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

<table>
<thead>
<tr>
<th>Code assigned:</th>
<th>2015.050a-db (to be completed by ICTV officers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short title:</td>
<td>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)</td>
</tr>
<tr>
<td>Modules attached</td>
<td>1 ✗ 2 ✗ 3 ✗ 4 ✗ 5 ✗ 6 ✗ 7 ✗ 8 ✗ 9 ✗ 10 ✗ (modules 1 and 10 are required)</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Code</th>
<th>2015.050aB</th>
<th>(assigned by ICTV officers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>To create 8 new species within:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genus:</td>
<td>G7c virus (new)</td>
<td></td>
</tr>
<tr>
<td>Subfamily:</td>
<td>Podoviridae</td>
<td></td>
</tr>
<tr>
<td>Family:</td>
<td>Podoviridae</td>
<td></td>
</tr>
<tr>
<td>Order:</td>
<td>Caudovirales</td>
<td></td>
</tr>
</tbody>
</table>

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 ECBP1| Escherichia phage ECBP1                               | 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 G7c virus 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.

To create a new genus within:

<table>
<thead>
<tr>
<th>Subfamily:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Family:</td>
<td>Podoviridae</td>
</tr>
<tr>
<td>Order:</td>
<td>Caudovirales</td>
</tr>
</tbody>
</table>

Code 2015.050bb (assigned by ICTV officers)

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

To name the new genus: G7c

Assigning the type species and other species to a new genus

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, G7civirus, 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:**


**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 \textit{G7c}virus and the genomic orphan N4.

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

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

Fig. 3. progressiveMauve alignment of the annotated genomes of members of the \textit{G7c}virus 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, IME11 and pSb-1 are not collinear with G7C.