This form should be used for all taxonomic proposals. Please complete all those modules that are applicable (and then delete the unwanted sections). For guidance, see the notes written in blue and the separate document "Help with completing a taxonomic proposal"

Please try to keep related proposals within a single document; you can copy the modules to create more than one genus within a new family, for example.

MODULE 1: **TITLE, AUTHORS, etc**

Code assigned:	***2015.050a-dB***	(to be completed by ICTV officers)

Short title: To create one (1) new genus, *G7cvirus*, including eight (8) new species within the family *Podoviridae*.
(e.g. 6 new species in the genus *Zetavirus*)

Modules attached (modules 1 and 10 are required)	**1** ☒	**2** ☒	**3** ☒	**4** ☐	**5** ☐
	6 ☐	**7** ☐	**8** ☐	**9** ☐	**10** ☒

Author(s):

Johannes Wittmann – Leibniz Institute DSMZ (Germany)
Andrew M. Kropinski – University of Guelph (Canada)
Jochen Klumpp – ETH Zurich (Switzerland)
Eugene Kulikov – Winogradsky Institute of Microbiology (Russia)
Andrey Letarov – Winogradsky Institute of Microbiology (Russia)
Evelien M. Adriaenssens – University of Pretoria (South Africa)

Corresponding author with e-mail address:

Andrew M. Kropinski Phage.Canada@gmail.com

List the ICTV study group(s) that have seen this proposal:

A list of study groups and contacts is provided at http://www.ictvonline.org/subcommittees.asp . If in doubt, contact the appropriate subcommittee chair (fungal, invertebrate, plant, prokaryote or vertebrate viruses)	Bacterial & Archaeal Virus Subcommittee

ICTV Study Group comments (if any) and response of the proposer:

Please note that we have chosen to refer to the new genera as *G7cvirus* rather than *G7clikevirus*, etc. since the Bacterial and Archaeal Virus Subcommittee of ICTV has voted overwhelmingly in favour of eliminating "*like*" from phage genus names. Currently in GenBank there are over 30 fully sequenced N4-like phage genomes, which while displaying similar genome length; and, the presence of a high molecular weight virion-associated RNA polymerase; are poorly related at the phylogenetic (Fig. 1), genomic and proteomic levels. At this time we do not want to propose higher orders, until a firm molecular basis can be proposed to define these.

Date first submitted to ICTV: June 2015
Date of this revision (if different to above):

ICTV-EC comments and response of the proposer:

MODULE 2: **NEW SPECIES**

creating and naming one or more new species.

If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be "unassigned" within a subfamily or family. Wherever possible, provide sequence accession number(s) for **one** isolate of each new species proposed.

Code	***2015.050aB***	(assigned by ICTV officers)

To create 8 new species within:

Genus:	***G7cvirus* (new)**
Subfamily:	
Family:	***Podoviridae***
Order:	***Caudovirales***

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no genus is specified, enter "**unassigned**" in the genus box.

Name of new species:	**Representative isolate: (only 1 per species please)**	**GenBank sequence accession number(s)**
Escherichia virus G7C	Escherichia phage vB_EcoP_G7C	HQ259105
Escherichia virus APEC7	Escherichia phage vB_EcoP_PhAPEC7	KF562340
Escherichia virus APEC5	Escherichia phage vB_EcoP_PhAPEC5	KF192075
Escherichia virus Bp4	Escherichia phage Bp4	KJ135004.2
Escherichia virus ECBP1	Escherichia phage ECBP1	JX415535
Escherichia virus IME11	Escherichia phage IME11	JX880034
Escherichia virus EC1UPM	Escherichia phage EC1-UPM	KC206276.2
Shigella virus Sb1	Shigella phage pSb-1	KF620435

Reasons to justify the creation and assignment of the new species:

- Explain how the proposed species differ(s) from all existing species.
  - If species demarcation criteria (see module 3) have previously been defined for the genus, **explain how the new species meet these criteria**.
  - If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.
- Further material in support of this proposal may be presented in the Appendix, Module 9

We have chosen 95% DNA sequence identity as the criterion for demarcation of species in this new genus. Each of the proposed species differs from the others with more than 5% at the DNA level as confirmed with the BLASTN algorithm.

Please note that we have chosen to refer to this new genus as *G7cvirus* rather than *G7clikevirus* since the Bacterial and Archaeal Virus Subcommittee of ICTV has voted overwhelmingly in favour of eliminating "*like*" and "*Phi*" from phage genus names.

MODULE 3: **NEW GENUS**

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	***2015.050bB***	(assigned by ICTV officers)

To create a new genus within:

Subfamily:	
Family:	***Podoviridae***
Order:	***Caudovirales***

Fill in all that apply.

- If the higher taxon has yet to be created (in a later module, below) write "**(new)**" after its proposed name.
- If no family is specified, enter "**unassigned**" in the family box

naming a new genus

Code	***2015.050cB***	(assigned by ICTV officers)

To name the new genus: *G7cvirus*

Assigning the type species and other species to a new genus

Code	***2015.050dB***	(assigned by ICTV officers)

To designate the following as the type species of the new genus

Escherichia virus G7c

Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered

The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). **Please enter here the TOTAL number of species (including the type species) that the genus will contain:**

8

Reasons to justify the creation of a new genus:

Additional material in support of this proposal may be presented in the Appendix, Module 9

N4-like phages are all podoviruses characterized by possessing a large single subunit RNA polymerase which is part of the virion structure. GenBank currently has over 30 phages which are N4-like, but a recent analysis of them revealed that they are genomically, proteomically and phylogenetically (Fig. 1) diverse; and, more importantly, none are closely related to coliphage N4.

BLASTN, CoreGenes and phylogenetic analyses all indicate that the proposed genus, *G7cvirus*, is cohesive and distinct from the other genera in the N4-superfamily of viruses. The next closest member is *Escherichia* phage N4 which shares only 43% DNA sequence identity. The phages of this genus possess genome of approx. 72 kb (43 mol%G+C), and encode 83 proteins and 0-2 tRNAs. They share >74% DNA sequence identity and >78% homologous proteins (Table 1). Figure 3 derived from progressiveMauve analysis reveals the overall sequence similarity between the members of this genus.

"By TEM study, the G7C virions resembled that of N4, about 70 nm in diameter with a small non-

contractile tubular tail, about 25 nm long. In G7C phage, the tail carries a ring-like structure with appendages folded onto the capsid wall and also a collar-like structure carrying another set of appendages similar to that visible on N4 particle" (Fig. 2) (3). This Russian isolate possesses a very limited host range. The genome of G7C has direct terminal repeats of 1160 bp.

PhAPEC5 and PhAPEC7 were isolated in Belgium and possess capsids 65 - 70 nm in diameter (4), and can lyse 9-10 out of 31 tested APEC strains. The only other virus in this group to be characterized by EM is *Shigella boydii* phage pSb-1 which displays a tail length and width of 12 and 19 nm, respectively, and the head diameter was 61 ± 4 nm (7). There is no data on whether it infects *E.coli* strains.

Origin of the new genus name:

Escherichia phage G7C

Reasons to justify the choice of type species:

The first virus of its type that was sequenced

Species demarcation criteria in the new genus:

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

We have chosen 95% DNA sequence identity as the criterion for demarcation of species in this new genus. Each of the proposed species differs from the others with more than 5% at the DNA level as confirmed with the BLASTN algorithm.

MODULE 10: **APPENDIX**: supporting material

additional material in support of this proposal

References:

1. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010; 5(6):e11147.

2. Turner D, Reynolds D, Seto D, Mahadevan P. CoreGenes3.5: a webserver for the determination of core genes from sets of viral and small bacterial genomes. BMC Res Notes. 2013; 6:140.

3. Kulikov,E., Kropinski,A.M., Golomidova,A., Lingohr,E., Govorun,V.,Serebryakova,M., Prokhorov,N., Letarova,M., Manykin,A.,Strotskaya,A. and Letarov,A. Isolation and characterization of a novel indigenous intestinal N4-related coliphage vB_EcoP_G7C. Virology 426 (2), 93-99 (2012)

4. Tsonos,J., Oosterik,L.H., Tuntufye,H.N., Klumpp,J., Butaye,P., De Greve,H., Hernalsteens,J.P., Lavigne,R. and Goddeeris,B.M. A cocktail of in vitro efficient phages is not a guarantee for in vivo therapeutic results against avian colibacillosis. Vet. Microbiol. 171 (3-4), 470-479 (2014)

5. Nho,S.W., Ha,M.A., Kim,K.S., Kim,T.H., Jang,H.B., Cha,I.S., Park,S.B., Kim,Y.K. and Jung,T.S.Complete Genome Sequence of the Bacteriophages ECBP1 and ECBP2 Isolated from Two Different *Escherichia coli* Strains. J. Virol. 86 (22), 12439-12440 (2012)

6. Fan,H., Fan,H., An,X., Huang,Y., Zhang,Z., Mi,Z. and Tong,Y. Complete Genome Sequence of IME11, a New N4-Like Bacteriophage. J. Virol. 86 (24), 13861 (2012)

7. Jun,J.W., Yun,S.K., Kim,H.J., Chai,J.Y. and Park,S.C. Characterization and complete genome sequence of a novel N4-like bacteriophage, pSb-1 infecting *Shigella boydii*. Res. Microbiol. 165 (8), 671-678 (2014)

Annex:

Include as much information as necessary to support the proposal, including diagrams comparing the old and new taxonomic orders. The use of Figures and Tables is strongly recommended but direct pasting of content from publications will require permission from the copyright holder together with appropriate acknowledgement as this proposal will be placed on a public web site. For phylogenetic analysis, try to provide a tree where branch length is related to genetic distance.

Fig 1. The virion RNA polymerases of the N4-like phages were aligned using ClustalW2 (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/) and the dnd file was rendered using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). The branching position of N4 is illustrated in red. The RNA polymerase of *Shigella* phage pSb-1 is not included since it contains numerous frameshift errors.

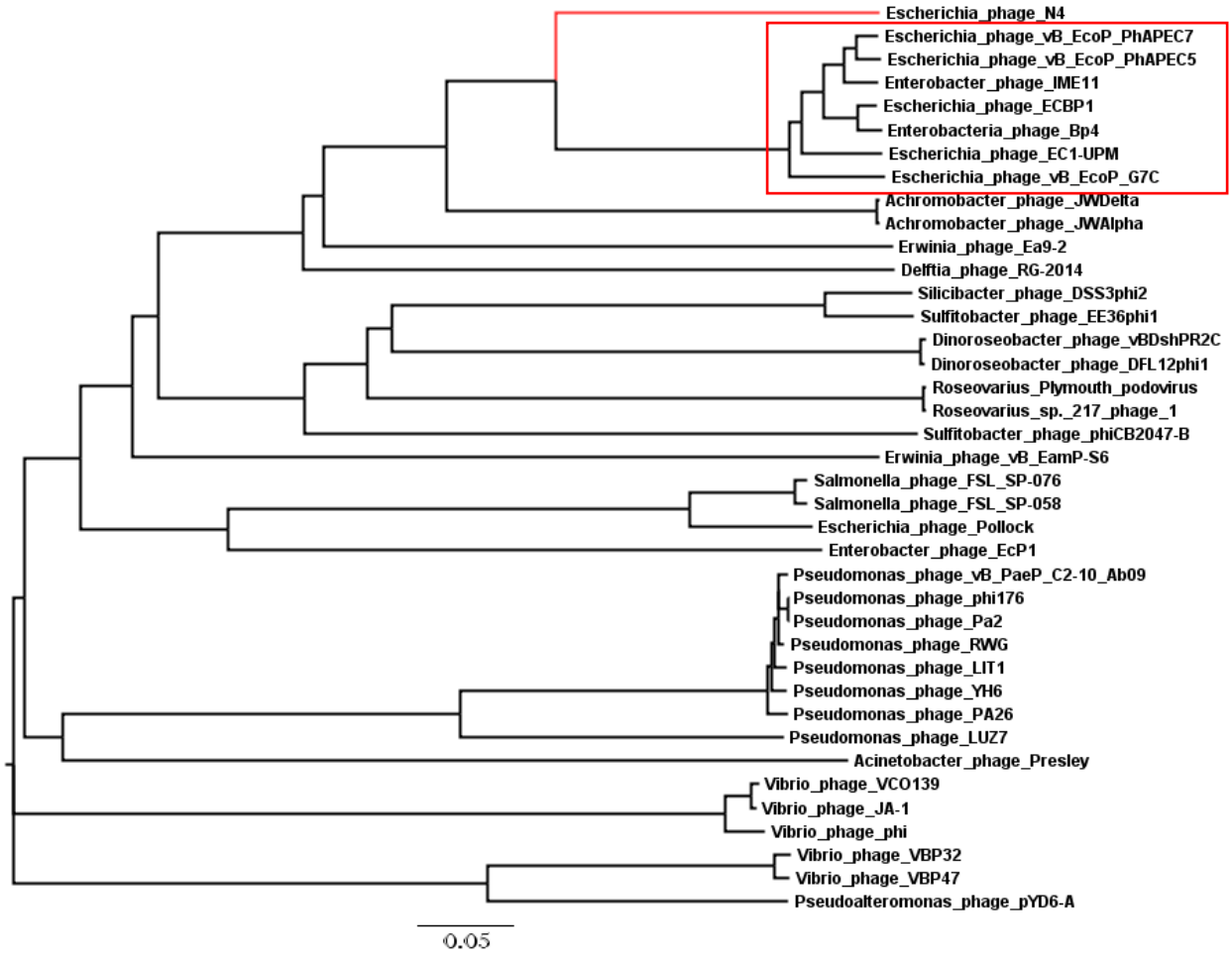

Fig. 2 Electron micrograph of uranyl acetate stained G7C (3).

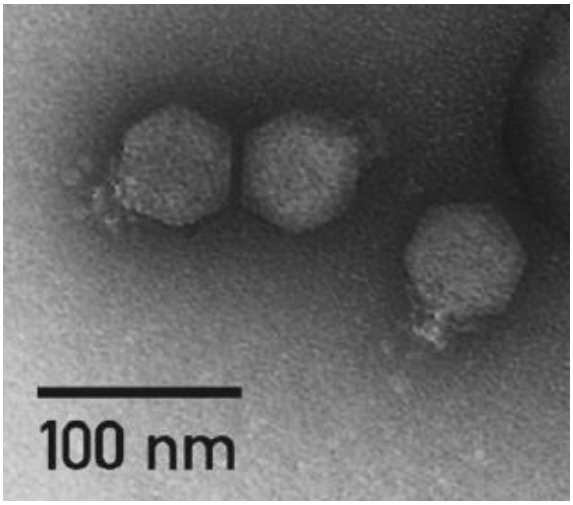

Table 1. Properties of the eight phages belonging to the genus *G7cvirus* and the genomic orphan N4.

Phage	GenBank accession No.	Genome length (kb)	Genome (mol%G+C)	No. CDS	No. tRNAs	DNA (% sequence identity)*	Proteome (% homologous proteins)**
vB_EcoP_G7C	HQ259105	72.92	43.4	79	0	100	100
vB_EcoP_PhAPEC7	KF562340	71.78	43.3	83	1	79	82.3
vB_EcoP_PhAPEC5	KF192075	71.25	43.5	83	1	76	78.5
Bp4	KJ135004.2	72.61	42.8	98***	2	78	84.8
ECBP1	JX415535	69.86	42.7	82	2	74	83.5
IME11	JX880034	72.57	43.1	91	0	74	82.3
EC1-UPM	KC206276.2	70.91	42.9	80	0	76	84.8
pSb-1	KF620435	71.63	42.7	103***	0	77	92.4
N4	EF056009	70.15	41.3	72	0	43	68.4

* Determined using BLASTN; ** Determined using CoreGenes (2); *** these genomes contain numerous frameshifts

Fig. 3. progressiveMauve alignment of the annotated genomes of members of the *G7cvirus* genus – from top to bottom: G7C, IME11, Bp4, EC1-UPM, PhAPEC5, PhAPEC7, pSb-1 and ECBP1(1). Colored blocks indicate the regions of 1 to 1 best alignment with rearrangement breakpoints in a different random color. The degree of sequence similarity between regions is given by a similarity plot within the colored blocks with the height of the plot proportional to the average nucleotide identity (Aaron Darling, personal communication). As can be seen the genomes of Bp4, IME11and pSb-1 are not collinear with G7C.


5000 10000 15000 20000 25000 30000 35000 40000 45000 50000 55000 60000 65000 70000
Escherichia phage vB_EcoP_G7C
Enterobacter phage IME11
Enterobacteria phage Bp4
Escherichia phage EC1-UPM
Escherichia phage vB_EcoP_PhAPEC5
Escherichia phage vB_EcoP_PhAPEC7
Shigella phage pSb-1
Escherichia phage ECBP1