



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

Marc H. V. Van Regenmortel¹

Received: 22 March 2019 / Accepted: 20 May 2019 / Published online: 18 June 2019
© Springer-Verlag GmbH Austria, part of Springer Nature 2019

Abstract

Properties useful for defining virus species are phenotypic properties of viruses that can be altered by a few mutations. Such properties include the natural host range, cell and tissue tropism, symptomatology, pathogenicity and mode of transmission. All these properties are not necessarily present in identical form in all the members of a species; therefore, a virus species is a polythetic class of viruses defined by a variable combination of several properties rather than by a single conserved property present in all the members of the species. This review will discuss current controversies about what virus species actually are as well as which names should be given to them. It will be emphasized that most species-defining properties are so-called relational properties that arise because viruses necessarily interact with biological partners such as vectors, hosts and immune systems. Although these relational properties are of utmost importance to laboratory and clinical virologists, they remain unknown if only the viral genome is available and the relational partners of the virus have not yet been identified. Since the International Committee on Taxonomy of Viruses (ICTV) in 2013 ratified a new definition of virus species, which no longer accepts that species are polythetic classes but instead are monophyletic groups, the implications of this new definition for viral taxonomy and nomenclature will be analyzed. In my private capacity, I also make the following recommendations regarding current debates on proposed new names for virus species as well as on the feasibility of assigning viral sequences found in metagenomic databases to individual species taxa in the current ICTV classification.

- 1) The ICTV should abandon the current rule that the names of virus species (for instance *Measles virus*) should differ from the virus name (measles virus) only by typography.
- 2) Non-Latinized binomial species names based on familiar virus and genus names should become the norm. This would obviate the need to create about 5000 hard-to-memorize Latinized species names.
- 3) Virus species are defined not by the intrinsic properties of virions and viral genomes but by the relational properties of viruses that arise from their interactions with host and vector partners. Since the hosts and vectors associated with nearly all viral sequences found in metagenomic databases are unknown, the phenotypic properties of the putative viruses also remain unknown,

and these viral sequences cannot be allocated to established species in the ICTV classification

Introduction

All biological classifications comprise conceptual creations of the human mind. The root of the word classification is “class”, a term that refers to all the classes of organisms or viruses that have “real” biological objects as their members. Taxonomy deals with taxa, which are the individual classes defined by taxonomists in their classification of animals, plants, microbes and viruses.

Class membership is the logical relation that facilitates the establishment of a bridge between ideas or mental constructs and the actual real objects to which such concepts refer. Biological taxonomies always use a hierarchy of classes or ranks such as species, genera and orders; the members of the lowest class, which is a species, are always automatically members of the higher classes above it. This class inclusion, a peculiarity of biological classification,

Handling Editor: Tim Skern.

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

¹ Medical University of Vienna, Vienna, Austria

obviates the need to repeat the properties used to define higher taxa when defining the lower taxa that they include. Since the species *Measles virus* is included in the genus *Morbillivirus*, the properties of the viruses that are members of the species include all the properties required for classifying them as members of the genus, in addition to the species-defining properties that allow the members of *Measles virus* to be distinguished from the members of other species in the same genus. Higher taxa such as families always have more virus members than lower taxa and require fewer properties (for instance, virion structure or replication strategy) than lower species taxa to meet the qualification for membership. The logical principle that increasing the number of qualifications decreases membership whereas decreasing that number increases the membership [1] is ignored by many virologists and has led to the erroneous claim that a single property is sufficient to define what are in fact spurious monothetic virus species [2–4].

Virus species as category, polythetic class and rank

All viruses are assigned to a species taxon, which is the lowest class in hierarchical biological classifications, and all the established virus species taxa form the species category, which is the class of all species taxa. The membership relation that individual viruses have to taxa is the same relation that taxa have to their respective categories, the members of the category species being all the species taxa. As discussed elsewhere [3], the species category was introduced in virus classification only in 1991 [5], when the following definition of virus species was endorsed by the ICTV: *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 included the notion of polythetic class that was already commonly used by taxonomists. Whereas monothetic classes are universal classes defined by one or two properties that are both necessary and sufficient for membership (for instance in a family), polythetic classes are defined by a variable combination of properties, none of which is a defining property that must necessarily be present in every member of the species. Every member of a polythetic class thus always shares only a limited number of these defining properties, which can be altered by a few mutations, such as the natural host range, cell and tissue tropism, pathogenicity, mode of transmission, and small differences in the genome. As a result, not all of these properties will necessarily be present in identical form in all members of the same species. This is the reason why species are not monothetic classes but are defined polythetically by a variable combination of a small number of

properties rather by a single conserved property necessarily present in all members of the species [6].

Gibbs [7] has repeatedly claimed that the ICTV had never in practice defined a virus species using the polythetic definition, an assertion that is patently untrue [8, 9]. Instead, Gibbs proposed that virus species should be defined monothetically by conserved nucleotide combination (NC) motifs present in their members, which they must have necessarily inherited from common ancestors [2]. The reason why that proposal is impossible to implement is that it would be necessary, before establishing a new species, to ascertain whether such a motif is present in all the members of the species and absent in others, a feat which obviously cannot be realized (this also applies to metagenomic sequences) [6, 10, 11]. A polythetic virus species must therefore first be demarcated and established by taxonomists by listing several of the properties of its members before it becomes possible to distinguish members from non-members by the presence of a sufficient number of species-defining properties. A new species cannot be established on the basis of a single NC motif, since all the members of a species need not necessarily have exactly the same sequence or possess a particular nucleotide motif. Only because the species *Tobacco mosaic virus* had previously been established were Gibbs et al. [12] able to demonstrate that the genome sequences of many of its members for instance possessed an NC motif of 47 nucleotides which could then be used as a diagnostic marker for identifying new members of the established species. However, if an NC motif is found to be present in a few, unidentified viral sequences, it cannot serve as a species-defining property on its own for creating a new species. It is interesting to note that Darwin considered that species were simply collections of related organisms that naturalists grouped together on the basis of a few shared features and that he believed that there were no clear-cut boundaries between individual species. His unwillingness to argue over the definition of species has been described as a modern solution to the species problem [13].

Although the ICTV in 2013 rejected the definition of a virus species as a conceptual polythetic class (see section 4), it recently accepted that a virus species is the lowest rank in virus classification below the ranks of genus, family and order [14]. Ranks are hierarchical categories invented by humans, such as the ranks of lieutenants and generals used in armies, and these ranks obviously should not be confounded with the hundreds of individual lieutenants and generals that exist in an army. Although all the members of lower biological classes are also members of higher classes (see section 1), this does not apply to the members of an army rank, since all the lieutenants are not also members of the rank of general!

Having accepted the notion that a virus species is a conceptual, man-made rank, the ICTV should in future be able

to accept that a species is also a conceptual class that has “real” members, which are the viruses studied by virologists.

Reification is one of the most common logical fallacies of human thinking

Confusing the abstract class of virus species with the concrete viral objects that are its members is an example of the common logical fallacy of reification (from the Latin *res* meaning an object), which consists in treating ideas, concepts and properties as if they were real physical entities and objects. Virus classification makes use of a number of categories and ranks that are conceptual constructions of the mind and not real physical objects. In the case of genera and families, virologists have no difficulty in accepting that these categories are conceptual constructions, and they do not confuse them with the real objects they handle in their laboratories. They know that a virus family, for instance, cannot be purified by centrifugation, sequenced, visualized by electron microscopy or used to infect a host, since it is an imaginary entity of the mind and not a physical object. It is therefore surprising that many virologists tend to view a virus species as a kind of virus that is more real than a family because they think it is able to infect certain hosts. This confusion between a species as an abstract category in a classification and as a concrete object is actually prevalent in the whole of biology, partly because the vast majority of organisms in botany, zoology and microbiology do not have common names different from the Latin name of the species to which these organisms belong. Consequently, human infections caused by organisms belonging to the species *Escherichia coli* are attributed to the species *E. coli* as if a conceptual taxonomic entity could infect living organisms. It seems many virologists are not always aware that they are the only biologists who could easily avoid this type of logical confusion because every virus has a common name that differs from the species name to which the virus belongs [4]. Unfortunately, the ICTV Executive Committee decided in 1998 to ratify the rule that the name of a virus species (for instance *Measles virus*) would in future be the same as the name of the corresponding virus (measles virus) but written in italics with the initial letter capitalized. This new rule nullified the advantage of having separate (but nearly identical) names for viruses and viral species, and virologists continued to write erroneously that the species *Measles virus* for instance had infected a certain host. It should be noted that some virologists argue that their unique ability to clearly distinguish between viruses as objects and virus species as classes should be abandoned (perhaps because of the difficulty of differentiating between measles virus and *Measles virus*), and they even suggest that the logical confusion between the names of organisms and the names of

species that is prevalent in the rest of biology should also be adopted in virology to bring it into line with other biological disciplines [15]. These virologists are probably influenced by the new ICTV species definition (see section 4), which states that virus species are groups of viruses (i.e., concrete viral objects) that are physically part of the species by a part-whole relationship. For more than 30 years, I have defended the generally accepted view that all biological classifications are made up of conceptual classes and that species classes have organisms or viruses as their members [1, 3]. I find it regrettable that the ICTV has introduced a new definition of species which claims that virus species consist of concrete groups of objects which in practice can be demarcated on the basis of a single genetic metric, and rejected the classic taxonomic definition of a species as a polythetic class [8, 19].

The new ICTV definition of virus species regards species as individuals instead of classes

Virus classification deals with abstract classes of viruses that are conceptual constructions of the mind, and class membership is the relation that creates a link between the two logical categories of an abstract class and of its concrete members or referents that are objects located in space and time. This membership relation is different from the part-whole relationship that exists between two concrete objects, one being a part of the other in the way a limb is part of a body. It is in fact impossible for a viral object to be part of a conceptual species construct, and it is also logically impossible for a thought or concept to be part of a material object [16 p. 11].

Gibbs and Gibbs [2] proposed that the term polythetic should be removed from the 1991 definition of virus species because they failed to understand the term, taking polythetic to mean variable [17]. They thought that members of a polythetic species could potentially share no species-defining properties at all that would differentiate them from members of another species, with the result that it would then be impossible to determine whether a species actually constituted a distinct replicating lineage defined by a property inherited from its ancestors [15]. Such an erroneous interpretation obviously disregarded the definition of a polythetic species as a class comprising members that always share several defining properties with no single property being necessarily present in all its members. It must be emphasized that the term polythetic describes a particular distribution of properties present in a species class and that the viruses that are its members cannot possess polythetic properties. If one confuses a particular distribution of properties present in a class with a constant property of all its members, one inevitably lands with paradoxes, since a polythetic class would then be synonymous with a monothetic class [3].

In 2013, the ICTV adopted the view of Gibbs and Gibbs [2] and accepted that a virus species could be defined monothetically by the presence in all its members of a common nucleotide combination (NC) motif considered to be a species-defining property that was both necessary and sufficient for demarcating and establishing a species class. This view confuses the use of a diagnostic marker for identifying members of a previously established species with the demarcation of a new species by taxonomists using a single species-defining property. It also ignores the logical principle that decreasing the number of qualifications for membership would greatly increase the number of species taxon members, which is absurd (see section 1). It is, of course, higher taxa such as families and orders, which are defined by very few defining properties, that have very large numbers of members.

The following species definition was subsequently ratified by the ICTV [18]: *A virus species is a monophyletic group of viruses whose properties can be distinguished from those of other species by multiple criteria.* Although it was acknowledged that these criteria could be genetic properties as well as other phenotypic properties, the removal of the concept of polythetic class made it in fact possible to demarcate and establish new species only on the basis of a single genetic metric. This approach, which had become very fashionable because of the availability of enormous databases of viral genome sequences, led to virus species being demarcated in the absence of any phenotypic differential properties. Furthermore, since a viral genus is also a monophyletic group, the new species definition had lost any specificity because it was also a definition of genus or of any other monophyletic group. As could be expected, the introduction of this new definition gave rise to a heated debate involving many senior virologists who disagreed with the proposed change [19]. This debate has been recorded in full on the ICTV website [19, 20].

Choosing a particular genetic metric for building up a classification depends, of course, on which regions of the genome sequence are considered to be most important, and such choices can be very arbitrary. The best-known example is the inflated number of begomovirus species that were established on the basis of an arbitrarily chosen percentage (89%) of pairwise sequence identity in the viral DNA-A sequences of these viruses [21]. When one plots the frequency distribution of pairwise sequence comparisons (PASC) of all members of a virus family, one obtains multimodal distributions where the peaks are usually attributed to groups of strains, species or genera, without taking into account the phenotypic and biological characteristics of the individual viruses found within these peaks [6]. Initially, more than one hundred so-called species of begomoviruses had been demarcated that infected the same hosts, produced very similar disease symptoms and possessed no

clearly distinguishable properties that would have justified labelling them distinct species instead of strains [3]. Currently, the ICTV has approved the creation of more than 300 begomovirus species, although it is well known that the frequent occurrence of recombination and reassortment in geminiviruses gives rise to chimeric viruses with polyphyletic genomes that cannot be correctly represented in a single PASC distribution [8]. There is actually no precise degree of genome difference that can be used as a cutoff point to differentiate between two species or two genera [6], and there is also no clear-cut criterion to decide how far back in time an evolutionary species can be traced as a separate species.

Are species units of classification, units of evolution, or fuzzy populations of viruses?

Evolution is the unifying theory that dominates contemporary biology. Hence, it is often claimed that species are units of evolution and not units of classification, although it is commonly accepted that lineages related by ancestry and descent are the units of classification. The reproductive isolation of sexually reproducing organisms that provides a mechanism for explaining the cohesion of such lineages seems to support that view. However, asexual lineages are very common in biology, for instance in prokaryotes and viruses, and since it is then difficult to define lineages that coincide with a plausible phylogeny, this has led to the conclusion that a species is a unit of classification rather than a unit of evolution [22]. It is, however, not the species themselves that evolve but the lineages. Viewing lineages as real individuals is thus another clear instance of reification [16 p. 237-262]. Evolutionary change is not change in the properties of individual organisms or viruses but change over time in the distribution of genetic and phenotypic properties in a certain population of biological entities. This is also why the relation of ancestry that links parents to offspring is difficult to apply to bacteria and viruses that commonly utilize lateral gene transmission.

The absence of clear-cut discontinuities between individual virus species has led to the view that they correspond to fuzzy populations of viruses [10] instead of being clearly defined, recognizable entities that could be logically demarcated on the basis of specific inclusion and exclusion criteria [15]. Of course this is not an argument for doing away with species, since individual colours are nevertheless identifiable in the continuum of colours observed in a rainbow. On the other hand, the categories of family and order do appear to have a better foundation because they are based on a few more distinctive, stable and unifying genetic and structural properties than is the case with species.

Non-Latinized binomial names for virus species

Although the first ICTV reports advocated a Latinized viral nomenclature, Latin virus names were never officially implemented, and the rules regarding the use of Latin were soon abandoned [23–25]. Over the years, for instance, nine different Latin names had been proposed for tobacco mosaic virus [26, 27], but the ICTV wisely did not pursue the coining of Latin virus names since virologists were strongly opposed to their use [28]. The dislike for Latin names was in part responsible for the slow adoption of the species concept in virology because it was assumed that the acceptance of species would inevitably lead to Latin names of virus species [27]. Species were introduced in viral taxonomy in 1991 when the 5th ICTV report had abolished the use of Latin names [5, 23].

There is nowadays a general agreement that a more cogent set of species names should be utilized by the ICTV. Non-Latinized binomial names (NLBNs) for virus species have been used for many years in ICTV reports [23, 27, 28] and by plant virologists in their books and scientific papers [3]. These NLBNs are created by replacing the word “virus” that occurs in all English names of viruses (e.g., measles virus) by the genus name to which the virus belongs and which also ends in “-virus” (*Measles morbillivirus*). Since all species names in biology are binomials, virologists would certainly find it easier to remember that measles virus is the name of the virus while *Measles morbillivirus* is the name of the species. This would thus eliminate the current confusion between the virus measles virus and the species *Measles virus*. Since NLBNs are superior to the current official species names for distinguishing viruses from species, a proposal was made in 2010 to generalize the use of such binomial names for all virus species [29], and many such NLBNs have since then been created [30] even though the ICTV had rejected their general use in 1998.

Recently, a new proposal was published in the form of a tentative thought experiment to test the feasibility of converting all the species names in the family *Arenaviridae* and the order *Mononegavirales* into Linnaean binomials by using the format of a genus name followed by a species epithet [31]. Such a system would require the creation of 4853 new Latinized species epithets that would have to follow the rules of Latin grammar. The results of such a system are compared with the current format of NLBNs in Table 1, which reverses the order of species name first, genus name second, used in current viral NLBNs species names.

The similarity between the virus names and the NLBNs makes it easy to memorize them. They also correspond to virus and genus names in current use, whereas the 4853 new Latin epithets that would have to be coined are likely to be hard to memorize and therefore unlikely to be welcome by virologists [31, 32]. The NLBNs have been used in virology since their introduction by Fenner [23] more than 40 years ago and follow the reverse order of genus first/species second used by Linnaeus. However, virologists have always used their own rules and code that differ from the traditions in the rest of biology [24–28] for instance regarding Latinization or the use of italics in taxonomic names, and the advantage of using species binomial names is the same, independently of the order genus/species that is used. The advantages and rationale for justifying the introduction of thousands of new Latin species names in virology because they follow the historic Linnaean Tradition in use in biology are certainly not self-evident.

Intrinsic properties and relational properties of objects should not be confused

It is often claimed that, at least in principle, all the properties of a virus are entirely encoded by its genome, although it is in fact impossible to infer all the phenotypic characteristics of a virus from its genome sequence. The main reason for this is that most species-defining properties are relational rather than intrinsic properties of viruses (see section 8). These relational properties, which are also called emergent properties, emerge when the virus interacts with a transmission vector or with the cellular tissues and immune system of the host, and they, of course, cannot be predicted from the viral genome when the vectors or the hosts remain unknown. Epigenetic factors together with alternative splicing and the discarding of introns always lead to unpredictable RNA transcripts and viral proteins [33]. Since these viral proteins interact with vector and host gene products through mechanisms that have not been elucidated, we cannot know the causal connections between viral genes and the development of viral phenotypes. Small DNA and RNA viruses are composites of replication and structural gene modules that are reorganized during their evolutionary histories, which leads to unpredictable changes in host range and other viral properties [34]. Furthermore, there is evidence that the presence of structurally homologous capsid proteins in many DNA and RNA viruses shows that the evolutionary history of structural genes is distinct from that of non-structural genes [35]. A further reason that prevents viral phenotypes from being predicted from genotypes is that single

Table 1 Current virus names compared with NLBNs and Latin binomial species names

Virus name	NLBN	Latin binomial species name
Adelaide River virus	Adelaide River ephemerovirus	Ephemerovirus fiumenadelaidense
Merino Walk virus	Merino Walk mammarenavirus	Mamarenavirus viamerinense

reference genomes are no longer considered to be adequate for representing the total amount of genetic variation present in a viral species, since the pangenome (i.e., the entire set of genes possessed by all the members of a species) of a virus species may contain as much as 10% of genes that are not shared by all the sequenced genomes in the species [36].

Can metagenomic viral sequences be attributed to species taxa?

Predicting viral phenotypes from genotypes is also made impossible if intrinsic and relational properties are not clearly differentiated. Whereas virions and nucleotide sequences have intrinsic properties that are inherent in these objects, such as their chemical constitution, size, structure, and internal organization, viruses in addition possess characteristic relational properties that arise because the virus necessarily interacts with biological partners such as vectors, hosts and immune systems. These relational properties, which take the form of transmission vectors, host ranges, disease symptoms, pathogenicity and immunoreactivity, are crucial for demarcating species although they remain unknown if only the viral genome is available and the relational partners of the virus have not been identified. With very few exceptions, this is presently the case for the vast majority of viral sequences found in metagenomic databases. Since relational species-defining properties only emerge when viral proteins interact with partners in the form of vector and host gene products, these properties cannot be inferred or predicted from the viral nucleotide sequence on its own (see section 7). This is the main reason why metagenomic viral sequences cannot be assigned to individual species taxa in the current ICTV classification.

Despite the above, it must be emphasized that an inability to allocate a sequence present in a metagenomic set of data to a particular virus species does not mean that metagenomic sequences cannot be allocated to higher taxa such as ICTV families or orders. Indeed, a viral genome can be unequivocally identified as belonging to a member of an established virus family because certain conserved structural features of virions and of replication mechanisms characteristic of a family tend to correlate with particular nucleotide sequences found in viral genomes [37]. Using the composite generalized Jaccard (CGJ) index together with the GRAViTY (Genome Relationships Applied to Virus Taxonomy) analytical method for evaluating whether currently assigned ICTV families of eukaryotic viruses were consistent with their genome sequences, Aiewsakun and Simmonds [38] demonstrated that there was a remarkable agreement between groupings of sequences and existing families. They showed that it is indeed possible in the case of eukaryotic viruses to reliably predict family memberships from genome

sequences alone. When this method was applied to the order *Caudovirales* of bacteriophage families, a better agreement was observed between genetic relatedness and subfamily assignments, indicating that the traditional morphology-based classification of prokaryotic viruses would have to be revised [38]. The claim that it is possible to partly incorporate certain metagenomic data in the current ICTV system of virus classification [39] has thus been confirmed for assigning virus genome sequences to some higher taxa on the basis of a few stable properties that correlate with particular nucleotide sequences. However, this is clearly not feasible in the case of species taxa that are defined polythetically by a variable combination of several relational, phenotypic properties instead of monophyletically by a single genetic metric.

Recommendations arising from this review

- 1) The ICTV should abandon the current rule that the names of virus species (e.g., *Measles virus*) differ only from the name of viruses (e.g., measles virus) by typography.
- 2) Non-Latinized binomial species names that have been used for more than 40 years and are based on familiar virus and genus names should be introduced. This will obviate the need to create about 5000, hard-to-memorize, Latinized species names.
- 3) Virus species are defined not by the intrinsic properties of virions and viral genomes but by the relational properties of viruses that arise by interactions with their host and vector partners. Consequently, viral phenotypes cannot be inferred from viral genome sequences found in metagenomic databases because in nearly all these cases the hosts and vectors of the putative viruses have not been identified.

Acknowledgements I am greatly indebted to Timothy Skern for his numerous suggestions that helped me to improve the first version of this review paper. I am also grateful to Florence Diemer (CNRS, Strasbourg, France) for secretarial help.

References

1. Buck RC, Hull DL (1966) The logical structure of the Linnaean hierarchy. *Syst Zool* 15:97–111
2. Gibbs AJ, Gibbs MJ (2006) A broader definition of the ‘virus species’. *Arch Virol* 151:1419–1422
3. Van Regenmortel MHV (2018) The species problem in virology. *Adv Virus Res* 100:1–18
4. 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

5. 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
6. Van Regenmortel MHV (2007) Virus species and virus identification: past and current controversies. *Infect Genet Evol* 7:133–144
7. Gibbs A (2017) ICTV online; taxonomy discussions: comments posted on May 9, 2017
8. Ball LA (2005) The universal taxonomy of viruses in theory and practice. In: Fauquet CM et al (eds) Eighth ICTV Report. Elsevier, Amsterdam, pp 3–8
9. Simmonds P, Aiewsakun P (2018) Virus classification- where do we draw the line? *Arch Virol* 163:2037–2046
10. Van Regenmortel MHV (2011) Virus species. In: Tibayrenc M (ed) Genetics and evolution of infectious diseases. Elsevier, Amsterdam, pp 3–19
11. Van Regenmortel MHV (2017) Recent developments in the definition and official names of virus species. In: Tibayrenc M (ed) Genetics and evolution of infectious diseases. Elsevier, Amsterdam, pp 1–23
12. Gibbs AJ, Armstrong J, Gibbs MJ (2004) A type of nucleotide motif that distinguishes tobamovirus species more efficiently than nucleotide signatures. *Arch Virol* 149:1941–1954
13. Ereshefsky M (2010) Darwin's solution to the species problem. *Synthese* 175:405. <https://doi.org/10.1007/s11229-009-9538-4>
14. ICTVonline, Master species list 34, rank, and ICTV Code, sections 3.3 and 3.23, October 2018
15. 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
16. Mahner M, Bunge M (1997) Foundations of biophilosophy. Springer, Berlin
17. Calisher CH (2018) Viruses do not have polythetic properties; species are polythetic classes and do not have any properties. *Arch Virol* 163:2033–2036
18. 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
19. King A (2012) Comments to Proposed Modification to Code Rule 3.21 (Defining Virus Species) ICTV Discussions. https://talk.ictvonline.org/ictv1/f/general_ictv_discussions-20/3930/comments-to-proposed-modification-to-code-rule-3-21-defining-virus-species. Accessed 28 Jan 2019
20. Van Regenmortel MHV, Ackermann H-W, Calisher CH, Dietzgen RG, Horzinek M, Keil Mahy BW, Martelli GP, Murphy FA, Pringle C, Rima BK, Skern T, Vetten H-J, Weaver SC (2013) Virus species polemics: 14 senior virologists oppose a proposed change to the ICTV definition of virus species. *Arch Virol* 158:1115–1119
21. Brown JK, Zerbini FM, Navas-Castillo J, Morlon E, Ramos-Sobrinho R, Silva JCF, Fiallo-Olivé E, Briddon RW, Fernandez-Zepeda C, Idris A, Malathi VG, Martin DP, Rivera-Bustamante R, Ueda S, Varsani A (2015) Revision of Begomovirus taxonomy based on pairwise sequence comparisons. *Arch Virol* 160:1593–1619
22. Dupré J (2001) In defence of classification. *Stud Hist Philos Biol Biomed Sci* 32:203
23. Fenner F (1976) The classification and nomenclature of viruses. Second Report of the International Committee on Taxonomy of Viruses. *Intervirology* 7:1–115
24. Matthews REF (1982) Classification and nomenclature of viruses. Fourth Report of the International Committee on Taxonomy of Viruses. *Intervirology* 17:1–200
25. Francki RIB, Fauquet CM, Knudson DL, Brown F (1991) Classification and nomenclature of viruses. Fifth Report of the ICTV. Springer, Vienna, New York
26. Francki RIB (1981) Plant virus taxonomy. In: Kurstak E (ed) Handbook of plant virus infections and comparative diagnosis. Elsevier, Amsterdam
27. Van Regenmortel MHV (2000) On the relative merits of italics, Latin and binomial nomenclature in virus taxonomy. *Arch Virol* 145:433–441
28. Matthews REF (1983) The history of viral taxonomy. In: Matthews REF (ed) A critical appraisal of viral taxonomy. CRC Press, Boca Raton, pp 1–35
29. Van Regenmortel MHV, Burke DS, Calisher CH, Dietzgen RG, Fauquet CM, Ghabrial SA, Jahrling PB, Johnson KM, Holbrook MR, Horzinek MC, Keil GM, Kuhn JH, Mahy BW, Martelli GP, Pringle C, Rybicki EP, Skern T, Tesh RB, Wahl-Jensen V, Walker PJ, Weaver SC (2010) A proposal to change existing virus species names to non-Latinized binomials. *Arch Virol* 155:1909–1919
30. Kuhn JH, Dürrwald R, Bao Y, Briese T, Carbone K, Clawson AN, de Risi JL, Garten W, Jahrling PB, Kolodziejek J, Rubbenstroth D, Schwemmler M, Stenglein M, Tomonaga K, Weissenböck H, Nowotny N (2015) Taxonomic reorganization of the family Bornaviridae. *Arch Virol* 160:621–632
31. Postler TS, Clawson AN, Amarasinghe GK, Basler CF, Bavari S et al (2017) Possibility and challenges of conversion of current virus species names to Linnaean binomials. *Syst Biol* 66:463–473
32. Van Regenmortel MHV (2017) ICTV online, Taxonomy Discussions, Comments posted on 14 February 2017. https://talk.ictvonline.org/ictv1/f/general_ictv_discussions-20/3940/comment-on-simmonds-et-al-2017-consensus-statement-viral-taxonomy-in-the-age-of-metagenomics-submitted-by-marc-van-regenmortel-14-02-2017. Accessed 28 Jan 2019
33. Van Regenmortel MHV (2016) Only viruses, but not their genome sequences, can be classified into hierarchical species and genus classes. *Curr Top Virol* 13:59–64
34. Simmonds P (2015) Methods for virus classification and the challenge of incorporating metagenomic sequence data. *J Gen Virol* 96:1193–1206
35. Koonin EV, Dolja VV (2013) A virocentric perspective on the evolution of life. *Curr Opin Virol* 3:546–557
36. Brito AF, Braconi CT, Weidmann M, Dichler M, Alves JM, Gruber A, Zanotto PM (2015) The Pangenome of the Anticarsia gemmatalis multiple nucleopolyhedrovirus (AgMNPV). *Genome Biol Evol* 8:94
37. Aiewsakun P, Simmonds P (2018) The genomic underpinnings of eukaryotic virus taxonomy: creating a sequence-based framework for family-level virus classification. *Microbiome* 6:38
38. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR et al (2017) Consensus statement: Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* 15:161–168
39. Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P (2018) Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. *J Gen Virol* 99:1331–1343

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