

Discussion

# Virus species and virus identification: Past and current controversies

M.H.V. Van Regenmortel\*

*Ecole Supérieure de Biotechnologie de Strasbourg, CNRS, Boulevard Sebastien Brandt,  
67400 ILLKIRCH Cedex 2, France*

Received 3 April 2006; accepted 4 April 2006  
Available online 19 May 2006

## Abstract

The basic concepts used in virus classification are analyzed. A clear distinction is drawn between viruses that are real, concrete objects studied by virologists and virus species that are man-made taxonomic constructions that exist only in the mind. Classical views regarding the nature of biological species are reviewed and the concept of species used in virology is explained. The use of pair-wise sequence comparisons between the members of a virus family for delineating species and genera is reviewed.

The difference between the process of virus identification using one or a few diagnostic properties and the process of creating virus taxa using a combination of many properties is emphasized. The names of virus species in current use are discussed as well as a binomial system that may be introduced in the future.

© 2006 Elsevier B.V. All rights reserved.

*Keywords:* Geminiviruses; Pair-wise sequence comparisons; Papillomaviruses; Picornaviruses; Polythetic class; Potyviruses; Quasi-species; Species; Virus classification; Virus nomenclature; Virus strain; Virus species

## 1. The nature of classification

Classifying objects is a human prerogative based on the capacity of the mind to conceptualize and recognize the presence of similar properties in individual objects. Properties and classes are related abstract entities: when a property is ascribed to an object, the object thereby becomes a member of a particular class defined by that property (Quine, 1987). If a virus has a positive strand RNA genome, it becomes automatically a member of the class of positive strand RNA viruses. Classifying viruses consists in inventing taxonomic classes like particular families or species and allocating individual viruses to these classes in order to achieve some order whereby similar viral agents are grouped together. It is important to realize that all taxonomic classes are abstract concepts, i.e. constructions fabricated by the mind and not real entities located in space and time. The failure to distinguish between real objects such as organisms and viruses and the mental constructions and abstractions needed to build up any classification system has been a fertile source of confusion in taxonomy (Van Regenmortel, 2003).

## 2. The nature of viruses

Viruses have been defined as molecular genetic parasites that use cellular systems for their own replication (Villarreal, 2005). Viruses are considered to be biological entities because they possess some of the properties of living systems like having a genome and being able to adapt to certain hosts and biological habitats. However, this does not mean that viruses should be regarded as microorganisms. Viruses do not possess some of the essential attributes of living organisms such as the ability to capture and store free energy and they lack the characteristic autonomy that arises from the presence of integrated, metabolic activities. A virus becomes part of a living system only after its genome has been integrated in an infected host cell. Most biologists accept that the simplest system that can be said to be alive is a cell (Mahner and Bunge, 1997). Only unicellular and multicellular organisms possess the property of being alive, while the organelles, macromolecules and genes found in cells are not themselves considered to be alive. The difference between viruses and various types of organisms is quite obvious when the functional roles of the proteins found in viruses and in organisms are compared. When

\* Tel.: +33 3 90 24 48 12; fax: +33 3 90 24 48 11.

E-mail address: [vanregen@esbs.u-strasbg.fr](mailto:vanregen@esbs.u-strasbg.fr).

proteins are divided in three functional categories corresponding to energy utilization, carriers of information and mediators of communication, the proportion of each protein class found in viruses is markedly different from that found in living organisms (Tamames et al., 1996; Patthy, 1999). Viruses have the highest proportion of proteins involved in information processes related to the control and expression of genetic information but have very few proteins of the energy and communication classes (Fig. 1). This distribution arises because viruses utilize the metabolic machinery of the host cell and rely entirely on the energy supply system of their hosts. In contrast, bacteria have the highest proportion of proteins of the energy class involved in small molecule transformations while animals have a high proportion of proteins involved in intra and intercellular communication (Patthy, 1999).

Some authors regard viruses as living microorganisms on the basis that they share with certain parasitic organisms the property of being obligate parasites. However, the dependency of viral genes on their cellular hosts is a totally different type of parasitism from the dependency shown, for instance, by bacteria that colonize the gut of certain animals. Obligate parasitism on its own is not a sufficient criterion for establishing that an entity is alive (Van Regenmortel, 2003).

Another important distinction is that between viruses and virus particles or virions. Only virions can be fully described by their intrinsic chemical and structural properties such as their mass, size, chemical composition, genome properties, etc. Viruses on the other hand, possess in addition so-called relational or emergent properties that are actualized only during transmission, infection and replication processes. These relational properties exist only by virtue of a relation with other entities like a host or a vector and they emerge only in the system as a whole. Confusing “virus” with “virion” is similar to confusing the entity “insect” which includes several different life stages, with a single one of these stages such as a pupa, a caterpillar or a butterfly.

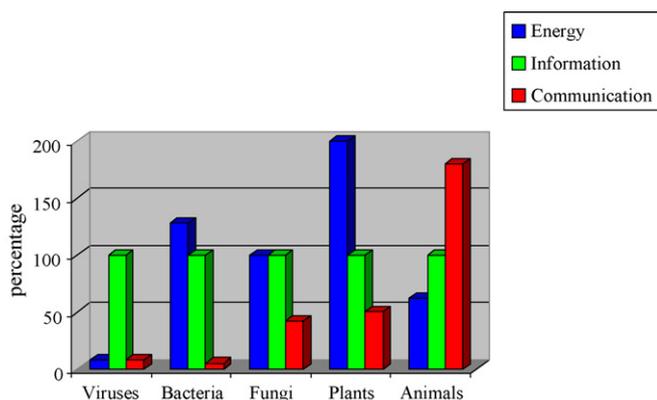


Fig. 1. The proteins of viruses and organisms have different functional roles. The vertical bars represent the proportion of proteins in the categories of energy utilization and carrier of information, relative to those in the category of mediator of information. Viruses have the highest proportion of proteins involved in information processes, i.e. in the control and expression of genetic information (from Patthy, 1999).

### 3. Classes and individuals

Distinguishing between real, tangible objects like viruses (i.e. concrete individuals) and mental constructs like virus species and genera (i.e. classes) that exist only in the mind is a basic requirement for clear thinking. Although a taxonomic class is defined by properties possessed by concrete objects, it is an abstract, conceptualized collection, i.e. a mental construct.

Classes used in biological classifications have a hierarchical structure, the main ranks being the species, genus, family, order and phylum. A class such as a particular species can only belong to one higher-rank class immediately above it, such as a particular genus and that genus in turn can only belong to one family. The logical relationship between two successive abstract classes in the hierarchy is known as class “inclusion”, whereby the lower-rank class is “included” in the higher-rank one. Class inclusion must be distinguished from the relationship of “class membership” which is the relationship used to establish a link between the abstract class and its members which are concrete individuals. Class membership is thus able to bridge two different logical categories, the abstract and the concrete (Buck and Hull, 1966; Mahner and Bunge, 1997, p. 230).

In a biological classification scheme, an individual organism or a virus can be a member of several abstract classes like a species or a genus. Class membership must be distinguished from the so-called “part-whole” relationship which exists only between two concrete objects, one being a part of the other in the way, for instance, that cells and organs are parts of an organism. It is not possible for a concrete object like a virus to be “part” of an abstract entity like a species (similarly a thought cannot be part of an object). The mixing of logical categories has led to much debate in viral taxonomy (Bos, 2003; Van Regenmortel, 2003).

A universal class, also known as an Aristotelian class, is defined by properties that are constant and immutable. This allows members of such a class to be recognized with absolute certainty since one or more property is necessarily present in every member of the class. Virus families, for instance, are universal classes because they consist of members, all of which share a number of defining properties that are both necessary and sufficient for class membership. Allocating a virus to a family is thus an easy task since a few structural or chemical attributes will suffice to allocate the virus to a particular family. For instance, all the members of the family Adenoviridae are non-enveloped viruses that have an icosahedral particle and double-stranded DNA, with projecting fibers at the vertices of the protein shell. Additional universal classes such as subfamilies, superfamilies and orders are also used in virology and they are also defined by characters that are necessarily present and allow an unambiguous allocation of the members belonging to such classes (Fauquet et al., 2005).

Unfortunately, not all properties of members of classes correspond to unambiguous predicates like the presence or absence of a DNA genome or of a particular type of particle. Many qualitative properties of concrete objects are inherently vague and do not provide precise demarcation lines. For

instance, a glass is said to be full when it is 95% or 98% full, since fullness admits of degrees and it is rarely stated that it is 100% full. In formal bivalent logic, however, such vagueness of descriptors of reality is absent and any state of affairs is presented following a mandatory dichotomy of the type: the glass is either full or not full. Such neat distinctions are rarely possible in the realm of empirical observations, which have to contend with vague terms and imprecise borderline cases (McNeill and Freiburger, 1993).

In the description of viruses, properties like the degree of genome sequence similarity between virus isolates or the nature of the symptoms induced by a virus in its host, whether mild or severe, tend to be inherently fuzzy and devoid of precise boundaries. As a result the classes that can be conceptualized on the basis of such properties will themselves be fuzzy and membership in the class will then be a matter of convention or stipulation rather than logical necessity. Virus species are precisely classes of this nature and this explains the difficulties one encounters when dealing with the species level in any biological classification. As pointed out by Schaefer and Wilson (2002) the fuzzy structure of biological populations makes it impossible for classification schemes based on mandatory all-or-nothing group membership to represent faithfully the blurred discontinuities that exist between individual groups.

A species class, the lowest rank in a biological classification, always consists of members that possess considerable variability because of the error-prone process of nucleic acid replication. This has led to the proposal that species correspond to so-called polythetic classes (Beckner, 1959). As discussed by Pigliucci (2003), a polythetic class is a cluster concept based on the concept of family resemblance described by Wittgenstein (1953). The members of a polythetic class exhibit overall similarity and have many properties in common. However, in contrast to a monothetic class, a polythetic class is not defined by a single property of the class but by a variable combination of properties (Fig. 2). The concept of polythetic class has been found useful for defining virus species in terms of fuzzy sets with hazy boundaries (Van Regenmortel, 1989, 1998).

Many biologists and philosophers have been reluctant to consider species as classes because they only recognized universal classes defined by constant properties that do not change with time. Since species change during evolution, it seemed to them that species have a spatiotemporal existence with a beginning and an end, and correspond to evolving biopopulations rather than abstract classes, immutable and timeless. According to the influential bionominalist school of thought, only concrete individuals really exist and species

should be regarded as individuals instead of abstract classes (Ghiselin, 1974; Hull, 1976). The species-as-individuals (SAI) thesis has many followers (see Ghiselin, 1997, pp. 37–61) as well as detractors (Mahner and Bunge, 1997, pp. 253–270) but in spite of its popularity, this thesis has not solved the species problem. To put it in a nutshell, the SAI interpretation amounts to the view that species are concrete entities with a definite position in space and time rather than abstract classes with an indefinite origin in time. Both schools of thought view the members of species as concrete individuals. It should be noted that bionominalists do not extend their interpretation of individuality to genera and families and they do accept that these taxa are classes. Obviously, classification schemes cannot exist without classes. The view that species are actually classes is the one that will be adhered to here.

#### 4. The nature of virus species

It is remarkable that in spite of the enormous developments of the biological sciences in the past century, there is at present no agreement on what a species actually is. The species problem has been discussed in thousands of publications (for a review see Mayr, 1982; Claridge et al., 1997) and the inability of biologists to reach agreement on the definition of species is clearly an indication that the nature of species is a philosophical problem (Pigliucci, 2003). We are able to conceptualize about the units of biological diversity in many different ways and this is illustrated by the fact that no less than 22 different species concepts have been recorded in different branches of biology (Mayden, 1997). The multiplicity of species definitions has given rise to a pluralistic approach (Mischler and Donoghue, 1982; Dupré, 1993; Hull, 1997; Ereshefsky, 1998) which was described by Kitcher (1984) as follows.

“Species are sets of organisms related to one another by complicated, biologically interesting relations. There are many such relations, which could be used to delimit species taxa. However, there is no unique relation which is privileged in that the species taxa it generates will answer to the needs of all biologists and will be applicable to all groups of organisms.”

The traditional view of species is that they correspond to groups of similar organisms that can breed among themselves and produce fertile offspring. The classical definition of biological species states that “species are groups of interbreeding natural population which are reproductively isolated from other such groups”. The reproductive isolation often simply reflects a geographic isolation. This definition is only applicable to organisms that reproduce sexually and it has limited value in the plant kingdom because of the high

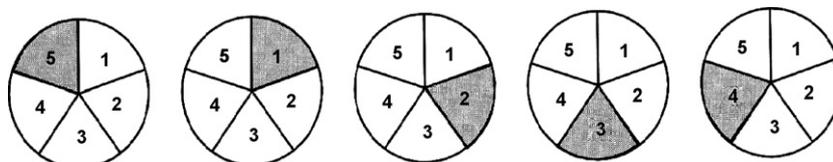


Fig. 2. Schematic representation of the distribution of defining properties (1 to 5) in five members of a polythetic class. Each member possesses several of these properties (4 out of 5) but no single defining property is present in all the members of the class. The missing property in each case is represented by the gray sector.

frequency of interspecies hybridization among plants. In order to make it applicable to asexual organisms, Mayr (1982) modified the definition of biological species as follows: “a species is a reproductive community of populations, reproductively isolated from others, that occupies a specific niche in nature”.

Some authors (Ghiselin, 1997, p. 305) reject the view that asexual organisms can form biological species but most biologists do not agree since it would render the species concept inapplicable to a considerable portion of the biological realm.

In the past, many virologists objected to the view that there could be virus species because they assumed that the only legitimate species concept was that of biological species defined by sexual reproduction, gene pools and reproductive isolation. As recounted in detail elsewhere (Van Regenmortel, 1989), the battle against the use of the species category in virology was very much an argument over semantics. Plant virologists who were most vocal against the introduction of virus species had been engaged for many years in the delineation of many separate plant viruses, but instead of calling them separate species, they referred to them as individual “viruses”. The inverted commas were meant to indicate that these entities corresponded to a separate taxon and not merely to a particular viral strain or isolate. As pointed out by Matthews (1985) these “viruses” really represented *de facto* species and plant virologists were in fact ahead of vertebrate virologists in delineating virus species (Van Regenmortel, 1989).

The absence of a satisfactory definition of virus species was also a reason why some virologists were reluctant to use the concept. Various definitions of virus species had been proposed over the years but none gained general acceptance. One definition stated that “virus species are strains whose properties are so similar that there seems little value in giving them separate names” (Gibbs and Harrison, 1976) while another stated: “A virus species is a concept that will normally be represented by a cluster of strains from a variety of sources, or a population of strains from a particular source, which have in common a set of stable properties that separate the cluster from other clusters of strains” (Matthews, 1982). Such definitions were not very helpful since they simply replaced undefined species by undefined strains and suggested that attributing names to viruses was the same activity as constructing a taxonomy.

Another definition proposed by Kingsbury (1985): “a virus species is a population of viruses sharing a pool of genes that is normally maintained distinct from the gene pools of other viruses”, was also unsatisfactory because many viruses reproduce entirely by clonal means and do not possess gene pools.

In 1991, the International Committee on Taxonomy of Viruses (ICTV) which is the body empowered by the International Union of Microbiological Societies to make decisions on matters of virus classification and nomenclature endorsed the following definition: a virus species is a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche (Van Regenmortel, 1989; Pringle,

1991). This definition means that no single property can be used as a defining property of a virus species because it is universally present in all members of the species and at the same time always absent in the members of other species. It is the inherent variability of the members of a virus species that prevents a single discriminating character such as a particular host reaction or a certain percentage of genome sequence dissimilarity to be used as a valid criterion for defining a species.

The polythetic principle has been criticized by Gibbs (2003) on the basis of an alleged contradiction between it and the concept of replicating lineage. Gibbs suggested that the ICTV was actually defining species monothetically, i.e. by properties shared by all individuals in a taxon, because it made use of genome sequence data.

However, a nucleotide motif or the entire nucleotide sequence of a virus genome is not a single, defining character and the ICTV, in fact, has continued to apply the polythetic principle when creating species taxa (Ball, 2005). This debate arises because the task of creating and defining classes and abstractions like species is confounded with the task of identifying the members of species, i.e. concrete viruses, using so-called diagnostic properties (Ghiselin, 1984; Van Regenmortel, 2000a, 2003, 2006; Van Regenmortel et al., 1997).

The Seventh and Eighth ICTV Reports (Van Regenmortel et al., 2000a; Fauquet et al., 2005) give information on which combination of properties have been used to create separate virus species within each genus. The major advantage of defining species as polythetic classes is that it makes it possible to accommodate, within a species, individual viruses that lack one or other characteristic that would normally be considered typical of the species.

The reference to replicating lineages in the definition of virus species indicates that species consist of related members that possess the replicating cohesiveness of clones. Species therefore are not simply objects with similar properties, devoid of a common origin, which can be classified phenetically. Since shared descent is a property that not only links the members of a particular species but also the members of different species and genera, using increasing genome sequence dissimilarities for demarcating species or genera is a subjective enterprise. There is indeed no precise degree of genome difference that can be used as a cut off point to differentiate between two species or two genera. The same problem exists with the evolutionary species concept advocated by some biologists (Wiley, 1978; Lovtrup, 1979) since there are no clearcut criteria for deciding how far back in time a species can be traced and became a separate species. There is no simple relationship between the extent of genome sequence similarity and the similarity in biological and phenotypic characteristics of a virus. Since the biological properties of viruses are the reason why virologists actually engage in species demarcation, it is counterproductive to try to differentiate between virus species only on the basis of genotypic or phylogenetic characteristics (Calisher et al., 1995).

The concept of ecological niche which is also included in the definition of virus species refers to biological properties such as host range, tissue tropism in the host and type of vector. The

niche is not a property of the environment but a property of the virus related to its habitat. In the absence of the virus, its ecological niche property is also absent and the notion of a vacant niche is thus meaningless (Colwell, 1992). A niche provides the needs that must be met for the virus to replicate and survive.

It is a common misconception that once the concept of virus species as a taxonomic category had been defined, it would be easy for virologists to decide if a particular virus isolate was a member of a certain species or not. The reason why this is not the case is that the definition applies only to the abstract concept of species as a category of classes and not to any particular species. Confusing the concept of species as a category of classes used in taxonomy with the concept of species as a class of similar viruses amounts to confusing the class of all chemical elements (a class of classes) with the class corresponding to single element such as gold defined by its atomic number 79. The element, of course, should also not be confused with a bar of gold!

In the genus *Lentivirus* of the Retroviridae, there are three species, *Human immunodeficiency virus 1* (HIV-1) *Human immunodeficiency virus 2* (HIV-2) and *Simian immunodeficiency virus*. HIV-1 and HIV-2 are further subdivided in groups which are themselves further subdivided in subtypes or clades representing different lineages and having some geographic associations (Damond et al., 2004). The propensity of HIV-1 strains to recombine has given rise to recombinant lineages, known as circulating recombinant forms which play an important role in the HIV-1 pandemic but greatly complicate the phylogenetic analysis (Robertson et al., 2000). The convention which views all widely divergent HIV-1 strains as belonging to a single species has an historical origin and is rather arbitrary since geographic and host range barriers to recombination between different clades could equally have justified the creation of additional species further down the lentivirus hierarchy.

Geographic and host range isolation is responsible for speciation in some virus families or genera but there are no general rules. Some viruses easily jump from one host to another but are limited in their geographic distribution, while others travel the world over but are constrained to a single host. In the family *Geminiviridae* the begomoviruses cluster according to geographic distribution into African and American viruses, presumably because of the inability of their insect vectors to fly over long distances. This geographical distribution is still clear-cut today despite frequent interspecies recombination events and increasing world-wide movement of infected plants (Stanley et al., 2005).

Another concept that has given rise to many misunderstandings is that of type species. It has been defined as follows: a type species is a species whose name is linked to the use of a particular genus name (Mayo et al., 2002). A type species is actually a name-bearing type corresponding to a class which itself is included in a particular genus class. A type species only typifies the use of the genus name but it is not typical in the sense of possessing the greatest number of characteristics of the genus. Confusion again arises if the type species as an abstract class is

not differentiated from the virus, which is a concrete, infectious entity. A virus species may have as one of its members a reference type virus which is kept in a culture collection and can be used for making comparisons with a newly isolated virus. Such a reference type is actually an isolate or a type strain and it should not be confused with a type species, which is an abstraction and not a virus culture (Mayo et al., 2002).

A further concept that has led to much confusion is that of quasi-species. Since RNA viruses have genomes that replicate in the absence of repair mechanisms and possess a mutation frequency per nucleotide site of  $10^{-3}$ – $10^{-5}$ , a clone of an RNA virus will always generate many thousands of mutants. The term quasi-species is used to refer to such populations of variable genomes which always include a so-called master sequence corresponding to the most fit genome sequence in a given environment (Holland et al., 1992; Domingo et al., 1995). The term quasi-species was introduced by Eigen (1987) to describe self-replicating RNA molecules which, because of mutations, do not consist of a single species of macromolecules. The term quasi-species in this chemical context was meant to indicate that the RNA molecules do not have a unique sequence. Virus species, in fact, are always quasi-species in the molecular sense and the term is most appropriately used for a population of variable genome sequences rather than a population of viruses. Because of the population structure of a quasi-species, natural selection in this case is no longer directed toward the single fittest variant but acts on the whole mutant distribution. Quasi-species theory is a complex field of population genetics and interested readers should consult Eigen (1996), Domingo and Holland (1997), Smith et al. (1997), Holmes and Moya (2002) and Domingo (2002).

Although levels of classification below the species level have no taxonomic significance and are not decided upon by the ICTV, it may be useful to clarify some of the terms virologists use to describe clusters of viral entities below the species category.

A viral strain is a biological variant of a given virus that is recognizable because it possesses some unique phenotypic characteristics that remain stable under natural conditions. Characteristics that allow strains to be recognized include (1) biological properties such as a particular disease symptom or a particular host, (2) chemical or antigenic properties, (3) the underlying genome sequence that is known to be correlated with the phenotypic uniqueness of the strain.

Strains that possess unique, stable antigenic properties are called serotypes, but such strains necessarily also possess unique structural, chemical and genome sequence properties that are correlated with the differences in antigenicity. As shown for instance in the case of the three serotypes of human poliovirus, serotypes also constitute stable replicating lineages since this is necessary for them to remain distinct over time (Drebot et al., 2002). Individual serotypes of mammalian viruses can be neutralized only by their own specific antibodies and not by antibodies specific for other serotypes. This inability of serotype-specific antibodies to cross-neutralize other serotypes is important for viruses submitted to the immunological pressure of their host but is irrelevant in the case of plant

viruses. In the latter case, the term serotype is used for any serologically distinguishable strain of a plant virus (Van Regenmortel, 1982).

If the only known difference between a “wild type” virus taken as reference and a particular variant is a difference in genome sequence, such a variant (or mutant) is not given the status of a separate strain since there is no recognizable distinct viral phenotype. For instance, the thousands of genetic variants that appear during replication of HIV in a given host are not given the label of strains.

The term “virus isolate” refers to any particular virus culture that is being studied and it may be a member of a genus or a species, a strain or an entity of unknown taxonomic status. A virus isolate is simply an instance of a particular virus and since it refers to a concrete entity and not to a class, the term has no taxonomic connotation.

### 5. Virus classification and pair-wise sequence comparisons

As thousands of sequences of viral genomes are continuously added to databases like GenBank (Bao et al., 2004) there is an increasing tendency to rely almost exclusively on genome data for virus classification. It is indeed tempting to assume that sequences may be able to provide the solution to all taxonomic questions (Calisher et al., 1995; Zanotto et al., 1996).

A particularly useful method for visualizing genome sequence similarities in the members of a virus family consists in plotting the frequency distribution of pair-wise identity percentages from all available genome sequences of viruses in the family. Such pair-wise sequence comparisons (PASC) produce multimodal distributions where the peaks are usually equated with clusters of sequences corresponding to groups of viral strains, species and genera (Van Regenmortel et al., 1997; Oberste et al., 1999). In fact, the various peaks actually represent the average degree of sequence identity between pairs of individual virus isolates that have been allocated either to different genera, to different species or to the same species.

Examples of PASC distributions are available on the website of the National Center for Biotechnology Information (NCBI) at the National Library of Medicine: <http://www.ncbi.nlm.nih.gov/sutils/pasc/viridity.cgi?textpage:main>.

Fig. 3A shows the frequency distribution of pair-wise identity percentages from nucleotide sequence comparisons of members of the species *Aphthovirus* in the family *Picornaviridae*. The two peaks correspond to serotypes and strains of foot-and-mouth disease virus. It is not known which structural constraints are responsible for the fact that antigenic differences among these serotypes are correlated with a narrow range of sequence variation.

Fig. 3B shows the PASC frequency distribution of members of different species in the genus *Enterovirus* of the family *Picornaviridae*. The peak at 60–65% nucleotide sequence identity arises from comparisons between members of different species.

Fig. 3C shows the overall PASC distribution when the nucleotide sequences of all the members of the *Picornaviridae*

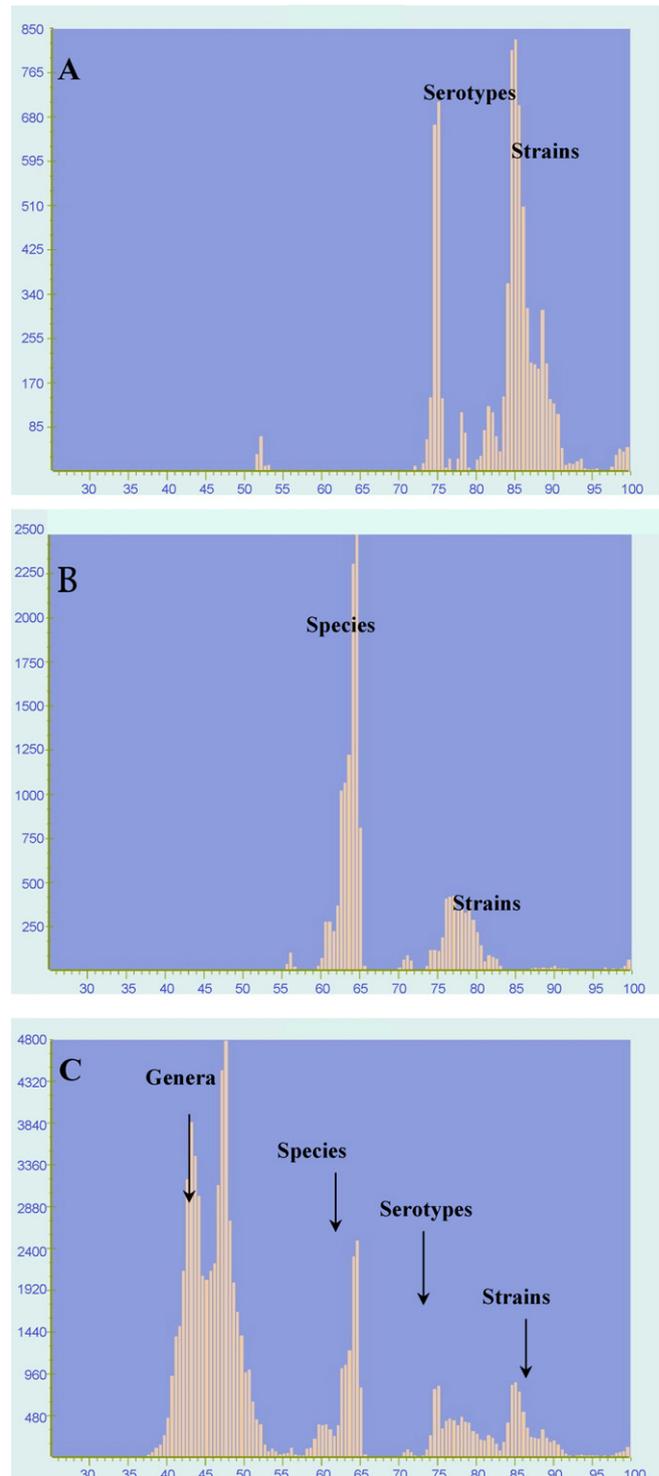


Fig. 3. (A) Frequency distribution of pair-wise identity percentages from nucleotide sequence comparisons of members of the species *Aphthovirus* in the family *Picornaviridae*. The two peaks correspond to serotypes and strains of foot-and-mouth disease virus. (B) Frequency distribution of pair-wise identity percentages from nucleotide sequence comparisons of members of different species in the genus *Enterovirus* of the family *Picornaviridae*. (C) Frequency distribution of pair-wise identity percentages from nucleotide sequence comparisons of all members of the family *Picornaviridae*. Note that in this case, there is a peak (40–53%) corresponding to comparisons between the members of different genera in the family.

are compared. The peak at 40–53% sequence identity arises from the comparisons between all the individual members of the different genera in the family. It is not at all clear which structural constraints are responsible for the fact that there are only three poliovirus serotypes but many aphthovirus serotypes and more than a hundred rhinovirus serotypes.

The classification of the large group of papillomaviruses has been in a state of flux for many years, the individual viruses being referred to as ill-defined papillomavirus “types” (De Villiers et al., 2004). However, when the PASC distribution of the sequence of the most conserved L1 gene of individual members of the family *Papillomaviridae* was plotted (Fig. 4A), two major peaks appeared. This allowed the papillomaviruses to be classified in terms of the two classical categories of species and genera. Members of the different genera share less than about 58% sequence identity whereas the members of any species within a genus share about 60–67% identity. It should be noted that there is no precise borderline value that could be used to separate the species and genus classes in terms of a certain

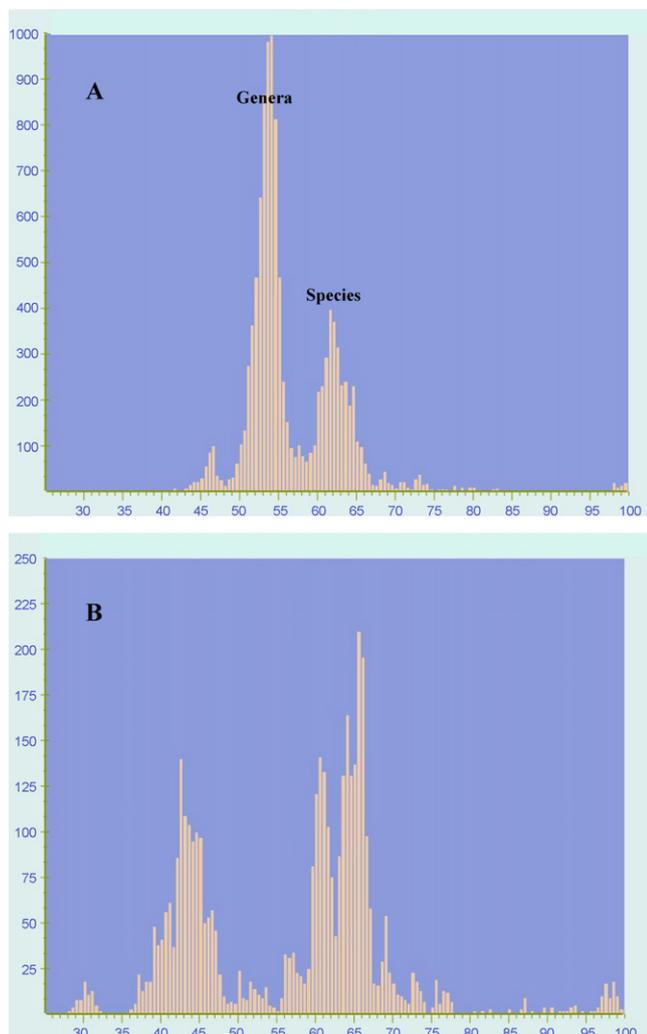


Fig. 4. (A) PASC distribution of nucleotide sequence comparisons of the L1 gene of members of the family *Papillomaviridae*. Members of the different genera share less than about 58% sequence identity. (B) PASC distribution of nucleotide sequence comparisons of members of the genus *Anellovirus*.

sequence identity percentage. There is also no precise borderline sequence identity value that separates individual species from more closely related clusters like strains or types. It is important to appreciate that pair-wise sequence comparisons of virus isolates cannot provide a definition or a reference sequence for recognizing individual species or genera.

The recently identified large group of ubiquitous TT viruses (Jelicic et al., 2004) is another interesting case. Although these viruses are presently classified as belonging all to the genus *Anellovirus* (Fauquet et al., 2005), their PASC distribution (Fig. 4B) shows two major peaks, suggesting that a classification in terms of genera (38–48% sequence identity) and species (56–70% sequence identity) is likely to be introduced in the future.

The family *Geminiviridae* is the largest family of plant viruses with four genera, 133 species and 62 tentative species listed in the Eighth ICTV Report (Fauquet et al., 2005). Geminiviruses have circular single-stranded DNA genomes in one or two components (DNA-A and DNA-B) that are packaged within twinned or geminate particles. Classifying geminiviruses is complicated because of the high frequency of recombination between different species (Padidam et al., 1999).

The genus *Begomovirus* of the *Geminiviridae* comprises 117 species whose members have been demarcated on the basis of a pair-wise sequence identity of less than 89% (Fauquet et al., 2003; Fauquet and Stanley, 2005). In the genus *Mastrevirus*, a cut-off figure of 75% sequence identity has been used for distinguishing members of different species and this higher percentage has led to the recognition of only 11 separate species. The PASC distribution of all 389 DNA-A sequences of members of the *Geminiviridae* is shown in Fig. 5. One interpretation of this distribution is to consider that the peak around 45–56% sequence identity arises from comparisons

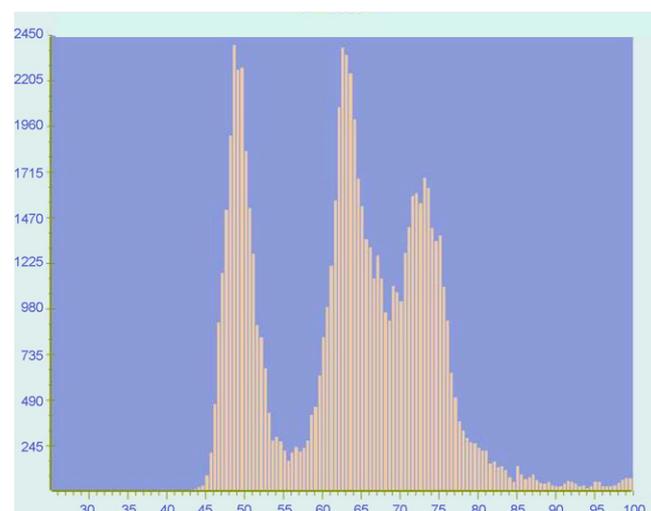


Fig. 5. PASC distribution of DNA-A nucleotide sequence comparisons of members of the *Geminiviridae*. The peak around 45–56% sequence identity arises from comparisons between members of different genera. The peaks around 56–80% sequence identity arise from comparison between members of different species. Different species of the genus *Begomovirus* possess a sequence identity of less than 89%, whereas species of the genus *Mastrevirus* possess a sequence identity of less than 75% (Fauquet and Stanley, 2005).

between members of different genera while the twin peaks around 56–80% arise from comparisons between members of different species. Other interpretations are also possible but decisions are made more difficult by the existence of recombination and lateral gene transfer among geminiviruses (C.M. Fauquet, personal communication). As is often found in PASC distributions, there are no precise borderline percentage identity values that make it possible to clearly differentiate members of the same species, of different species and of different genera in the family *Geminiviridae*.

The plant virus family *Potyviridae* contains viruses with single stranded RNA genomes encapsidated in filamentous particles about 700 nm long. Six genera and more than 120 species are currently recognized in the family (Fauquet et al., 2005). All pair-wise comparisons have been made using the nucleotide (nt) and amino acid (aa) sequences of each fully sequenced entire open reading frame and of each separate gene of every virus in the family (Adams et al., 2005). Fig. 6 shows the pairwise comparisons made between all available coat protein (CP) sequences (1220 sequences, 743590 comparisons). The comparisons between CP sequences from members of the same species gave two small peaks around 90–99% nt identity. Members of different species have less than about 76% nt identity, while members of different genera have less than about 55% nt identity (Fig. 6). Once again, the demarcating criteria used to distinguish members of different genera from members of different species and members of the same species were somewhat fuzzy. Additional data such as host vectors had to be used to justify the current classification scheme within the family *Potyviridae* (Adams et al., 2005). Although pair-wise sequence comparisons are useful for delineating clusters of viruses that are increasingly distant in genome sequence, such comparisons focus only on the

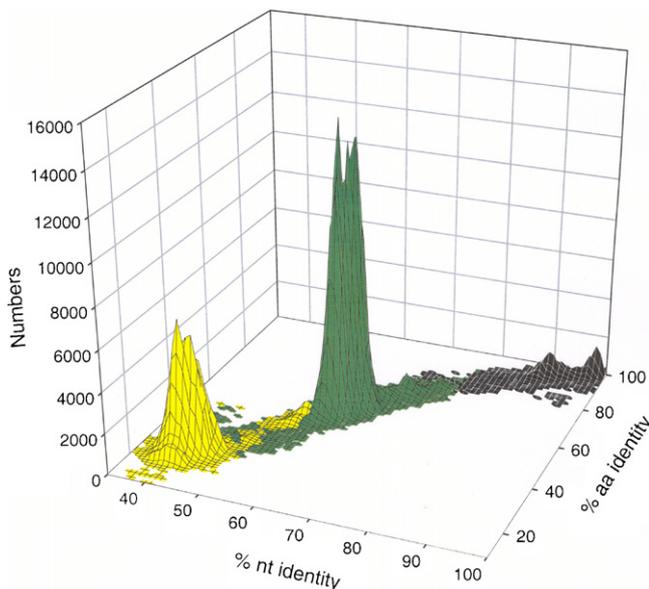


Fig. 6. PASC distributions of nucleotide (nt) and amino acid (aa) sequence comparisons of the coat protein of members of the family *Potyviridae*. Comparisons between isolates of the same species are in black, between isolates of different species in green and between isolates of different genera in yellow (from Adams et al., 2005).

properties of viral genomes and underrate the relevance of biological properties for classification purposes. Nucleotide sequences on their own cannot provide clear cut quantitative criteria for distinguishing between strains, species and genera, using similar identity percentage values in different families. Phenotypic characters of the viruses themselves must be included in a polythetic mix of defining properties in order to achieve a sensible classification that answers the needs of all interested parties involved in the study of viruses.

## 6. Virus identification and the diagnosis of virus diseases

Although virus taxonomy is sometimes downplayed as a theoretical exercise for satisfying tidy minds, the development of a sound virus classification is actually of great practical importance. A coherent and well-accepted classification and nomenclature is required, for instance, for detecting and recognizing the agents responsible for emerging epidemics, whether in humans, livestock or plants. Once the taxonomic status of a virus has been established, it can be identified and named in an unambiguous manner and this is essential for error-free communication between virologists. Furthermore, when a newly found virus is shown to be a member of a particular species or genus, many of the properties of the virus such as its pathogenesis, mode of transmission and host range can be predicted by analogy with those of other members of the taxon.

It is important not to confuse the process whereby a virologist identifies a virus isolate as a member of a species with the process whereby a taxonomist creates and defines a virus species using the polythetic principle. It is the presence of a covariant set of shared properties in most members of a species that makes it possible to predict many of the properties of a new virus once it has been identified as a member of a particular species. In order to identify a virus, use is made of one or a few diagnostic properties (Ghiselin, 1984; Van Regenmortel, 1990, 2000a). These properties, however, must not be confused with the multiple defining properties of a species. A single diagnostic tool like a monoclonal antibody specific for the members of a single species (Dekker et al., 1997) or a nucleotide combination motif (Gibbs et al., 2004) may be sufficient to distinguish the members of two closely related species such as *Tobacco mosaic virus* and *Tomato mosaic virus* in the genus *Tobamovirus*. Although such diagnostic properties make it possible to identify which viruses are members of a particular species, they do not allow an abstract class like a species to be defined (Ghiselin, 1974; Van Regenmortel, 2003, 2006; Van Regenmortel and Mahy, 2004; Van Regenmortel et al., 1997). A nucleotide motif or the reactivity of a monoclonal antibody cannot be used to define a virus species since the definition entails the presence of a number of shared properties representing a polythetic distribution. If a virus species were defined by a single character and nothing else, the identification of a virus as a member of a certain species would not be very informative. It is the failure to distinguish between the definition of an abstraction and the identification of a concrete viral entity that led Gibbs (2003) to assert that virus species can

be defined monothetically rather than polythetically. What is true, of course, is that a virus can sometimes be identified by a single diagnostic property. It cannot be overemphasized that identifying a virus as a member of a species is not the same task as providing a definition for the species class to which the virus belongs (Van Regenmortel, 2006).

To distinguish between viruses belonging to different species of a genus, one obviously cannot rely on properties that are shared by all the members of the genus like virion morphology, genome organization or method of replication. Properties that are useful for differentiating members of different species are the natural host range, cell and tissue tropism, pathogenicity, mode of transmission, certain physicochemical and antigenic properties of virions and genome sequence similarities (Van Regenmortel et al., 1997). Unfortunately, many of these properties can be altered by a few mutations and they may therefore vary in different members of the same species. This is the main reason why species demarcation is not an easy task and often requires drawing boundaries across a continuous range of genetic and phenotypic variability. Deciding whether a virus isolate corresponds to a strain of an established species or belongs to a separate species remains in many cases an elusive task that is more a matter of opinion or adjudication than logical necessity. The fuzziness and absence of sharp boundaries between species is now widely accepted (Rossello-Mora and Amann, 2001; Van Regenmortel, 1998; Schaefer and Wilson, 2002; Hanage et al., 2005) but it does not warrant abandoning the species concept. It would be equally non-sensical to refuse to distinguish colours because of the continuous nature of the spectrum of electromagnetic waves (Van Regenmortel, 1990).

Since its creation in 1966, the ICTV on behalf of the International Union of Microbiological Societies has been responsible for developing the official virus taxonomy using family, genus and species taxa. ICTV reports describing the current state of the taxonomy have appeared at regular intervals (Murphy et al., 1995; Van Regenmortel et al., 2000a,b). The eighth report which lists three orders, 73 families, 287 genera and 1950 species was published in 2005 (Fauquet et al., 2005). Another important task of the ICTV has been to develop a universal virus database, the ICTVdb, which stores descriptive data on virus isolates formatted for the DELTA computer system (Buchen-Osmond et al., 2000). The creation of this database has been a slow process. It took the ICTV nearly 10 years to adopt a standard set of descriptors for entering data in the system but the development of the database is now gaining momentum. Information on the ICTVdb is available at: <http://www.ncbi.nlm.nih.gov/ICTVdb>.

Many virologists, especially those who view the creation of taxonomic taxa as a theoretical endeavour of little practical value, regard the building of the database as the most useful activity of the ICTV. Unambiguous virus identification is based on precise descriptions of viruses at the strain and isolate levels since this will then allow newly isolated viruses to be compared with those in the database. The ICTV has been criticized in the past for not specifically linking the species described in its reports to the names and descriptions of one or more viruses that are members of these species (Gibbs, 2003). In the Eighth

ICTV report, the names of viruses belonging to each species are now mentioned and the ICTVdb which is linked to GenBank, can be consulted to obtain information on the genomic and phenotypic properties of these viruses. It must be emphasized that all the virus sequences entered in GenBank necessarily correspond to single isolates of viruses and in many cases they correspond to members of species recognized by the ICTV. On the other hand, as many as 3142 virus sequences recorded in GenBank are erroneously attributed to so-called “species” which do not exist in the ICTV master list (Fauquet et al., 2005; Fauquet and Fargette, 2005). The reason for this is that the Entrez forms which are used to enter data in GenBank automatically label each sequence as being that of a virus species rather than of an isolate. Since a virus species cannot have a sequence (Van Regenmortel, 2003), this has led to considerable confusion. Attempts are currently being made to remedy this situation (Fauquet and Fargette, 2005).

The diagnosis of viral diseases and the identification of their aetiological agents is facilitated by the availability of reference materials that are chosen by international speciality groups and are kept in culture collections such as the American Type Culture Collection in Rockville, Maryland (Murphy, 1996). However, assigning a particular virus isolate to the status of reference material has, so far, not been considered a taxonomic responsibility entrusted to the ICTV.

## 7. The names of viruses and virus species

The names of virus orders, families, subfamilies and genera are written in italics with a capital initial letter and they have the following endings: *virales* for orders, *viridae* for families, *virinae* for subfamilies and *virus* for genera. After the category species was accepted as the lowest taxonomic class to be used in viral taxonomy (Pringle, 1991; Van Regenmortel et al., 1991), I proposed to the Executive Committee of the ICTV in 1998 that two changes should be introduced in the names of virus species. The one proposal was to adopt the common English names of viruses as species names with the initial letter capitalized, and to italicize these names to provide a visible sign that species correspond to taxonomic classes, just like italicized genera and families. This proposal was accepted by the ICTV (Mayo and Horzinek, 1998; Van Regenmortel, 1999) and this italicized typography is now applied in all virology journals and books. One advantage of this typography is that official virus species recognized by the ICTV are easily differentiated from other viral entities like viral strains, types, serotypes and variants that are written in Roman.

The second proposal I made was to adopt the non-latinized binomial system which had been used unofficially by plant virologists for many years (Albouy and Devergne, 1998; Bos, 2000). In this system, the word *virus* appearing at the end of a virus name is replaced by the genus name which also ends in *virus*. This gives rise to names like *Tobacco mosaic tobamovirus* and *Measles morbillivirus* which have already been used in earlier ICTV reports (Fenner, 1976; Matthews, 1979, 1982) and have the advantage of giving additional information on the properties of the viruses (Van Regenmortel, 2000b). In 1998,

many animal virologists were opposed to such a binomial system and the ICTV Executive Committee rejected the proposal.

Some authors have criticized the ICTV for introducing the current species nomenclature and typography (Bos, 1999, 2000, 2002; Gibbs, 2000) and advocated instead that binomial Latin names should be used for virus species on the grounds that this would bring virus nomenclature in line with the Codes of Biological Nomenclature used for organisms (Agut, 2002; Bos, 1999; Gibbs, 2003). Most virologists, however, do not regard viruses as organisms and are strongly opposed to the latinization of virus names (Matthews, 1983, 1985; Milne, 1984; Van Regenmortel, 1989, 2000a,b). It is indeed, hard to imagine that virologists would have preferred the creation *de novo* of 1950 latinized species names rather than the current system based on well-known English names (Ball, 2005).

Since the major journals in virology are written in English, all virologists are conversant with the English names of viruses. When non-virologists communicate in languages other than English, they will use the common names of viruses that exist in these languages and would need to refer to italicized, species names very rarely. The use of English instead of Latin species names is hardly an impediment at a time when English has become the language of international scientific and technical communication and is more widely understood than Latin (Van Regenmortel, 2003).

One objection to the italicized typography of species names raised by Bos (1999, 2000) was that it makes it impossible to distinguish the English name of a plant like nasturtium (belonging to the species *Tropaeolum majus*) from the Latin name *Nasturtium* which is a different species altogether (i.e. water-cress). However, this is a wholly theoretical objection (Dijkstra and Khan, 2002; Van Regenmortel, 2000b; Van Regenmortel and Fauquet, 2002) since there are no species names for which such confusion could arise.

A more serious problem in the current way of naming species is that the name of a species (for instance *Measles virus*) differs only in typography from the name of the virus, i.e. measles virus. This means that authors in their articles need to be clear on whether they intend to refer to the species or to the virus, a distinction many authors find difficult to make. There is indeed a widespread tendency to confuse viruses with species (Drebot et al., 2002; Calisher and Mahy, 2003; Van Regenmortel, 2003, 2006) and to write, for instance, that the species *Cucumber mosaic virus* has been isolated from a tobacco plant, is transmitted by an aphid vector, is the causal agent of a mosaic disease and has been sequenced. All such statements are incorrect since species are abstract taxonomic constructions of the mind. Virus species do not have hosts, vectors or sequences, only viruses do. In the example above, it is the virus, cucumber mosaic virus rather than the species *Cucumber mosaic virus* that is meant. It would also have been correct to say that it was a strain or an isolate of *Cucumber mosaic virus* (rather than the species itself) that was being studied, since this would have made it clear that one was referring to a concrete individual rather than to an abstraction.

One way to facilitate the distinction between virus and species names would be to change the current monomial

names of virus species into non-latinized binomials as had been proposed to the ICTV in 1998 (Van Regenmortel, 2000b, 2001). This would make it easier to distinguish between the species *Measles morbillivirus* and the virus: measles virus, between *Mumps rubulavirus* and mumps virus, between *Hepatitis A hepatovirus* and hepatitis virus A, between *Tobacco mosaic tobamovirus* and tobacco mosaic virus, etc.

This system could be implemented without problems for about 98% of all virus species names (Van Regenmortel and Fauquet, 2002). Opponents to the binomial system tend to focus on a few exceptional cases where genera have no international genus names or the existing names do not follow ICTV rules (e.g. *Influenza virus A*). However, these few problematic cases could be resolved and they hardly justify the rejection of the binomial system as a whole.

Some authors are opposed to the binomial system because they believe that it amounts to changing all the established virus names that virologists have been using for many years. This is, of course, not so since the common or vernacular names of viruses in English or other languages will remain the same. The only change that would be introduced concerns the names of species taxa, for which no names existed in the past.

Another criticism of the binomial system is that the last word in the species name would change whenever the classification was changed and the species was allocated to another genus. An example is *Beet necrotic yellow vein furovirus*, which was moved to the genus *Benyvirus* and would become *Beet necrotic yellow vein benyvirus* (Dijkstra and Khan, 2002). However, this may well represent an advantage rather than a disadvantage since it would draw attention to the taxonomic change. Furthermore, the beginning of the species name would remain the same, which would facilitate its retrieval in an alphabetical list. Binomial names are simply a contraction of the species name in italics with the genus name in italics replacing the word *virus* in the original species name. Obviously binomial names should also be italicized.

This issue of whether non-latinized binomial names (NLBNs) should become the official species names of viruses has been debated within the ICTV for many years (Van Regenmortel, 2000a,b, 2001). In 1998, there was a large majority of ICTV Executive Committee members opposed to NLBNs, but in 2004, 50% of the Executive Committee members were in favour of the system. The ICTV has always found it difficult to poll the representative opinion of virologists worldwide because very few virologists bother to express their views on matters of taxonomy (Matthews, 1983; Van Regenmortel et al., 2000b). Even within the ICTV, only a minority of the 82 Study Groups responded when asked about their opinion on NLBNs (Ball and Mayo, 2004). In order to overcome such inertia, the ICTV found it necessary in the past to ratify decisions by accepting that a no answer vote by the ICTV membership was a vote in favour.

In 2002, efforts were made to canvass the opinion of virologists regarding their acceptance of binomial names. The results of two ballots showed that a sizeable majority (80–85%) of the 250 virologists who expressed an opinion were in favour of NLBNs (Mayo, 2002; Van Regenmortel and Fauquet, 2002).

Three years later, the ICTV which clearly is moving very cautiously, has still not clarified its position regarding the possible implementation of a non-latinized binomial system of species names. At present, the ICTV policy apparently is to allow individual Study Groups to adopt such names for the viruses they handle and not to enforce the system for all virus species. It is to be hoped that the new ICTV Executive Committee established at the 13th International Congress of Virology, held in San Francisco in August 2005, will address the issue in the near future.

In a scientific paper, it is necessary to refer only once, for instance in the Materials and methods section, to the taxonomic species name written in italics while in the remainder of the text the virus can be referred to by its common name written in Roman. Since it is the common names that are used repeatedly in the text, they are the ones that require acronyms rather than the names of virus species (Van Regenmortel and Mahy, 2004). Since the names of viruses are mostly the same as the names of the species they belong to (except for the typography) there is sometimes confusion whether the abbreviations refer to the virus or to the species. The first lists of recommended abbreviations that were published referred to the names of viruses (Hull et al., 1991; Fauquet and Martelli, 1995) but, unfortunately, subsequent lists referred to the names of species (Fauquet and Pringle, 1999). The acronyms should in fact be used only for the names of viruses and they will not be affected if binomial names of virus species were introduced.

## References

- Adams, M.J., Antoniw, J.F., Fauquet, C.M., 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Arch. Virol.* 150, 459–479.
- Agut, H., 2002. Back to Latin and tradition: a proposal for an official nomenclature of virus species. *Arch. Virol.* 147, 1465–1470.
- Albouy, J., Devergne, J.C., 1998. *Maladies à virus des plantes ornementales*. Editions INRA, Paris.
- Ball, L.A., 2005. The universal taxonomy of viruses in theory and practice. In: Fauquet, C.M., et al. (Eds.), *Eighth ICTV Report*. Elsevier, pp. 11–16.
- Ball, L.A., Mayo, M.A., 2004. Report from the 33rd Meeting of the ICTV executive committee. *Arch. Virol.* 149, 1259–1263.
- Bao, Y., Federhen, S., Leipe, D., Pham, V., Resenchuk, S., Rozanov, M., Tatusov, R., Tatusova, T., 2004. Guest commentary. National Center for Biotechnology Information Viral Genomes Project. *J. Virol.* 79, 7291–7298.
- Beckner, M., 1959. *The Biological Way of Thought*. Columbia University Press, New York.
- Bos, L., 1999. The naming of viruses: an urgent call to order. *Arch. Virol.* 144, 631–636.
- Bos, L., 2000. Structure and typography of virus names. *Arch. Virol.* 145, 429–432.
- Bos, L., 2002. International naming of viruses a digest of recent developments. *Arch. Virol.* 147, 1471–1477.
- Bos, L., 2003. Virus nomenclature; continuing topicality. *Arch. Virol.* 148, 1235–1246.
- Buchen-Osmond, C., Blaine, L. and Horzinek, M., 2000. The universal virus database of ICTV (ICTVdB). In: Van Regenmortel M.H.V., Fauquet, C.M., Bishop, D.H.L., et al. (Eds.), *Seventh ICTV Report*. San Diego Academic Press pp. 19–24.
- Buck, R.C., Hull, D.L., 1966. The logical structure of the Linnaean hierarchy. *Syst. Zool.* 15, 97–111.
- Calisher, C.H., Mahy, B.W.J., 2003. Taxonomy: get it right or leave it alone. *Am. J. Trop. Med. Hyg.* 68, 505–506.
- Calisher, C.H., Horzinek, M.C., Mayo, M.A., Ackermann, H.W., Maniloff, J., 1995. Sequence analyses and a unifying system of virus taxonomy: consensus via consent. *Arch. Virol.* 140, 2093–2099.
- Claridge, M.F., Dawah, H.A., Wilson, M.R. (Eds.), 1997. *Species: The Units of Biodiversity*. Chapman and Hall, London.
- Colwell, R.K., 1992. Niche: a bifurcation in the conceptual lineage of the term. In: Keller, E.F., Lloyd, E.A. (Eds.), *Keywords in Evolutionary Biology*. Harvard University Press, Cambridge, Mass, pp. 241–248.
- Damond, F., Worobey, M., Campa, P., et al., 2004. Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. *Aids Res. Hum. Retrovir.* 20, 66–672.
- Dekker, E.L., Dore, I., Porta, C., Van Regenmortel, M.H.V., 1997. Conformational specificity of monoclonal antibodies used in the diagnosis of tomato mosaic virus. *Arch. Virol.* 95, 191–203.
- De Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.-U., Zur Hausen, H., 2004. Classification of papillomaviruses. *Virology* 324, 17–27.
- Dijkstra, J., Khan, J.A., 2002. The current naming of plant viruses: a critical appraisal. *Arch. Virol.* 147, 2251–2253.
- Domingo, E., 2002. Quasispecies theory in virology. *J. Virol.* 76, 463–465.
- Domingo, E., Holland, J.J., 1997. RNA virus mutations and fitness for survival. *Ann. Rev. Microbiol.* 51, 151–178.
- Domingo, E., Holland, J.J., Biebricher, C., Eigen, M., 1995. Quasispecies: the concept and the word. In: Gibbs, A.J., Calisher, C.H., Garcia-Arenal, F. (Eds.), *Molecular Basis of Virus Evolution*. Cambridge University Press, Cambridge, p. 171.
- Drebot, M.A., Henchal, E., Hjelle, B., LeDuc, J.W., Repik, P.M., Roehrig, J.T., Schmaljohn, C.S., Shope, R.E., Tesch, R.B., Weaver, S.C., Calisher, C.H., 2002. Improved clarity of meaning from the use of both formal species names and common (vernacular) virus names in virological literature. *Arch. Virol.* 147, 2465–2471.
- Dupré, J., 1993. *The Disorder of Things. Metaphysical Foundations of the Disunity of Science*. Harvard Univ. Press, Cambridge, MA.
- Eigen, M., 1987. New concepts for dealing with the evolution of nucleic acids. *Cold Spring Harbor Symp. Quant. Biol.* 52, 307–319.
- Eigen, M., 1996. On the nature of viral quasispecies. *Trends Microbiol.* 4, 216–218.
- Ereshefsky, M., 1998. Eliminative pluralism. In: Hull, D.L., Ruse, M. (Eds.), *The Philosophy of Biology*. University press, Oxford, pp. 348–368.
- Fauquet, C.M., Fargette, D., 2005. International Committee on Taxonomy of Viruses and the 3, 142 unassigned species. *Virology Journal*. 2, 64–73.
- Fauquet, C.M., Martelli, G.P., 1995. Updated ICTV list of names and abbreviations of viruses, viroids and satellites infecting plants. *Arch. Virol.* 140, 393–413.
- Fauquet, C.M., Pringle, C.R., 1999. Abbreviations for vertebrate virus species names. *Arch. Virol.* 144, 1865–1880.
- Fauquet, C.M., Stanley, J., 2005. Revising the way we conceive and name viruses below the species level: a review of geminivirus taxonomy calls for new standardized isolate descriptors. *Arch. Virol.* 150, 2151–2179.
- Fauquet, C.M., Bisaro, D.M., Bridson, R.W., Brown, J.K., Harrison, B.D., Rybicki, E.P., Stenger, D.C., Stanley, J., 2003. Revision of taxonomic criteria for species demarcation in the family *Geminiviridae*, and an updated list of begomovirus species. *Arch. Virol.* 148, 405–421.
- Fauquet C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A., 2005. *Virus Taxonomy. Classification and Nomenclature of Viruses*. Eighth Report of the ICTV, Elsevier, San Diego.
- Fenner, F., 1976. *The Classification and Nomenclature of Viruses*. Second Report of the International Committee on Taxonomy of Viruses. *Intervirology* 7, 1–115.
- Ghiselin, M.T., 1974. A radical solution to the species problem. *Syst. Zool.* 23, 536–544.
- Ghiselin, M.T., 1984. “Definition” “character” and other equivocal terms. *Syst. Zool.* 33, 104–110.
- Ghiselin, M.T., 1997. *Metaphysics and the Origin of Species*. New York State University Press, New York.
- Gibbs, A., 2000. Virus nomenclature descending into chaos. *Arch. Virol.* 145, 1505–1507.
- Gibbs, A.J., 2003. Virus nomenclature, what next? *Arch. Virol.* 148, 1645–1653.

- Gibbs, A.J., Armstrong, J.S., Gibbs, M.J., 2004. A type of nucleotide motif that distinguishes tobamovirus species more efficiently than nucleotide signatures. *Arch. Virol.* 149, 1941–1954.
- Gibbs, A., Harrison, B., 1976. *Plant Virology. The Principles.* Edward Arnold, London, pp. 292.
- Hanage, W.P., Fraser, C., Spratt, B.G., 2005. Fuzzy species among recombinogenic bacteria. *BMC Biol.* 3, 6–17.
- Holland, J.J., De la Torre, J.C., Steinhauer, D.A., 1992. RNA virus populations as quasispecies. In: Holland, J.J. (Ed.), *Genetic Diversity of RNA Viruses.* Springer-Verlag, New York.
- Holmes, E.C., Moya, A., 2002. Is the quasispecies concept relevant to RNA viruses? *J. Virol.* 76, 460–462.
- Hull, D.L., 1976. Are species really individuals? *Syst. Zool.* 25, 174–191.
- Hull, D.L., 1997. The ideal species concept and why we can not get it. In: Claridge, M.F., Dawah, H.A., Wilson, M.R. (Eds.), *Species: The Units of Biodiversity.* Chapman and Hall, London, pp. 381–424.
- Hull, R., Milne, R.G., Van Regenmortel, M.H.V., 1991. A list of proposed standard acronyms for plant viruses and viroids. *Arch. Virol.* 120, 151–164.
- Jelcic, I., Hotz-Wagenblatt, A., Hunziker, A., Zur Hausen, H., De Villiers, E.-M., 2004. Isolation of multiple TT virus genotypes from spleen biopsy tissue from a hodgkin's disease patient: genome reorganization and diversity in the hypervariable region. *J. Virology.* 78, 7498–7507.
- Kingsbury, D.W., 1985. Species classification problems in virus taxonomy. *Intervirology* 24, 62–70.
- Kitcher, P., 1984. *Species.* Philos. Sci. 51, 308–333.
- Lovtrup, S., 1979. The evolutionary species: Fact or fiction? *Syst. Zool.* 28, 386–392.
- Mahner, M., Bunge, M., 1997. *Foundations of Biophilosophy.* Springer-Verlag, Berlin.
- Matthews, R.E.F., 1979. Classification and nomenclature of viruses. Third Report of the International Committee on Taxonomy of Viruses. *Intervirology* 12, 132–296.
- Matthews, R.E.F., 1982. Classification and nomenclature of viruses. Fourth Report of the International Committee on Taxonomy of Viruses. *Intervirology* 17, 1–200.
- Matthews, R.E.F., 1983. The history of viral taxonomy. In: Matthews, R.E.F. (Ed.), *A Critical Appraisal of Viral Taxonomy.* CRC Press, Boca Raton, Florida, pp. 1–35.
- Matthews, R.E.F., 1985. Viral taxonomy for the non-virologist. *Ann. Rev. Microbiol.* 39, 451–474.
- Mayden, R.L., 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge, M.F., Dawah, H.A., Wilson, M.R. (Eds.), *Species: The Units of Biodiversity.* Chapman and Hall, London, pp. 381–424.
- Mayo, M.A., 2002. ICTV at the Paris ICV: results of the plenary session and the binomial ballot. *Arch. Virol.* 147, 2254–2260.
- Mayo, M.A., Horzinek, M.C., 1998. A revised version of the International Code of Virus Classification and Nomenclature. *Arch. Virol.* 143, 1645–1654.
- Mayo, M.A., Maniloff, J., Van Regenmortel, M.H.V., Fauquet, C.M., 2002. The type species in virus taxonomy. *Arch. Virol.* 147, 1271–1274.
- Mayr, E., 1982. *The Growth of Biological Thought; Diversity, Evolution and Inheritance.* Harvard University Press, Cambridge, MA.
- McNeill, D., Freiburger, P., 1993. *Fuzzy Logic.* Simon and Schuster, New York.
- Milne, R.G., 1984. The species problem in plant virology. *Microbiol. Sci.* 1, 113–122.
- Mischler, B.D., Donoghue, M.J., 1982. Species concepts: a case for pluralism. *Syst. Zool.* 31, 491–503.
- Murphy, F.A., 1996. Viral taxonomy. In: Fields, B.N., Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology.* 3rd ed. Lippincott-Raven Publishers, Philadelphia, pp. 15–57.
- Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., Summers, M.D. (Eds.), 1995. *Virus Taxonomy.* Sixth Report of the ICTV, Springer, Vienna.
- Oberste, M.S., Maher, K., Kilpatrick, D.R., Pallansch, M.A., 1999. Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *J. Virol.* 73, 1941–1948.
- Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* 265, 218–225.
- Patthy, L., 1999. *Protein Evolution.* Blackwell Science, Oxford.
- Pigliucci, M., 2003. Species as family resemblance concepts: the (dis-)solution of the species problem? *Bioessays* 25, 596–602.
- Pringle, C.R., 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.
- Quine, W.V., 1987. *Classes versus properties.* In: Quiddities, Belknap Press, Cambridge, MA, pp. 2–24.
- Robertson, D.L., Anderson, J.A., Bradac, J.K., et al., 2000. HIV-1 nomenclature proposal. *Science* 288, 55–57.
- Rossello-Mora, R., Amann, R., 2001. The species concept of prokaryotes. *FEMS Microbiol. Rev.* 25, 39–67.
- Schaefer, J.A., Wilson, C.C., 2002. The fuzzy structure of populations. *Can. J. Zool.* 80, 2235–2241.
- Smith, D.B., McAllister, J., Casino, C., Simmonds, P., 1997. Virus “quasispecies”: making a mountain out of a molehill? *J. Gen. Virol.* 78, 1511–1519.
- Stanley, J., Bisaro, D.M., Bridden, R.W., et al., 2005. Family: *Geminiviridae.* In: Fauquet, C.M., et al (Eds.), *Eighth ICTV Report,* Elsevier, Academic Press, San Diego, pp. 301–326.
- Tamames, J., Ouzounis, C., Sander, C., Valencia, A., 1996. Genomes with distinct function composition. *FEBS Lett.* 389, 96–101.
- Van Regenmortel, M.H.V., 1982. *Serology and Immunochemistry of Plant Viruses.* Academic Press, New York, 302 pp.
- Van Regenmortel, M.H.V., 1989. Applying the species concept to plant viruses. *Arch. Virol.* 104, 1–17.
- Van Regenmortel, M.H.V., 1990. Virus species, a much overlooked but essential concept in virus classification. *Intervirology* 31, 241–254.
- Van Regenmortel, M.H.V., 1998. From absolute to exquisite specificity. Reflections on the fuzzy nature of species, specificity and antigenic sites. *J. Immun. Meth.* 216, 37–48.
- Van Regenmortel, M.H.V., 1999. How to write the names of virus species. *Arch. Virol.* 144, 1041–1042.
- Van Regenmortel, M.H.V., 2000a. Introduction to the species concept in virus taxonomy. In: Van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., McGeogh, D.J., Maniloff, J., Mayo, M.A., Pringle, C.R., Wickner, R.B., (Eds.), *Virus Taxonomy.* Seventh ICTV Report. San Diego: Academic Press, pp. 3–16.
- Van Regenmortel, M.H.V., 2000b. On the relative merits of italics, Latin and binomial nomenclature in virus taxonomy. *Arch. Virol.* 145, 433–441.
- Van Regenmortel, M.H.V., 2001. Perspectives on binomial names of virus species. *Arch. Virol.* 146, 1637–1640.
- Van Regenmortel, M.H.V., 2003. Viruses are real, virus species are man-made taxonomic constructions. *Arch. Virol.* 148, 2481–2488.
- Van Regenmortel, M.H.V., 2006. Virologists, taxonomy and the demands of logic. *Arch. Virol.* 151, in press.
- Van Regenmortel, M.H.V., Fauquet, C.M., 2002. Only italicized species names of viruses have a taxonomic meaning. *Arch. Virol.* 147, 2247–2250.
- Van Regenmortel, M.H.V., Mahy, B.W.J., 2004. Emerging issues in virus taxonomy. *Emerg. Infect. Dis.* 10, 8–13.
- Van Regenmortel, M.H.V., Maniloff, J., Calisher, C.H., 1991. The concept of virus species. *Arch. Virol.* 120, 313–314.
- Van Regenmortel, M.H.V., Bishop, D.H.L., Fauquet, C.M., Mayo, M.A., Maniloff, J., Calisher, C.H., 1997. Guidelines to the demarcation of virus species. *Arch. Virol.* 142, 1505–1518.
- Van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., McGeogh, D.J., Maniloff, J., Mayo, M.A., Pringle, C.R., Wickner, R.B. (Eds.), 2000a. *Taxonomy of viruses.* Seventh ICTV Report. San Diego: Academic Press.
- Van Regenmortel, M.H.V., Mayo, M.A., Fauquet, C.M., Maniloff, J., 2000b. Virus nomenclature: consensus versus chaos. *Arch. Virol.* 145, 2227–2232.
- Villarreal, L.P., 2005. *Viruses and The Evolution of Life.* ASM Press, Washington DC, pp. 395.
- Wiley, E., 1978. The evolutionary species concept reconsidered. *Syst. Zool.* 27, 17–26.
- Wittgenstein, L., 1953. *Philosophical Investigations.* Macmillan, New-York.
- Zanotto, P.M., de, A., Gibbs, M.J., Gould, E.A., Holmes, E.C., 1996. A reevaluation of the higher taxonomy of viruses based on RNA polymerases. *J. Virol.* 70, 6083–6096.