Problems of classification in the family *Paramyxoviridae*

Bert Rima¹ · Peter Collins² · Andrew Easton³ · Ron Fouchier⁴ · Gael Kurath⁵ · Robert A. Lamb⁶ · Benhur Lee⁷ · Andrea Maisner⁸ · Paul Rota⁹ · Lin-Fa Wang¹⁰

Received: 22 August 2017 / Accepted: 15 January 2018 / Published online: 25 January 2018
© Springer-Verlag GmbH Austria, part of Springer Nature 2018

Abstract

A number of unassigned viruses in the family *Paramyxoviridae* need to be classified either as a new genus or placed into one of the seven genera currently recognized in this family. Furthermore, numerous new paramyxoviruses continue to be discovered. However, attempts at classification have highlighted the difficulties that arise by applying historic criteria or criteria based on sequence alone to the classification of the viruses in this family. While the recent taxonomic change that elevated the previous subfamily *Pneumovirinae* into a separate family *Pneumoviridae* is readily justified on the basis of RNA dependent -RNA polymerase (RdRp or L protein) sequence motifs, using RdRp sequence comparisons for assignment to lower level taxa raises problems that would require an overhaul of the current criteria for assignment into genera in the family *Paramyxoviridae*. Arbitrary cut off points to delineate genera and species would have to be set if classification was based on the amino acid sequence of the RdRp alone or on pairwise analysis of sequence complementarity (PASC) of all open reading frames (ORFs). While these cut-offs cannot be made consistent with the current classification in this family, resorting to genus-level demarcation criteria with additional input from the biological context may afford a way forward. Such criteria would reflect the increasingly dynamic nature of virus taxonomy even if it would require a complete revision of the current classification.

In attempts to classify the seven viruses in the family *Paramyxoviridae*, which are not assigned to a genus, and after the decision by the International Committee on Taxonomy of Viruses (ICTV) executive committee to allow classification of viruses for which only the nucleotide sequences are available, it became clear that the currently available criteria are insufficient to unambiguously place viruses in specific

taxa in the family *Paramyxoviridae*. The seven unassigned viruses are Nariva, Mossman, Tupaia paramyxo- and Salem viruses as well as the more recently isolated beilong, J and Tailam viruses. These have all been described as morbilli-like viruses on the basis of amino acid sequence motifs of their P proteins [5].

The classical criteria used for assigning a virus to a given genus and species in the family *Paramyxoviridae* have

Handling Editor: Sead Sabanadzovic.

✉ Bert Rima
b.rima@qub.ac.uk

¹ Centre for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, The Queen's University of Belfast, Belfast, Northern Ireland, UK

² Respiratory Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

³ School of Life Sciences, University of Warwick, Coventry, UK

⁴ Department of Viroscience, Erasmus Medical Centre, Rotterdam, The Netherlands

⁵ US Geological Survey Western Fisheries Research Centre, Seattle, WA, USA

⁶ Department of Molecular Biosciences, Howard Hughes Medical Institute, Northwestern University, Evanston, IL, USA

⁷ Icahn School of Medicine at Mount Sinai, New York, NY, USA

⁸ Institute of Virology, Philipps University Marburg, Marburg, Germany

⁹ National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA

¹⁰ Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore

developed historically and were based on the properties of a virus which could be assessed by biochemical assays and SDS-PAGE before sequencing of genomes and proteins was available. One of the first criteria used to differentiate viruses in this family was the presence of neuraminidase and/or haemagglutination activity as these required no knowledge of protein or nucleic acid compositions and could be assessed with tests that preceded molecular characterisation of the viruses. These tests distinguished, for example, the current genus *Respirovirus* (members of which have both haemagglutination and neuraminidase activity) from *Morbillivirus* (in which neuraminidase activity is absent). However, haemagglutination was not always demonstrable in all of the members of, for example, the morbilliviruses. Several single reports exist in the literature which show haemagglutination

activity in morbilliviruses with very specific sets of red blood cells under very specific conditions (e.g. [9], but the frequent lack of confirmation made this criterion inconclusive. The absence or presence of the enzymatic neuraminidase activity has been a more resilient criterion that clearly and reliably distinguished the morbilliviruses from the other then known paramyxoviruses. In the past host range was also used as a strong determinant (Table 1). However, when it became clear that most viruses classified in the family have a strong potential to make host range jumps, the use of this as a phylogenetic determinant became problematic. As more recent metagenomics has shown the presence of many similar and directly related viruses in animals and humans, host range has become an even less clear criterion. Recent studies on potential reservoir hosts, especially bat species,

Table 1 Comparison of properties of unassigned viruses with those currently recognized virus species in the family *Paramyxoviridae*

Genus name	Virus names	Hosts from which isolated	PASC identity values ¹	ATUs ²	Protein expressed from P/V/C gene and size of P protein	Editing ³	C ⁴ protein
<i>Morbillivirus</i>		various mammals	53-74	none	P ~520 aa	+1G = V	C
	feline morbillivirus	cat	45	none	P 491 aa	+1G = V	C
Unassigned	Nariva virus	Zygodontomys brevicauda (rodent)	40	none	P 516 aa	+1G = V	ORF
	Mossman virus	Rattus leucopus	38	none	P 542 aa	+1G = V	ORF
	Salem virus	horse	37.8	none	V	+6-13G = P P = 356 aa	ORF
	Tupaia paramyxovirus	Tupaia belangeri (tree shrew)	36	none	P 527 aa	+1G = V	ORF
	J virus	Mus musculus	35	+SH-TM	P 496 aa	+1G = V +2G = W	ORF
	Tailam virus	Rattus andamensis	34.9	+SH-TM	P 496 aa	+1G = V +2G = W	ORF
	beilong virus	Rattus? (HKM cells)	34	+SH-TM	P 496 aa	+1G = V +2G = W	ORF
<i>Henipavirus</i>		various mammals	33	none	P 707 aa	+1G = V +2G = W Not in Cedar virus	C
<i>Respirovirus</i>		various mammals	30	none	P 568 aa	+1G = V	C
<i>Aquaparamyxovirus</i>	Atlantic salmon paramyxovirus	fish	30	none	P 575 aa	+1G = V +2G = W	C
<i>Ferlavirus</i>	Fer-de-Lance virus	snake	30	+U	V	+2G = P (429 aa)	no
<i>Rubulavirus</i>		various	25	None or +SH	V	+2G = P P 388-409 aa	no
<i>Avulavirus</i>		birds	23	none	All P (small) except type 11 = V P 393-446 aa	types 2,3,5 edited V and W; type 11 +2G=P 447(small)	no

¹PASC values of either species in comparison to all morbillivirus species or members of a genus compared to the genus morbillivirus

²ATUs = Additional transcription units in excess of the normally found N-P-M-F-{G/H(N)}-L set

³number of G residues inserted by co-transcriptional editing that gives rise to V or P or W proteins

⁴C protein is identified (C) or ORF is present in the sequence (ORF)

have indicated that there are many previously unknown paramyxoviruses [12]. When the ICTV study group attempted to apply the historically used set of criteria to classify seven unassigned species in the family *Paramyxoviridae* it became clear that using host species does not assist in reaching a rational classification (Table 1) as five of the viruses were isolated from rodents whilst the Salem virus and Tupaia paramyxovirus were isolated from a horse and tree shrew respectively.

A third important historical criterion developed in the classification of members of the *Paramyxoviridae* family was the size of the P protein and this was augmented by criteria based on the presence of an overlapping reading frame encoding a C protein and whether or not the virus genome generated P or V protein encoding mRNAs in the absence of co-transcriptional editing. Co-transcriptional editing of mRNA allows the generation of mRNAs that access alternative open reading frame (ORF) by insertion of non-templated G residues (see Table 1). The size of the P protein (smaller in the *Rubulavirus* genus than in other genera in SDS-PAGE) and the presence of an overlapping ORF encoding a C protein led to the distinction between the rubulaviruses, which encode V protein from unedited transcripts and the respiroviruses, which require editing and the consequential frame shift for the synthesis of V protein. However, application of these criteria to the classification of the seven unassigned viruses again led to inconsistencies. Annotation of the Salem virus complete genome sequence indicates that it directly encodes a V protein and has a small P protein (< 460 aa), consistent with membership of the *Rubulavirus* genus, yet many other sequence characteristics are similar to the other unassigned viruses and to members of the genus *Morbillivirus*. Reliance on the need for editing to express the V protein for classification also leads to inconsistencies if this criterion is applied to members of the genus *Avulavirus*, which themselves were distinguished from those belonging to the genus *Rubulavirus* on the basis of their avian host range. Most of the avulaviruses encode P proteins from unedited transcripts except avian paramyxovirus 11, which is proposed from sequence data to require insertion of two G residues during co-transcriptional editing to generate mRNA expressing the P protein. Thus, neither criteria on co-transcriptional editing nor the size of the P protein are applied consistently in the current classification. The avula- and rubulaviruses also stand apart from the others because of the absence of an overlapping ORF at the 5' end of the coding sequence for P/V that encodes a C protein, a feature they share with members of the *Ferlavirus* genus and Salem virus.

The genome organisation in terms of the number and order of transcription units (TUs) were applied as further criteria when more complete genome sequences became available for the members of the family (Figure 1). The

criterion of numbers of TUs was violated immediately when the rubulaviruses were not separated into two groups i.e. those with 7 TUs e.g. mumps virus and parainfluenza virus type 5 (PIV5), which encode a small hydrophobic (SH) protein and the others with 6 TUs, which lacked the SH TU and protein. Members of the genus *Ferlavirus* have an additional TU between the N and P/V genes. The unedited mRNA copied from their genome template also encodes a V rather than the P protein (as in the rubulaviruses) and the P protein itself is of an intermediate size of 429 amino acid residues in length.

Collectively the seven unassigned viruses have been described as morbillivirus-like [5]. However, the suggestion for the classification of the unassigned species in a separate genus creates a similar inconsistency already accepted within the classification of the rubulaviruses. Beilong-, J- and Tailam viruses should then be distinguished from Nariva-, Mossman- and Tupaia paramyxovirus as well as from Salem virus on the basis that their genome sequences show the presence of two extra TUs encoding an SH protein and a trans-membrane protein of as yet undetermined functions (Figure 1).

It can be seen that all the criteria that were used historically, have in one case or more been applied inconsistently in the current classification in this family. Continuing classification on these criteria is likely to lead to further inconsistencies that will undermine the classification process. The potential number of paramyxoviruses that are to be newly classified may well be very great as there is vast sequence diversity in wildlife reservoirs [12]. For many of the new viruses the only information available will be their nucleotide sequence and inferred genome organisation. It is unlikely that the historically used parameters, especially biological parameters such as the presence of neuraminidase or haemagglutination activity or receptor usage and even less likely pathogenesis, if any, will be determined for these viruses. Moreover, in many cases when "virus sequences" are derived from metagenomic studies there will not even be an infectious agent to study. As a consequence of the recent decision by the ICTV executive committee to allow classification of such sequences [11], the study groups may be required to re-examine and possibly develop new schemes for classification in their respective families.

PASC analysis

One currently applied methodology in virus classification is pairwise amino acid sequence comparison (PASC). PASC analysis is currently used at the NCBI web site as an aid in virus classification. It aligns sequences using BLAST of all the proteins in a virus to those of all other viruses in the comparison and then computes a score of identity to all the

Fig. 1 The order of the transcription units in the *Paramyxoviridae*. The boxes and intergenic, leader and trailer sequences are not to scale but only the order is indicated. Abbreviations N: nucleocapsid protein; P phosphoprotein, V V protein; C C protein; M membrane or

matrix protein, F fusion protein; SH small hydrophobic protein, tM transmembrane protein; H haemagglutinin protein; HN haemagglutinin-neuraminidase protein; G attachment protein and L large protein (RdRp)

other viruses in the family or genus [2]. However, applying it in the family *Paramyxoviridae* highlights the conundrum that must be solved.

Using comparisons between the currently recognized members of the genus *Morbillivirus* it is clear that the current classification is validated by PASC if the cut off points for species and genus are chosen carefully. The six older members within the current genus have identities between 53 and 68 in between species comparisons. A comparison of canine distemper virus (CDV) with phocine distemper (PDV) generates a small peak in the histogram with a value of 74 (Table 1). These two viruses are known to be very closely related but CDV and PDV were assigned to different species, largely based on the host from which they were isolated. However, in subsequent analysis it became clear that the host ranges of these viruses overlapped. The other outlier is the newly classified feline morbillivirus (FeMV) with a PASC value of 45. PASC analysis thus validates

the classification of the current species within the genus if a lower cut-off point of greater than 74 and less than 84 (defined by the variation between two members of the species small ruminant morbillivirus) is applied for the definition of a species and around 45 for the lowest identity value in PASC analysis for a virus to belong to the genus *Morbillivirus*.

However, difficulties arise when comparing the seven unassigned viruses with established members of the *Morbillivirus* and *Henipavirus* genera (Table 2). This clearly shows that using PASC to make decisions here does not allow the setting of clear cut-off points and indicates that the beilong, J and Tailam viruses have slightly higher PASC values in comparisons to the henipaviruses whilst Salem, Mossman, Nariva and tupaia viruses have the highest value when compared against the morbilliviruses. It is relevant here that Hendra virus was originally given the name equine morbillivirus on the basis of several

Table 2 PASC comparison of unassigned paramyxoviruses with specific species in the genera *Henipavirus* and *Morbillivirus*

Comparator 1	Acc number	Comparator 2	Highest PASC Identity value + virus species
beilong virus	DQ100461	J virus	51.3
beilong virus	DQ100461	Tailam virus	67.0
J virus	AY900001	henipaviruses	36.9 Mojiang
		morbilliviruses	35.7 measles
beilong virus	DQ100461	henipaviruses	35.9 Nipah
		morbilliviruses	34.4 PPRV
Tailam virus	JN689227	henipaviruses	35.7 Nipah
		morbilliviruses	34.8 RPV
Nariva virus	FJ362497	henipaviruses	35.4 Nipah
		morbilliviruses	40.4 CeMV
Mossman virus	AY286409	henipaviruses	34.5 Nipah
		morbilliviruses	38.5 CDV
Salem virus	JQ697837	henipaviruses	32.5 Nipah
		morbilliviruses	37.9 PDV
Tupaia paramyxovirus	AF079780	henipaviruses	35.1 Mojiang
		morbilliviruses	36.1 measles
feline morbillivirus	JQ411014	henipaviruses	32.7 Nipah
		morbilliviruses	47.2 PPRV

similarities with the morbilliviruses. The PASC comparisons between J virus, beilong and Tailam viruses range from 0.51 to 0.67 (for the beilong/J virus and beilong/Tailam virus comparison respectively). The comparison of these three viruses with the other unassigned viruses shows identity values of around 0.33. Thus, classification based on PASC analysis would require the generation of at least two new genera between the morbilliviruses and the henipaviruses: one containing Mossman, Nariva, Salem and tupaia paramyxovirus closer to the morbilliviruses and the other containing J, beilong and Tailam viruses closer to the henipaviruses. However, to sustain this classification arbitrary and different cut-off points would have to be set to distinguish them from the *Morbillivirus* genus or the *Henipavirus* genus. This underscores the need for continued input from expert study groups to set such arbitrary demarcation criteria.

PASC analysis clearly separates the rubulaviruses and avulaviruses from the morbilliviruses, but it is also clear that the PASC identity comparisons between the *Ferlavirus* and *Aquaparamyxovirus* genera show these to be closer to the respiroviruses than to the morbilliviruses and even further from the rubulaviruses and avulaviruses, with identities ranging around 30 (Table 3). Table 3 also shows that assigning the Fer-de-Lance virus and Atlantic salmon paramyxovirus to separate genera based on the host from which they were isolated and also on genome organisation, has generated inconsistencies if PASC identity values are used as a sole criterion.

Table 3 PASC comparison of members of specific genera in the *Paramyxoviridae*

Comparator 1	Comparator 2	Identity value
Atlantic salmon paramyxovirus	respiroviruses	40.8
	morbilliviruses	31
	rubulaviruses	28.7
Fer-de-Lance virus	respiroviruses	39
	morbilliviruses	32.6
	rubulaviruses	31.5
avulaviruses	rubulaviruses	37
rubulaviruses	morbilliviruses	30

Overall it seems clear that PASC analysis alone will not help in placing the unassigned viruses in what is a continuum between the ferla-, aquaparamyxo-, respiro-, morbilli- and henipaviruses. The rubula- and avulaviruses are clearly separate from this continuum, but form one between them.

RdRp motifs as classification criteria

Comparisons of amino acid sequences of the RdRps can provide criteria for the establishment of a new family in the order *Mononegavirales* (Rima, in preparation) and with the current number of viruses this still leads to the recognition of distinct groupings. Dissimilarities in the RdRp amino acid sequences were one of the reasons, together with a number

of other specific characteristics, for classifying the *Pneumoviridae* as a different family in the order *Mononegavirales* [10]. Motifs and sequences in RdRp motif analysis also demonstrate the differences in the *Paramyxoviridae* between on the one hand the rubula- and avulaviruses and on the other the 5 remaining recognized genera and the unassigned viruses. This distinction is also demonstrated clearly in a phylogenetic tree based on the comparisons of the complete amino acid sequences of the L or RdRp proteins of all 65 currently known species in the family (Figure 2) including a recently newly identified bank vole virus [1]. However, again using such trees for classification into lower taxa such as the genus and species may require setting arbitrary taxon-specific cut-off criteria.

The seven unassigned morbilli-like viruses contain all the RdRp sequence motifs that are characteristic of all the current members of the morbillivirus genus, though the lengths of the insertions in the hinge regions identified in the L protein varies for each of the viruses. Furthermore, the genome organisation differs in at least in three of them (Figure 1) as there are two additional transcription units encoding an SH and a transmembrane protein. Biological differences like these may prompt classification of J, beilong and Tailam viruses into a separate genus. However, the classification of Salem virus versus Nariva, Mossman and bank vole viruses again raises questions about where to set cut-off points; these are issues for consideration by future expert study groups.

Basing the classification of the *Paramyxoviridae* on RdRp sequences should also take account of a number of facts namely that (i) the data bases contain sequences which have internal deletions, (ii) in some cases conserved motifs are only found in alternate reading frames due to single nucleotide deletions and insertions in the sequences reported in the data banks and (iii) recent experience with deep sequencing indicates that the traditional Sanger sequencing methods in virus samples in which defective interfering (DI) particles predominate probably has established sequences of DI RNAs rather than the authentic standard virus RdRp sequence. Hence, it is paramount that well curated sequences only should be used for this analysis.

Receptor usage as a criterion for classification

A recent publication [13] suggested that receptor usage could be considered as a potential criterion for classification of viruses. Whilst this may work for viruses such as morbilliviruses and henipaviruses for which the receptors have been identified it is more difficult to apply this to viruses which use sialic acid as the primary attachment module. In their consideration of this proposal Zeltina and coworkers [13] already mention exceptions and difficulties in this approach.

The amino acid motif NRKSCS is found consistently at the start of propellor blade 2 in the haemagglutinin-neuraminidase proteins of viruses that have neuraminidase activity [7]. Table 4 shows the equivalent sequences in all the paramyxoviruses. It is clear that the cysteine residue is conserved in all the viruses and that the motif is absent in those viruses that are known to use protein receptors such as the morbilli- and henipaviruses. However, the NRKSCS motif is also absent in a number of viruses that have been classified as rubulaviruses based on the size of the P protein and sequence motifs in the RdRp. It would also be unclear without further biochemical studies whether the NRKSCS to NRRSCS mutation in viruses such as beilong, J and Tailam viruses would render the protein inactive. It is unlikely that this type of information will ever become established for many of the viruses that will require classification and which are known only on the basis of their nucleotide sequence information making it unattractive as a main criterion for classification.

Other criteria

The close relationship in terms of sequence similarity between J and beilong viruses with the henipaviruses (Table 2) led to the suggestion of a further criterion namely the ability of plasmids expressing the N, P and L proteins of one of these viruses to successfully rescue minigenome constructs of any of the other viruses. J and beilong viruses can be “cross-rescued” but were unable to rescue henipa (Nipah) virus minigenomes. In the past it has been shown that three species of morbillivirus CDV, measles virus (MV) and rinderpest virus (RPV) did “cross-rescue” [3], so this would argue to place J and beilong viruses in the same genus. The application of this criterion as well as another relating to potential superinfection immunity established by persistent or acute infection of one virus to another [8] may have application in well studied viral systems but would most probably never be realizable in the plethora of virus sequences that the ICTV study group is now asked to classify. These criteria cannot be applied because of lack of data for the “recognized but unassigned” morbillivirus-like viruses that pose the current question in classification.

Conclusions

Ideally, the classification of viruses should allow a relatively expert reader to immediately recognize a set of likely characteristics of the virus. This provides meaning to a statement that a specific virus is classified in a given taxon. The more genera and species introduced into a classification scheme the lower the immediacy of the recognition. There

Fig. 2 Molecular Phylogenetic analysis by Maximum Likelihood method of 65 members of the *Paramyxoviridae* based on the amino acids sequences of the L or RdRp protein. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [4]. The tree with the highest log likelihood (-173347.50) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT

model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 1995 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [6]. Printed in red are those viruses which are still formally to be classified (color figure online)

Table 4 Amino acid sequences at the start of propellor blade 2 in the attachment protein (H/HN/G)

Genus	Virus	Size of H(N) or G in amino acid residues	Sequence	Receptor usage
<i>Ferlavirus</i>	Fer-de-Lance virus	HN 564	NRKSCS	?
<i>Aquaparamyxovirus</i>	Atlantic salmon paramyxovirus	HN 576	NRKSCS	?
<i>Respirovirus</i>	Sendai virus	HN 575	NRKSCS	Neuram. Ac.
	Human parainfluenza virus 3	HN 572	NRKSCS	Neuram. Ac.
	Bovine parainfluenza virus 3	HN 572	NRKSCS	Neuram. Ac.
	Porcine parainfluenza 1	HN 576	NRKSCS	?
<i>Henipavirus</i>	Hendra virus	G 604	TIHHCS	Ephrin-B2
	Nipah virus	G 420	TVYHCS	Ephrin-B2
	Cedar virus	G 622	QVINCV	?
	Mojiang virus	G 625	IINSCA	?
	Kumasi virus	G 632	NYHSCT	?
<i>Morbillivirus</i>	Measles virus wt	H 617	DFSNCM	SLAM/nect
	Measles Ed-Zag vac	H 617	DLSNCM	SLAM/nect/CD46
	Canine distemper virus	H 607	KTKVCT	SLAM/nect
	Canine distemper virus vac	H 607	KAKVCT	SLAM/nect/?
	Phocine distemper virus	H 607	NTKICT	SLAM/nect
	Rinderpest virus	H 609	ELETCM	SLAM/nect
	Peste des petits ruminants virus	H 609	DYRSCL	SLAM/nect
	Porpoise morbillivirus	H 604	GLNFCL	?
	Feline morbillivirus	H 595	GMESCT	?
<i>Unassigned</i>	Nariva virus	“H” 657	AYDGCA	?
	Mossman virus	“G” 632	VFDGCS	?
	Tupaia paramyxovirus	“H” 665	NLRDCS	?
	Salem virus	“G” 620	LSGKCT	?
	Beilong virus	“G” 734	NRRSCT	?
	Tailam virus	“G” 1052	NRRSCT	?
	J virus	“G” 709	NRRSCS	?
<i>Rubulavirus</i>	Mumps virus	HN 582	NRKSCS	Neuram. ac
	Porcine rubulavirus	HN 576	NRKSCS	?
	Bat mumps virus	HN 582	NRKSCS	?
	Mapuera virus	HN 582	NRKSCS	?
	Simian virus 41	HN 568	NRKSCS	?
	Parainfluenza virus 2	HN 571	NRKSCS	?
	Parainfluenza virus 4A	HN 579	NRKSCS	?
	Parainfluenza virus 4B	HN 574	NRKSCS	?
	Parainfluenza virus 5	HN 532	NRKSCS	Neuram.ac
	Menangle virus	“HN” 595	PVRTCS	?
	Tioman virus	“HN” 593	QARGCS	?
	Teviot virus	“HN” 595	QTRGCS	?
	Achimota virus 1	“HN” 595	VTYQCS	?
	Achimota virus 2	“HN” 583	FRRGCS	?
	Tuhoko virus 1	“HN” 580	WLRSCS	?
	Tuhoko virus 2	“HN” 588	VSRQCS	?
	Tuhoko virus 3	“HN” 582	RLYHCS	?
	Sosuga virus	“HN” 582	RLYHCS	?
<i>Avulavirus</i>	Avian paramyxovirus 1 (NDV)	HN 571	NRKSCS	Neuram.ac
	Avian paramyxovirus 2	HN 580	NRKSCS	?
	Avian paramyxovirus 3	HN 577	NRKSCS	?
	Avian paramyxovirus 4	HN 569	NRKSCS	?

Table 4 (continued)

Genus	Virus	Size of H(N) or G in amino acid residues	Sequence	Receptor usage
	Avian paramyxovirus 5	HN 574	NRKSCS	?
	Avian paramyxovirus 6	HN 613	NRKSCS	?
	Avian paramyxovirus 7	HN 569	NRKSCS	?
	Avian paramyxovirus 8	HN 577	NRKSCS	?
	Avian paramyxovirus 9	HN 579	NRKSCS	?
	Avian paramyxovirus 10	HN 575	NRKSCS	?
	Avian paramyxovirus 11	HN 583	NRKSCS	?
	Avian paramyxovirus 12	HN 614	NRKSCS	?
	Avian paramyxovirus 13	HN 579	NRKSCS	?
	Avian paramyxovirus 14	HN 580	NRKSCS	?
	Avian paramyxovirus 15	HN 579	NRKSCS	?
	Avian paramyxovirus 16	HN 618	NRKSCS	?
	Avian paramyxovirus 17	HN 599	NRKSCS	?
	Avian paramyxovirus 18	HN 591	NRKSCS	?
	Avian paramyxovirus 19	HN 587	NRKSCS	?

In quotation marks the nomenclature used in the data bank submissions

is a general and immediate perception of what *Mononegavirales* are (ignoring the rhabdoviruses assigned to the genus *Dichornavirus*). The implications of being classified as a paramyxovirus in terms of strategy of gene expression, replication, numbers and likely types of genes and even part of the gene order are generally recognised as all are variations of the basic genome order N–P/V–M–F–G + H(N)–L. However, with classification in lower taxa such as genus and species none of the criteria that are discussed here lead to a classification that is easily understood or compliant with the classical ones. This particularly the case for the “viruses” for which nothing else would be known than only a full length genome sequence (assuming that the sequence is correct and not based on some defective virus).

With the avalanche of metagenomic data, we understand that classical criteria are no longer up to the task. Sequence based methodologies informed by the relevant biological context may be the only practical way forward. Even so, any sequence-based criteria would have to be fluid. It is inevitable that immediacy of recognition will be lost, if it is necessary to know what specific motif variations in the RdRp or what position in a phylogenetic tree are used as criteria for classification in a specific genus or species. While it might be possible to generate numerous new genera and species on the basis of nucleotide or protein sequence variation in a description of the variation and possible evolutionary relationships within the family, the usefulness of this activity and its acceptance by the community of virologists remains an open question in the near term. In the long term, as the datasets become more complete, and machine learning algorithms more sophisticated, sequenced based methodologies

for virus classification might more accurately reflect the true evolutionary relationships within this family.

Funding The authors received no funding specific to this paper as they volunteer as members of the ICTV Study Group.

Compliance with ethical standards

Conflict of interest They all declare that they have no conflict of interest in relation to this article.

Human and animal rights statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Alkhovsky S, Butenko A, Eremyan A, Shchetinin A (2017) Genetic characterisation of bank vole virus (BaVV), a new paramyxovirus isolated from kidneys of bank voles in Russia. Arch Virol. <https://doi.org/10.1007/s00705-017-3639-z>
2. Bao Y, Chetvernin V, Tatusova T (2014) Improvements to pairwise sequence comparison (PASC): a genome-based web tool for virus classification. Arch Virol 159:3293–3304. <https://doi.org/10.1007/s00705-014-2197->
3. Brown DD, Collins FM, Duprex WP, Baron MD, Barrett T, Rima BK (2005) “Rescue” of mini-genomic constructs and viruses by combinations of morbillivirus N, P and L proteins. J Gen Virol 2005(86):1077–1081
4. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282
5. Karlin D, Belshaw R (2012) Detecting remote sequence homology in disordered proteins: discovery of conserved motifs in

- the N-termini of mononegavirales phosphoproteins. PLoS One 7(3):e31719. <https://doi.org/10.1371/journal.pone.0031719>
6. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
 7. Langedijk JP, Daus FJ, van Oirschot JT (1997) Sequence and structure alignment of Paramyxoviridae attachment proteins and discovery of enzymatic activity for a morbillivirus hemagglutinin. J Virol 71:6155–6167
 8. Ludlow M, McQuaid S, Cosby SL, Cattaneo R, Rima BK, Duprex WP (2005) Measles virus superinfection immunity and receptor redistribution in persistently infected NT2 cells. J Gen Virol 86:2291–2303
 9. Gould EA, Cosby SL, Shirodaria PV (1976) Salt dependent haemagglutinating measles virus in SSPE. J Gen Virol 33:139–142
 10. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Wang L, ICTV Report Consortium (2017) ICTV virus taxonomy profile: *Pneumoviridae*. J Gen Virol 98(12):2912–2913. <https://doi.org/10.1099/jgv.0.000959>
 11. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, Davison AJ, Delwart E, Gorbalenya AE, Harrach B, Hull R, King AM, Koonin EV, Krupovic M, Kuhn JH, Lefkowitz EJ, Nibert ML, Orton R, Roossinck MJ, Sabanadzovic S, Sullivan MB, Suttle CA, Tesh RB, van der Vlugt RA, Varsani A, Zerbini FM (2017) Consensus statement: virus taxonomy in the age of metagenomics. Nat Rev Microbiol 15(3):161–168. <https://doi.org/10.1038/nrmicro.2016.177>
 12. Thibault PA, Watkinson RE, Moreira-Sota A, Drexler JF, Lee B (2017) Chapter one—zoonotic potential of emerging paramyxoviruses: knowns and unknowns. Adv Virus Res 98:1–55. <https://doi.org/10.1016/bs.aivir.2016.12.001> (Epub 2017 Feb 2)
 13. Zeltina A, Bowden TA, Lee B (2016) Emerging paramyxoviruses: receptor tropism and zoonotic potential. PLoS Pathog 12(2):e1005390. <https://doi.org/10.1371/journal.ppat.1005390>