim that “sequence-only viruses are still viruses”, this makes little sense, since sequence-only viruses are actually only sequences if there is nothing else to describe except nucleic acid sequences. Unfortunately, their conclusion that this leads to a seemingly irreconcilable dilemma does not point to a workable solution.

References

1. Buck RC, Hull DL (1966) The logical structure of the Linnaean hierarchy. *Syst Zool* 15:97–111. <https://doi.org/10.2307/2411628>
2. Gibbs AJ, Gibbs MJ (2006) A broader definition of the ‘virus species’. *Arch Virol* 151:1419–1422. <https://doi.org/10.1007/s00705-006-0775-2>
3. Adams MJ, Lefkowitz EJ, King AMQ, Carstens EB (2013) Recently agreed changes to the International Code of Virus Classification and Nomenclature. *Arch Virol* 158:2633–2639
4. Simmonds P, Adams MJ, Benkő M et al (2017) Consensus statement: Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* 15:161–168

5. Van Regenmortel MHV (2016) Classes, taxa and categories in hierarchical virus classification: a review of current debates on definitions and names of virus species. *Bionomina* 10:1–21
6. Stamos DN (2003) The species problem. Biological species, ontology and the metaphysics of biology. Lexington Books, Oxford
7. Ereshefsky M (2009) Darwin's solution to the species problem. *Synthese* 175:405–425. <https://doi.org/10.1007/s11229-009-9538-4>
8. Claridge MF (2010) Species are real biological entities. In: Ayala FJ, Arp R (eds) *Contemporary debates in philosophy of biology*. Wiley, Chichester, pp 91–109
9. Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge MF, Dawah HA, Wilson MR (eds) *Species: the units of biodiversity*. Chapman and Hall, London, pp 381–424
10. Milne RG (1984) The species problem in plant virology. *Microbiol Sci* 1:113–122
11. Wildy P (1971) Classification and nomenclature of first report of the international committee on nomenclature of viruses. *Monographs in virology*, vol 5. Karger, Basel
12. Greuter W, Garrity G, Hawksworth DL et al (2011) Draft BioCode (2011). Principles and Rules regulating the naming of organisms. New draft, revised in November 2010. *Bionomina* 3:26–44. <https://doi.org/10.1164/bionomina.3.1.3>
13. Fenner F (1976) The classification and nomenclature of viruses. Second Report of the International Committee on Taxonomy of Viruses. *Intervirology* 7:1–115. <https://doi.org/10.1159/000149938>
14. Matthews REF (1985) Viral taxonomy. *Microbiol Sciences* 2:74–75
15. Pringle CR (1991) The 20th meeting of the executive committee of the ICTV. Virus species, higher taxa, a universal database and other matters. *Arch Virol* 119:303–304. <https://doi.org/10.1007/BF01310680>
16. Francki RIB, Fauquet CM, Knudson DL, Brown F (1991) Fifth report of the international committee on taxonomy of viruses. *Arch Virol Suppl* 2:450. <https://doi.org/10.1007/978-3-7091-9163-7>
17. Hull DL (1976) Are species really individuals? *Syst Zool* 25:174–191. <https://doi.org/10.2307/2412744>
18. Ball LA (2005) The universal taxonomy of viruses in theory and practice. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) *Virus taxonomy. Eight Report of the International Committee on Taxonomy of Viruses*. Elsevier Academic Press, Amsterdam, pp 3–8. <https://doi.org/10.1016/b978-0-12-249951-7.50003-1>
19. Van Regenmortel MHV (2018) Viral species, viral genomes and HIV vaccine design: is the rational design of biological complexity a utopia? *Arch Virol* 163:2047–2054. <https://doi.org/10.1007/s00705-018-3955-y>
20. Mayo MA, Horzinek MC (1998) A revised version of the International Code of Virus Classification and Nomenclature. *Arch Virol* 143:1645–1654. <https://doi.org/10.1007/s007050050406>
21. Matthews REF (1971) *Plant virology*. Academic Press, San Diego
22. Brunt A, Crabtree K, Gibbs A (1990) *Viruses of tropical plants*. CAB International, Wallingford
23. Bos L (1999) *Plant viruses, unique and intriguing pathogens—a textbook of plant virology*. Backhuys Publishers, Leiden
24. Matthews REF (1979) Classification and nomenclature. Third Report of the International Committee on Taxonomy of Viruses. *Intervirology* 7:1–115
25. Matthews REF (1982) Classification and nomenclature of viruses. Fourth Report of the International Committee on Taxonomy of Viruses. *Intervirology* 17:1–200. <https://doi.org/10.1159/000149278>
26. Ereshefsky M (2011) Mystery of mysteries: Darwin and the species problem. *Cladistics* 27:67–79
27. Simmonds P (2018) A clash of ideas - the varying uses of the 'species' term in virology and their utility for classifying viruses in metagenomic datasets. *J Gen Virol* 99:277–287
28. Van Regenmortel MHV, Burke DS, Calisher CH et al (2010) A proposal to change existing virus species names to non-Latinized binomials. *Arch Virol* 155:1909–1919
29. Bukreyev AA, Chandran K, Dolnik O et al (2014) Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Arch Virol* 159:821–830. <https://doi.org/10.1007/s00705-013-1846-9>
30. Kuhn JH, Dürrwald R, Bao Y et al (2015) Taxonomic reorganization of the family Bornaviridae. *Arch Virol* 160:621–632. <https://doi.org/10.1007/s00705-014-2276-z>
31. Van Regenmortel MHV (2019) Solving the species problem in viral taxonomy: recommendations on non-latinized binomial species names and on abandoning attempts to assign metagenomic viral sequences to species taxa. *Arch Virol* 164:2223–2229
32. King A, Adams M, Lefkowitz E, Carstens E (2012) ICTV Proposal 2011.002sg. In: 2011.002auG.A.v9.statute_and_code_changes.pdf (page 15) http://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/general-2008/4444.aspx. Accessed on 29 July 2020
33. Quine WV (1990) *Classes versus properties. Quiddities: an intermittently philosophical dictionary*. Penguin Books, London, pp 22–24
34. Frisrup K (1992) Character: Current Usages. In: Keller EF, Lloyd EA (eds) *Keywords in evolutionary biology*. MA (Harvard University Press), Cambridge, pp 45–51
35. Mahner M, Bunge M (1997) *Foundations of biophilosophy*. Springer, Berlin, pp 11–12. <https://doi.org/10.1007/978-3-662-03368-5>
36. https://talk.ictvonline.org/ictv1/f/general_ictv_discussion_s-20/3930/comments-to-proposed-modification-to-code-rule-3-21-defining-virus-species. Accessed on 29 July 2020
37. Van Regenmortel MHV, Ackermann H-W, Calisher CH et al (2013) Virus species polemics: 14 senior virologists oppose a proposed change to the ICTV definition of virus species. *Arch Virol* 158:1115–1119. <https://doi.org/10.1007/s00705-012-1583-5>
38. Van Regenmortel MHV (2017) Only viruses, but not their genome sequences, can be classified in hierarchical species and genus classes. *Curr Top Virol* 13:59–67
39. Siddell SG, Walker PJ, Lefkowitz EJ et al (2020) Binomial nomenclature for virus species: a consultation. *Arch Virol* 165:512–525. <https://doi.org/10.1007/s00705-020-04555-0>
40. Postler TS, Clawson AN, Amarasinghe GK et al (2017) Possibility and challenges of conversion of current virus species names to Linnaean binomials. *Syst Biol* 66:463–473
41. Hull R, Rima B (2020) Virus taxonomy and classification: naming of virus species. *Arch Virol*. <https://doi.org/10.1007/s00705-020-04748-7>
42. Simmonds P, Aiewsakun P (2018) Virus classification—where do you draw the line? *Arch Virol* 163:2037–2046. <https://doi.org/10.1007/s00705-018-3938-z>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.