Taxonomy and Classification of Viruses

CORNELIA BÜCHEN-OSMOND

76

MODERN VIRUS TAXONOMY

The number of viruses as pathogens or silent passengers of other organisms, from bacteria to the dominant mammal, is very large. As we explore new niches for life and as the sensitivity and specificity of detection techniques improve, the list of viruses expands. Today, the International Committee on Taxonomy of Viruses (ICTV) recognizes about 1,550 virus species (20), but some 30,000 virus strains and isolates are being tracked by virologists in different fields of biology. The ICTV is the "international court" of experts that rules on names and relationships of all viruses, but only to the level of species.

Virus taxonomy has been described as "a polarizing subject, based on the opinionated use of data" (12). Recognizing this intrinsic problem, ICTV took early initiatives in electronic data management (1) and has supported the development of a universal virus database, ICTVdB, since 1991. The last sentence in Melnick's chapter, "Taxonomy and Classification of Viruses," in the previous edition of this Manual (11) referred to the recent development of ICTVdB. Although originally designed as a tool for taxonomic research, ICTVdB will soon be available as a continually updated reference source for clinical and other uses. This chapter is dedicated to Melnick and takes up his cue. It briefly outlines recent developments in virus taxonomy, describes the decimal code of classification that has emerged in the construction of the database, and illustrates the utility of the database with a summary of viruses causing human diseases.

While this chapter was being prepared, ICTVdB was augmented with symptom data for viruses that infect humans and is now congruent with the World Health Organization International Code of Diseases (22) (ICD-10) (http://www.who.int/whosis/icd10/).

RECENT DEVELOPMENTS IN VIRUS TAXONOMY

Virus taxonomy is changing rapidly, with changes ranging from the trivial (use of italics for species names) to profound reorganization driven by the explosion of sequence information. The universal system of viral taxonomy now accepts Linnean-like classification at the levels of order, family, subfamily, genus, and species. The suffix "virales" identifies an order. Families are identified by the suffix "-viridae," subfamilies are identified by the suffix "-virinae," and genera are identified by the suffix "-virus." The importance of distinguishing subspecies, strains, and isolates in vaccine development, diagnostics, etc., is recognized, but these lower levels are not formally classified by ICTV. Over the last 5 years the subgeneric classification of some viruses infecting humans has been drastically reorganized (Table 1). Few genera have escaped the attention of lumpers and splitters, primarily because sequence information has helped to refine earlier classifications.

International Disease Code

In the early days of virus taxonomy, clinical and pathogenic properties, as well as ecological and transmission characteristics, were the main characteristics used to classify viruses. These remain important taxonomic parameters that can be used in the database because they have been coded in the International Code of Diseases (ICD-10), as is well illustrated by the example of viruses causing hepatitis (Table 2). If the virus infection is enteric and the primary tissue tropism of the infection targets the liver leading to hepatitis, the disease is ranked according to severity. In ICD-10 the letters A and B are reserved for certain infectious and parasitic diseases. All viral hepatitis diseases in which the primary infection is restricted to the liver tissue are found in the section B15-19. For example, acute hepatitis A has been assigned the code B15, whereas B16 is reserved for acute forms of hepatitis B. B18 is used only for chronic forms of hepatitis. The digit after the decimal point is used to classify the form and severity of the disease. B15.0 and B 16.0 are used to describe a severe form of hepatitis that has developed into a coma hepaticum; B1 5.9 and B16.9, on the other hand, refer to a less severe form of hepatitis, which is not accompanied by a coma hepaticum. If the virus infection is not enteric and the primary target tissue is not the liver, the disease is coded according to other criteria even though the infection may eventually result in a liver disease.

Criteria for Virus Classification

As electron microscopy became a more widely accessible tool, the morphology of the virus particle (Fig. 1) became a prime characteristic for virus classification. Although most viruses have been seen in the electron microscope, these

TABLE 1 Families in which there have been major changes since 1995 in the generic and subgeneric classification of viruses that infect humans

Taxonomic change	Families
From species level to serotypes/strains	*Adenoviridae, Astroviridae, Bunyaviridae, Flaviviridae, Picornaviridae, Reoviridae, Retroviridae, Togaviridae*
Few changes	*Arenaviridae, Coronaviridae, Flaviviridae, Hepadnaviridae, Orthomyxoviridae, Parvoviridae, Poxviridae*
Major reorganizations leading to creation of new families or genera	*Caliciviridae, Filoviridae, Herpesviridae, Paramyxoviridae, Retroviridae, ,Rhabdoviridae, Polyomaviridae,*[a] *Papillomaviridae*[a]

[a] Formerly Papovaviridae.

virions (infectious particles) now tend to be better known by their chemical and genomic makeup, the complex disease symptoms in their hosts, and their vectors and geo-graphic distribution. Virion stability (determined by varying pH and temperature, exposure to lipid solvents, detergents, etc.) and virion antigenicity (determined by various serological methods) also emerged as essential criteria for classification of viruses. Sequencing of the viral genome is now often done very early in identification protocols, even in developing countries. Although in most cases, sequencing places a virus in a specific taxonomic slot, virion morphology and serological test results remain important criteria for identifying an unknown virus (4). In some cases sequence analysis cannot resolve a virion to the level necessary for unequivocal clinical diagnosis and serological tests must be used to differentiate between closely related serotypes (e.g., differentiating between *Japanese encephalitis virus*, *Murray Valley encephalitis virus*, and *West Nile virus*).

Characteristics Used To Describe Viruses

Today, the primary criteria used to differentiate virus orders, families, and genera are as follows:

- the type and organization of the viral genome
- the strategy of viral replication
- the structure of the virion

The species concept in viral taxonomy is predicated on the fact that viruses are "biological entities and not simply chemicals" (19), but the concept has been particularly fuzzy. Since 1991, ICTV has accepted the definition that "a virus species is a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche." Van Regenmortel (19) lists the following characters for discriminating between virus species:

- relatedness of genome sequence
- natural host range
- cell and tissue tropism
- pathogenicity and cytopathology
- mode of transmission
- physicochemical properties of virions
- antigenic properties of viral proteins

Most databases deal with a relatively limited range of data attributes, such as nucleic acid or amino acid sequences, but the value of more complex (from a database perspective) text annotation is increasingly recognized. As is clear from the information given above, a huge range of diverse biological information has to be accommodated in ICTVdB, and all virus attributes are covered in its character list (http://ictvdb.bio2.columbia.edu/chars.htm). Only partly populated, ICTVdB already lists more than 2,000 virus descriptions (items http://ictvdb.bio2.columbia.edu/ICTVdB/) constructed from about 2,400 characters, some of which (such as host range) have up to 2,000 options for selection (Table 3). By the time available data on virus isolates and strains are entered, the number of descriptions will be close to a million.

BANKING DIVERSE DATA IN ICTVDB

Developed at the Australian National University, and now based at Columbia University, ICTVdB owes much to the support of the U.S. National Science Foundation and sponsorship by the American Type Culture Collection. ICTVdB has grown in concept and capability to become a major reference resource and research tool. It was originally constructed to facilitate the more objective analysis of systematic and evolutionary relationships, but it can be used to generate character or distance matrices and to explore relationships using all virus properties. A few of the data management issues resolved in ICTVdB are outlined here to assist user interaction with the database.

Structural Concepts in ICTVdB

The database uses the DELTA (DEscription Language for TAxonomy) system (7), developed at Division of Entomology, Commonwealth Scientific and Industrial Research Organisation, by Michael Dallwitz (6), which has now been adopted as a world standard for data exchange in taxonomy. A distinctive feature of DELTA is its capacity to store an extraordinary diversity of data and to translate these data into natural language for traditional reports and Web publication. On the input side, the capacity of DELTA to handle very large datasets one item at a time is ideally suited to a long list of virus properties (character list), often accompanied by extensive text comments and images.

Another advantage of ICTVdB is a user-friendly, online data entry capability enabling peer review of new information, ranging rom molecular properties of a virus to its geographic distribution and host range. Such diverse information, with intrinsic dependencies between genomic data, protein composition, particle structure, and infectivity, places particular demands on the flat-file system of DELTA. These have been met by building a dependency network in data specification files and by blending data from diverse sources (see below). For example, ICTVdB does not contain sequence data itself but has links to genomic and protein databanks. Conversion of ICTVdB data from

TABLE 2 The continuing importance of clinical properties, symptoms, and disease designation (as coded in ICD-10) in the taxonomy of viruses causing hepatitis

Virus name	Signs and symptoms	ICD-10 code	ICD disease description
Hepadnaviridae			
Orthohepadnavirus			
Hepatitis B virus (HBV)	Acute hepatitis which may	B16.2	Acute hepatitis B without Deltavirus
	progress to chronic hepatitis,		with hepatic coma
	liver cirrhosis, and primary		
	hepatocellular carcinoma		
		B16.9	Acute hepatitis B Without Deltavirus
			without hepatic coma
		B18.1	Chronic hepatitis B without Deltavirus
Hepatitis B virus and *Hepatitis*	Acute hepatitis which may	B16.1	Acute hepatitis B with Deltavirus
delta virus	progress to chronic hepatitis		without hepatic coma
	with superinfection of Deltavirus	B16.0	Acute hepatitis B with Deltavirus with hepatic coma
		B18.0	Chronic hepatitis B with Deltavirus
Picornaviridae	Acute hepatitis	B15.0	Hepatitis A with hepatic coma
Hepatovirus		B15.9	Hepatitis A without hepatic coma
Hepatitis A virus (HAV)			
Not assigned to a family	Acute hepatitis	B17.1	Acute hepatitis E
"Hepatitis E-like viruses"			
Hepatitis E virus (HEV)			
Flaviviridae	Acute hepatitis	B17.1	Acute hepatitis C
Hepacivirus		B18.2	Chronic viral hepatitis C
Hepatitis C virus			
GB virus A	Acute and chronic hepatitis	B17.8	Other acute viral hepatitis
GBV-A-like agents		B18.8	Other chronic viral hepatitis
GB virus *B*	Acute and chronic hepatitis	B17.8	Other acute viral hepatitis
		BB.8	Other chronic viral hepatitis
Hepatitis *GB* virus C	Acute and chronic hepatitis	B17.8	Other acute viral hepatitis
		B18.8	Other chronic viral hepatitis
Hepatitis *G* virus	Acute and chronic hepatitis	B17.8	Other acute viral hepatitis
		B18.8	Other chronic viral hepatitis
Subviral agent unassigned to a virus	Acute and chronic hepatitis	B16.0	Acute hepatitis B with Deltavirus
family			(coinfection) with hepatic coma
Deltavirus			
Hepatitis delta virus		B16.1	Acute hepatitis B with Deltavirus
			(coinfection) without hepatic coma
		B17.0	Acute Deltavirus (super) infection of
			Hepatitis B carrier
		B18	Chronic viral hepatitis B with Deltavirus
			hepatitis B carrier
		B18.0	Chronic viral hepatitis B with Deltavirus
Herpesviridae	Cytomegalovirus mononucleosis, infectious mononucleosis	B25.0 (K 77.0)	Pneumonitis
Betaherpesvirinae		B25.1 (J 17.1)	Hepatitis
Human herpesvirus 5 (HHV-5)		B25.2 (K87.1)	Pancreatitis
		B25.8	Other cytomegalovirus diseases
		B25.9	Disease, unspecified
		B27.0	Gammaherpesvirus mononucleosis
		B27.1	Mononucleosis
Flaviviridae	Hepatitis, fever	A95.0	Sylvatic yellow fever
Flavivirus		A95.1	Urban yellow fever
Yellow fever virus (YFV)		A95.9	Yellow fever, unspecified

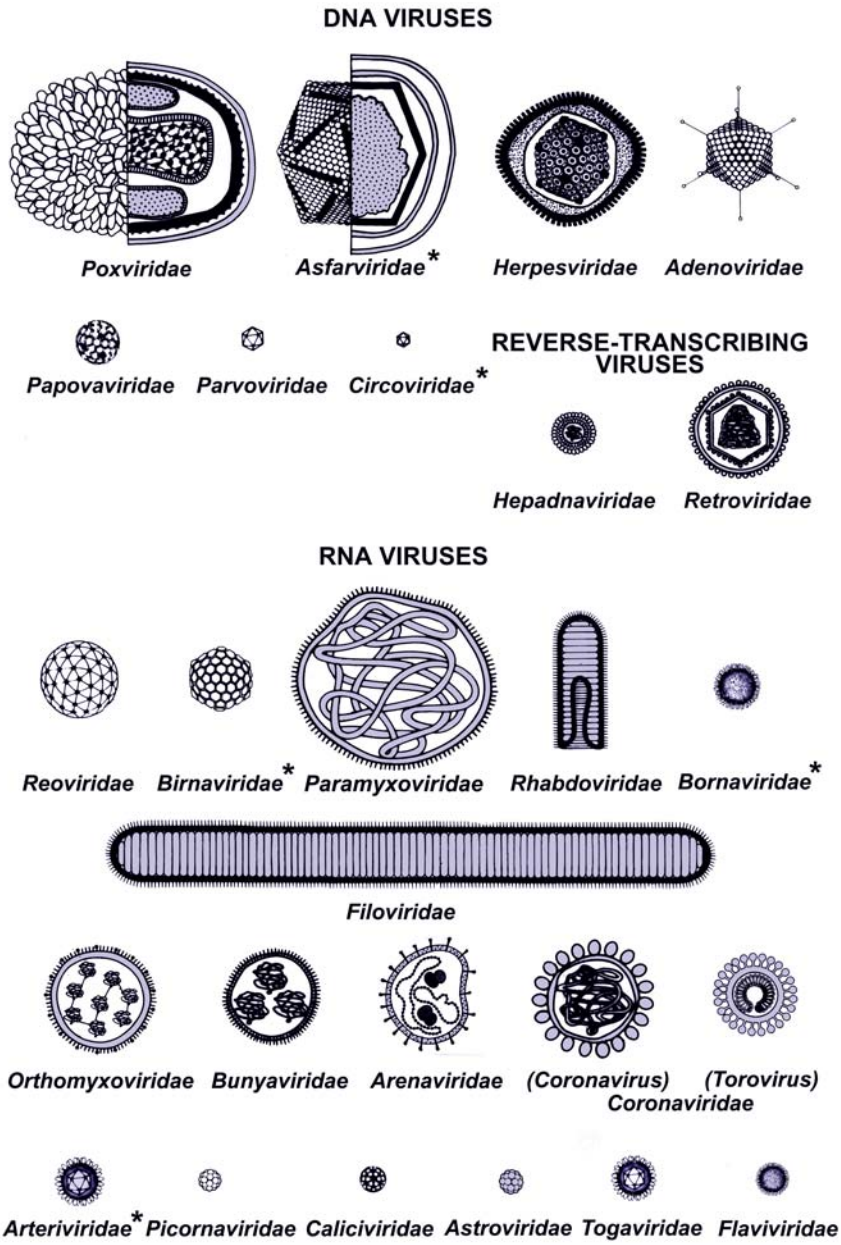

FIGURE 1 Shapes and sizes of viruses of families that include animal, zoonotic, and human pathogens. Names marked by an asteric indicate viruses only known to infect animals. The virions are drawn to scale, but artistic license has been used in representing their structure. In some, the cross-sectional structure of capsid and envelope is shown, along with a representation of the genome; with the very small virions, Only their sire and symmetry -are depicted. Viruses in families marked with an asterisk are not known to infect humans. (Reprinted from reference I2a with permission of Elsevier Science [USA]).

DELTA into NEXUS format On extensible file format for systematic information and construction of phylogenetic trees, commonly used by sequence databases; sequence data can he downloaded from all major sequence data banks in this format for tree analyses) was deemed essential for the comprehensive phylogenetic analyses that are widely usedto monitor the evolution of viruses in relation to emerging diseases. If a database is to accept the latest data from all branches of virology and place these diverse data into contemporary taxonomic context, it will most commonly deal with information at the level of strains and isolates. Ideally, ICTVdB will serve virus taxonomy "from the bottom up" with primary data from researchers who describe their viruses using rich and diverse semantics,

TABLE 3 The character list of ICTVdB [a]

Section heading	Character no. [b]
Isolate designation and definition	1-31
Classification	32-92
Virion properties	71-1135
Genome organization	1136-1594
Antigenicity	1595-1654
Biological properties	1655-2375
Taxonomic structure	2376-2390
Comments, references, and contributors	2391-2415

[a] Available at http://life.bio2.columhia/chars.htm.

[b] The characters are numbered in consecutive order of their logical appearance in the list. The character numbers are not unique identifiers but change when new characters are added to the list or the character, are reordered. In the example in the table, the designation of a virus isolate takes up 31 separate characters. The questions defined in the characters 1-31 range from the isolation date (one character each for day, month, and year) over the location of the isolation to the isolation host and person(s) and institution involved in the isolation.

reflecting geographic and linguistic factors. At the same time, the database must accept revisions and consolidations "from the top down" as the consensus in virus taxonomy reflects this new information.

Most new information in virology is generated at the molecular level and is deposited in sequence databases; however, significant events in virology tend to be associated with "host jumping," epidemics, and environmental disturbances. All of this information can be accommodated in and retrieved from ICTVdB. From the outset, DELTA was designed not only to generate identification keys but also to translate its data into natural-language hard copy, for translation onto the Web in HTML format, and for translation into many languages used around the world.

The Decimal Code in ICTVdB

The core infrastructure of ICTVdB is its distinctive "table of contents," the ICTV Index to Virus Classification (previously known as Index Virum), a list of ICTV -approved and unapproved virus names (http://life.bio2.columbia.edu/Ictv/). Constructed as a taxonomic database from classifications that do not employ normal Linnean binomial nomenclature, ICTVdB has several distinctive features that are not usually used in systematics but were introduced of necessity (2). Chief among these is its decimal code (3). Because virus names are changed frequently (8, 21), contain diverse linguistic and geographic elements, and are usually coupled to a disease or its symptoms, virus nomenclature (18) presents challenging semantic problems for a database.

The decimal code was introduced because the peculiar nomenclature used in virology defies direct and systematic interrogation in a database, and a decimal code (analogous to the code of enzyme nomenclature) seemed to offer a simple resolution to many problems. For example, at one and the same time the decimal code affords unequivocal identification of a virus to the level of strain or isolate and indicates its taxonomic context. Although the code was initially alphanumeric, new genera are now simply added numerically. Application of the code to the recently revised taxonomy of Poliovirus is illustrated in Table 4, where, for example, the emotive issue (14) of the relegation of such widely used species names as Poliovirus 1, 2, and 3 to serotypes has been justified by pairwise comparison of genomic data.

Most importantly, the decimal code serves as a file name for database outputs as well as an accession number for external linkage to ICTVdB. It is used as a file name for transposing ICTVdB to the Web and also serves as a surrogate accession number used by sequence databases such as EMBL and SWISS-PROT to link to ICTVdB. The increasing focus on such lower-level taxonomic information and more frequent taxonomic revision requires a decimal code expanded to 19 digits that should cope with even the most ambitious splitters in the taxonomic community. Many virologists are finding the decimal code useful, and it has been adopted by the Springer Index of Viruses (I6), but this invention of necessity is by no means universally accepted among virus taxonomists. Eventually, the decimal code will be used in much the same way as sequence accession numbers and will need to be assigned and cited before publication of information leading to proposals for new taxa.

TABLE 4 The decimal code applied to taxonomic revisions of Poliovirus and constructed to anticipate the explosion of lower-level data (serotypes, strains, and isolates)

Level	Extended decimal code
Order	00. = (not assigned to an order)
Family	00.05. = Picornaviridae
Subfamily	00.052.0. = (not assigned)
Genus	00.052.0.01. = Enterovirus
Serogroup	Superseded by new species concept
Species (type species)	00.052.0.01.001. = Poliovirus
Species	00.052.0.01.007. = Poliovirus
Subspecies	00.052.0.01.007.00. = (not assigned)
Serotype	00.052.0.01.007.00.001. = Poliovirus 1
	00.052.0.01.007.00.002. = Poliovirus 2
	00.052.0.01.007.00.003. = Poliovirus 3
Strain, isolate	00.052.0.01.007.00.001.001. = PV-1" Mahoney
	00.052.0.01.007.00.002.001. = PV-2 Lansing
	00.052.0.01.007.00.003.001. = PV-3 Leon/37

Operational Aspects of ICTVdB

One of the important operational considerations in ICTVdB is that at critical points in the character list, binary statements such as virus particle with or without envelope are used to establish dependencies so that only the subsequently valid characters need to be considered. These dependencies provide the internal linkage hierarchy in the data and direct the search path during interrogation; among other things, they reveal errors during data entry. The dependencies are very important for the decision-making process during identification and data comparison, and some multistate characters in key positions (e.g., plant or animal virus) can control the applicability of up to 2,000 characters down the line.

Other important operational considerations in ICTVdB include the strategies developed to facilitate communication across semantic boundaries when one is dealing with data from diverse disciplines such as bacteriology, agriculture, and veterinary and medical sciences, each of which has evolved a distinctive vocabulary. Although terms have been standardized within ICTVdB, these standards cannot be imposed on virologists in all disciplines, and they cannot be imposed retrospectively on the literature. Therefore, semantic equivalents are given where needed and subsumed within dependencies. For example, the left-hand column of Table 5 lists character #96, which handles the semantic equivalence of tegument = inner lipid protein membrane and capsid = head of a tailed phage. The right-hand columns of Table 5 show relevant dependencies, so that character 96 option 4 or 5, for example, accesses only characters 621 to 669, handling the semantic equivalence of head and capsid. The dependencies build the internal hierarchy of the database.

Images of virus particles are used in several ways in ICTVdB. For example, text descriptions of key morphological characters become much more precise when they are linked in the character list to representative vignettes from electron micrographs. Thin-section electron micrographs of infected tissues are also used to illustrate virus infection cycles and host pathology. Descriptions of all viruses generated from ICTVdB will be enhanced by electron micrographs of the type species, irrespective of the presentation format selected. Not surprisingly, images of virus particlesare among the most frequently accused files in ICTVdB on the Web. A large image file is more instructive to users, but in the database it is functionally equivalent to numerous characters, like "virus 75-50 nm in diameter" in the case of Rotavirus. File size considerations and access paths dictate that image files be stored outside the main data set, in either local files or files accessed on the Internet.

The personal computer-based ICTVdB is presented as a natural-language translation on the Web, using HTML conversion for the DELTA formatted data. This Web environment is essential for universal access, interactive data entry and interrogation, and interoperability with other databases. Currently, a plethora of accessories are available, many of which are standard components of DELTA (e.g., Web Intkey, an interactive identification program). Others, like the data entry forms, Java applets, and scripts used to display directory trees, have been developed specifically for use in ICTVdB. In the future, interoperability will be vastly improved by XML tagging. Just as it is certain that the flow of new information about viruses will not slow, it is also certain that new technology available to ICTVdB to handle these data will emerge.

CURRENT TAXONOMY OF VIRUSES INFECTING HUMANS

The most recent taxonomic decisions (2000) have been incorporated into the ICTV *Index to Classification* in ICTVdB, and the nomenclature of the 23 families of viruses that infect humans is shown in Table 6. The table summaries the important characteristics for differentiation of these viruses, listed according to the accepted taxonomic order based on the nature of the genome.

As the designation of diseases, signs, and symptoms becomes more standardized throughout the world through the use of ICD-10, it will be a relatively simple matter to assign unequivocal nomenclature and taxonomic status by blending with ICTVdB. Table *7* is a comprehensive current reconciliation of data from ICTVdB and ICD-10 (13; ICD-10-Diagnosenthesaurus, version 3.0). No doubt the taxonomy of this table will be continually revised, but it will remain updated within ICTVdB (http://life.bio2.columbia.edu/Ictv/ICD-10.htm).

TABLE 5 Semantic equivalent and dependencies among the major morphological properties of virus particles [a]

Virion or phage <components>	Dependent character blocks	
	Section heading	Character no.
1. An envelope <inner and outer envelope>/	Envelope	530-575
2. A surface membrane	Surface membrane	576-620
3. A tegument <inner Lipid protein membrane>/	Tegument	857-880
4. A head <of phage treated as isometric capsid>/	Head	621-669
5. A capsid <including outer capsid/	Capsid (coat protein)	621-669
6. Spikes <on capsid structure>/	Spikes	670-681
7. Inner capsid/	Inner capsid	682-704
8. A tail <of phage treated as elongated capsid>/	*Tail*	750-772
9. Fibers <at the tail>/	Fibers	773-784
10. A nucleocapsid/	Nucleocapsid	705-749
11. A nucleoid <term for nucleocapsid in	Nucleoid	705-749
12. A core/	Core	785-833
13. Lateral bodies/	Lateral bodies	881-904
14. A matrix/	Matrix	834-856

[a] Many viruses consist of a capsid only, while others are complex and are constructed from several components.

TABLE 6 Summary of important characteristics used to differentiate families (and one genus) of viruses infecting; humans

Classification		Virion properties						
			Virion				Genome	
Code and group	Virus (disease)	Envelope	Shape	Size (nm)	Nucleocapsid Symmetry	Nature[a]	Structure	Size (kb or kbp)
00.058. *Poxviridae*	Smallpox virus, molluscum contagiosum virus		Brick-shaped or oval	250 x 200 x 200	Complex	dsDNA, linear	Monopartite, inverted terminal repeats, ends covalently closed	130--375
00.031. *Herpesviridae*	l Herpes simplex virus (chickenpox, shingles, zoster), cytomegalovirus (mononucleosis), Epstein Barr virus (kissing disease), Kaposi's sarcoma-associated herpesvirus		Spherical	150	Icosahedral	dsDNA, linear	Monopartite, terminal and internal reiterated repeats, forming 2 covalently linked components, making 2 or 4 isomeric forms	125-235
00.001. *Adenoviridae*	Human adenovirus A to F (enteric infections, diarrhea, respiratory infections)		Isometric	70-90	Icosahedral	dsDNA, linear	Monopartite, inverted terminal repeat,, ends covalently linked protein	28-45
00.099. *Papillomaviridae*	Papillomavirus (warts)		Isometric	55	Icosahedral	dsDNA, circular	Monopartite strands circular supercoiled	6.8-8.4
00.047. *Polyomaviridae*			Isometric	40	Icosahedral	dsDNA, circular	Monopartite, strands circular supercoiled	4.7-5. 3
00.050. *Parvoviridae*	B19 virus (exanthema in children)		Isometric	25	Icosahedral	ssDNA, linear	Monopartite, population mainly (-), same (+); sequence with palindromic ends that allow circularization during replication	5
00.030. *Hepadnaviridae*	Hepatitis B virus	+	Spherical	30--34	Icosahedral	dsDNA, circular	Monopartite, circular dsDNA with regions of ssDNA	3.2
00.061. *Retroviridae*	Human immunodeficiency virus types 1 and 2	+	Spherical	80-100	Icosahedral	ssDNA, (+), linear	Monopartite, diploid, each 5' end H-bonded, 5' end capped, 3' end poly(A)	7--1 1
00.060. *Reoviridae*	Reovirus (respiratory, enteric infections), rotavirus A and B (diarrhea, enteric infections)		Isometric	60-80	Icosahedral	dsRNA, linear	10-12 segments, depending on genus	16-27
01.025. *Filoviridae*	Ebola virus, Marburg virus	+	Filamentous, pleomorphic	790-970 x 80	Helical	ssRNA, (-), linear	Monopartite	19.1
01.048. *Paramyxoviridae*	Parainfluenza virus, mumps virus, measles virus, lit man respiratory syncytial virus, hendravirus	+	Pleomorphic	150-300	Helical	ssRNA, linear	Monopartite	18-20
01.062. *Rhabdoviridae*	Rabies virus		Bullet shaped	180 x 75	Helical	ssRNA, linear	Monopartite	13-16
00.046. *Orthomyxoviridae*	Influenza virus types A-C	+	Pleomorphic		helical	ssRNA, (--), linear	6-8 segments, depending on genus	10-13.6

(Continued on next page)

TABLE 6 Summary of important characteristics used to differentiate families (and one genus) of viruses infecting humans (*Continued*)

Classification		Virion properties						
		Virion				Genome		
Code and group	Virus (disease)	Envelope	Shape	Size (nm)	Nucleocapsid symmetry	Nature [a]	Structure	Size (kb or kbp)
00.011. *Bunyaviridae*	California encephalitis virus, Lit Crosse virus, Hantaan virus, Sin Nombre virus, Crimean-Congo hemorrhagic fever virus	+	Spherical, pleomorphic	80-120	Helical	ssRNA,(--), linear	Monopartite, negative- or ambisense with sticky ends Him allow circularization during replication	11-12
00.003. *Arenaviridae*	Lassa virus, lymphocytic choriomeningitis virus, Guanarito virus, Junin virus, Machupo virus, Sabia virus	+	Spherical	110-130	Helical	ssRNA,(-), circular	Monopartite, negative- or ambisense with sticky ends that allow circularization during replication	10-14
03.019. *Coronaviridae*	Human coronavirus (respiratory and gastrointestinal infections)		Spherical, pleomorphic	80-220	Helical	ssRNA, (+), linear	Monopartite, nested set of transcription, 3' end poly(A), 5' end capped	20-30
00.052. *Picornaviridae*	Human enterovirus types A-D, poliovirus, rhinovirus types A and B, hepatitis A virus, parechovirus (human echovirus)		Isometric	28- 30	Icosahedral	ssRNA,(+), linear	Monopartite, 3' end poly(A), 5' end covalently linked protein (VPg)	7.2 8 4
00.012. *Caliciviridae*	Norwalk virus, Sapporo virus, hepatitis E virus		Isometric	35-39	Icosahedral	ssRNA, (+), linear	Monopartite, 3' end poly(A), 5' end covalently linked protein (VPg)	7.4-7.7
00.005. *Astroviridae*	Human astrovirus (gastroenteric and enteric infections)		Isometric	27__3₁1	Icosahedral	ssRNA,(+), linear	Monopartite, 3' end poly(A), 5, end not known	7-8
00.073. *Togaviridae*	Ross River virus, Chikungunya virus, O'nyong-nyong virus, rubella virus	+	Spherical	70	Icosahedral	ssRNA,(+), linear	Monopartite, 3' end poly(A), 5' end capped	9.7--11.8
00.026. *Flaviviridae*	Tick-borne encephalitis virus; dengue virus; Japanese encephalitis virus, Murray Valley virus; St. Louis encephalitis virus, West Nile virus, hepatitis C: virus, hepatitis G virus, hepatitis GB virus	+	Spherical	45-60	Icosahedral	ssRNA,(+), linear	Monopartite, Flavivirus 5' end capped, Pestivirus, Hepacivirus 5' end (?)	9.5-12.5
00.022.0.01. *Deltavirus*	Hepatitis deltavirus (aggravates hepatitis 13 virus infection)	+	Spherical	36-43	Helical	ssRNA,(-), circular	Monopartite, transcription similar to viroids; requires Hepdnaviridae for replication (similar to satellite viruses)	1.7
90.001. Prions	Creutzteldt-Jakob disease, kuru, Gerstmann-Straussler-Schenker syndrome, fatal familial insomnia	N A[b]	Rods	Protein PrP^{SC}	NA	NA	No nucleic acid, self-replicating infectious prion protein (Pro') with a molecular weight (M,) of 33,000-35,000	NA

[a] dsDNA; double-stranded DNA; ssDNA, single-stranded DNA; dsRNA, double-stranded RNA; ssRNA, single-stranded RNA; (+), positive stranded; (-), negative stranded.

[b] NA, nut applicable.

TABLE 7 Reconciliation of comprehensive, current taxonomy from ICTVdB with transmission, symptom, and disease designation from ICD-10 and important fact sheers on the web[a]

Virus code and name	Transmission	Signs and symptoms	ICD-10 and URL to important disease fact sheets
01. Mononegavirales [b]			
01.025. *Filoviridae*	Biosafety Level 4		http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/filoviruses.htm
01.025.0.01. "Marburg-like viruses"			http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/marburg.htm
01.025.0.01.001. *Marburg virus* (MARV)	Direct contact with blood or body fluids; droplet and aerosol infection may occur	Hemorrhagic fever	A98.1 Marburg virus disease
01.025.0.02. "Ebola-like viruses"	As above		http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/ebola.htm
01.025.0.02.005. *Cote d'Ivoire Ebola virus* (CIEBOV)	Direct Contact	Hemorrhagic fever	A98.4 Ebola virus disease
01.025.0.02.002. *Reston Ebola* virus (REBOV)	Direct Contact	Hemorrhagic fever	A98.4 Ebola virus disease
01.025.0.02.003. *Sudan Ebola virus* (SEBOV)	Direct contact	Hemorrhagic fever	A98.4 Ebola virus disease
01.025.0.02.004. *Zaire Ebola virus* (ZEBOV)	Direct Contact	Hemorrhagic fever	A98.4 Ebola virus disease

[a]The hull length of Table 7 can he accessed at http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/ICD-10.htm

[b]Order for viruses wick a minus-strand RNA genome, such as *Filoviridae*.

Even though most viruses affecting humans have been identified, new viruses continue to emerge or reemerge (5, 9, 15) in response to the ever-increasing mobility of the dominant mammal and its disturbance of ever more remote environments. Humankind is now an environmental factor of global proportions. As its activities increasingly contribute to global warming and climatic change, we can expect changes in the patterns of distribution of virus vectors (17) that will challenge new hosts in new environments. Although viruses are a mere twig on the evolutionary tree, they will remain the master explorers of evolutionary space (10).

Thus far, ICTVdB has been a single-investigator project, with a great deal of goodwill and support for software development. The principal impediments to its usefulness and sustainability are common to most biological databases. First, filling out of the database requires commitment from the virological community to data entry and update. It is pleasing to have some researchers deposit new virus data in ICTVdB at the same time as they deposit sequence data elsewhere. Hopefully, this will become routine, but it remains difficult to extract existing data and to engage the expertise of busy senior scientists. Second, support beyond the development phase, now largely completed, requires a shift from public funding to a commercial context. A subscription database seems a plausible path forward, one that could see a consortium of database professionals working to maintain ICTVdB. Whatever happens, it is hoped that a public-domain shop window can be retained, responsive to significant developments in virology and accessible to all parties.

REFERENCES

l. **Atherton, J. G., I. R. Holmes, and E. H. Jobbins.** 1983. *ICTV Code for the Description of Virus Characters*, vol. 14. Karger, Basel, Switzerland.

2. **Büchen-Osmond, C., L. Blaine, and M. C. Horzinek.** 2000. The universal virus database of ICTV (ICTVdB), p. 19-24. In M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. Carstens, M. K. Estes, S. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. Wickner (ed.), *Virus Taxonomy. Classification and Nomenclature of Viruses. Seventh Report of the International Committee on Taxonomy*. Academic Press, Inc., San Diego, Calif.

3. **Büchen-Osmond, C., and M. Dallwitz.** 1996. Towards a universal virus database — progress in the ICTVdB. *Arch. Virol.* **141**:392-399.

4. **Calisher, C.** 2001. Identification of arboviruses and certain rodent-borne viruses. *Emerg. Infect. Dis.* **7**:756-758. [Online. Available from: URL: http://www.cdc.gov/ncidod/eid/vol7no4/calisher.htm.]

5. **Cleaveland, S., M. K. Laurenson, and L. H. Taylor.** 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and risk of emergence. *Philos. Trans. R. Soc. London Ser. B* **356**:991-999.

6. **Dallwitz, M. J.** 1980. A general system for coding taxonomic descriptions. *Taxon* **29**:41-46.

7. **Dallwitz, M. J., T. A. Paine, and E. J. Zurcher.** 1993. *User's Guide to the DELTA System: a General System for Processing Taxonomic Descriptions*, 4th ed. CSIRO Division of Entomology, Canberra, Australia.

8. **Gibbs, A. J.** 2000. Virus nomenclature descending into chaos. *Arch. Virol.* **145**:1505-1507.

9. **Heymann, D.** 1997. Emerging infectious diseases. *World Health* **1**:4-6.

10. **Keese, P., and A. Gibbs.** 1991). Plant viruses: master explorer of evolutionary space. *Curr. Opin. Genet. Dev.* **3**:873-877.

11. **Melnick, J. L.** 1999. Taxonomy and classification of viruses, p. 835-842. In P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Yolken (ed.), *Manual of Clinical Microbiology,* 7th ed. ASM Press, Washington, D.C.

12. **Murphy, F. A.** 1995. Introduction to the universal system of virus taxonomy, p. 1-14. In F. A. MMurphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.), *The Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses*. SpringerVerlag, Vienna, Austria.

12a. **Murphy, F. A.** 1999. The nature of viruses as etiologic agents of veterinary and zoonotic diseases, p. 30. In F. A. Murphy, R. P. J. Gibbs, M. C. Horzinek, and M. J. Studdert (ed.), *Veterinary Virology*, 3rd ed. Academic Press, San Diego, Calif.

13. **National Centre for Classification in Health**. 2000. *ICD-10-AM: the International Statistical Classification of Diseases and Related Health Problems, 10th revision, Australian modification*, 2nd ed. National Centre for Classification in Health, Sydney, Australia.

14. **Pringle, C. R.** 1999. Virus taxonomy at the XIth International Congress of Virology, Sydney, Australia 1999. *Arch. Virol.* **144**:2065-2070.

15. **Taylor, L. H., S. M. Latham, and M. E. J. Woolhouse.** 2001. Risk factors for human disease emergence. *Philos. Trans. R. Soc. London Ser. B* **356**:983-989.

16. **Tidona, C. A., G. Darai, and C. Büchen-Osmond** (ed.). 2001. The Springer Index *of* Viruses. Springer-Verlag KG, Heidelberg, Germany.

17. **Tong, S., P. Bi, K. Parton, J. Hobbs, and A. J. Mc-Michael.** 1998. Climate variability and transmission of epidemic polyarthritis. *Lancet* **351**:1100.

IS. **van Regenmortel, M. H. V.** 1999. How to write the names of virus species. Arch. Virol. 144:1041-1042.

19. **van Regenmortel, M. H. V.** 2000. Introduction to the species concept in virus taxonomy, p. 3-16. In M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. Carstens, M. K. Estes, S. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. Wickner (ed.), Virus Taxonomy. Classification and Nomenclature *of* Viruses. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, Inc., San Diego, Calif.

20. **van Regenmortel, M. H. V., C. M. Fauquet, D. H. L. Bishop, E. Carstens, M. K. Estes, S. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. Wickner (ed.).** 2000. *Virus Taxonomy. Classification and Nomenclature of Viruses. Seventh Report of the International Committee on Taxonomy of Viruses.* Academic Press, Inc., San Diego, Calif.

21. **van Regenmortel, M. H. V., M. A. Mayo, C. M. Fauquet, and J. Maniloff.** 2000. Virus nomenclature: consensus versus chaos. *Arch. Virol.* **145**:2227-2232.

22. **World Health Organization.** 1999. ICD-10: *The International Statistical Classification of Diseases and Related Health Problems*, 10th ed, vol. 2001. World Health Organization, Geneva, Switzerland.