This Word module should be used for all taxonomic proposals.



Please complete **Part 1** and:

either **Part 3** for proposals to create new taxa or change existing taxa

or **Part 2** for proposals of a general nature.

Submit the completed Word module, together with the accompanying Excel module named in Part 3, to the appropriate ICTV Subcommittee Chair.

For guidance, see the notes written in blue, below, and the help notes in file Taxonomic\_Proposals\_Help\_2018.

**Part 1:** **TITLE, AUTHORS, etc**

|  |  |  |  |
| --- | --- | --- | --- |
| **Code assigned:** | ***2018.002M*** | | (to be completed by ICTV officers) |
| **Short title:** One new genus (*Caligrhavirus*), including 3 new species, in the family *Rhabdoviridae*. | | | |
|  | | | |
| **Author(s):** | | | |
| Peter J. Walker, peter.walker@uq.edu.au  Are Nylund, are.nylund@uib.no  Arnfinn L. Okland, arnfinn.okland@uib.no | | | |
| **Corresponding author with e-mail address:** | | | |
| Peter J. Walker, [peter.walker@uq.edu.au](mailto:peter.walker@uq.edu.au) | | | |
| **List the ICTV study group(s) that have seen this proposal:** | | | |
|  | | ICTV *Rhabdoviridae* Study Group | |
| **ICTV Study Group comments (if any) and response of the proposer:** | | | |
| Proposal supported by a majority of Study Group members (8 supporters and 4 non-responders) | | | |
|  | | | |
| Date first submitted to ICTV: | | | June 6, 2018 |
| Date of this revision (if different to above): | | |  |

|  |
| --- |
| **ICTV-EC comments and response of the proposer:** |
|  |

**Part 3:** **PROPOSED TAXONOMY**

|  |
| --- |
| **Name of accompanying Excel module:** 2018.002M.N.v1.Caligrhavirus |

**Supporting material:**

The new genus *Caligrhavirus* is proposed to accommodate three currently unclassified rhabdoviruses that have been detected in caligid copepods. Each virus is proposed to be assigned to a new species within the new genus.

Lepeophtheirus salmonis rhabdovirus 9 (LSalRV-9) and Lepeophtheirus salmonis rhabdovirus 127 (LSalRV-127) were each detected in salmon lice (*Lepeophtheirus salmonis* Krøyer, 1837*­*) infesting farmed Atlantic salmon (*Salmo salar* Linnaeus, 1758) on the west coast of Norway (1).

Caligus rogercresseyi rhabdovirus (CRogRV) was detected in sea lice (*Caligus rogercresseyi*) infesting farmed Atlantic salmon at Los Lagos on the coast of Chile, in 2016 (2).

No isolates are available for any of these viruses. For LSalRV-127, rhabdovirus-like particles (bacilliform of bullet-shaped, 55 nm diameter and up to 425 nm in length) have been detected by transmission electron microscopy in thin sections of necrotic tissues of salmon louse samples; cross-striated nucleocapsids with helical symmetry were also observed (1). High levels of LSalRV-127 RNA were also detected in infected salmon louse tissues by real-time PCR by *in situ* hybridization using genome sense (-) of mRNA sense (+) hybridization probes (1).

Near-complete genomes containing complete coding sequences and partial terminal sequences have been determined for each of the viruses (1, 2). The genomes range in length from approximately 11.5 kb to 11.7 kb, containing the five canonical rhabdovirus structural protein genes (*N*, *P*, *M*, *G*, and *L*) (**Figure 1**). The CRogRV genome contains an additional long ORF (*U1*) between the *M* gene and *G* gene within a single transcriptional unit bounded by conserved transcription initiation and transcription termination/polyadenylation sequences; ORF *U1* encodes a putative 15.3 kDa acidic protein of unknown function.

Based on well-supported ML trees generated from complete L protein sequences, caligrhaviruses form a monophyletic clade that is distinct from all currently established genera (**Figure 2**). Nucleotide sequence identity (p-distance) between caligrhavirus genomes is relatively low (26.7% to 33.1%) (**Table 1**). Amino acid sequence identities are also relatively low (<30% in the N and G proteins and <50% in the L proteins) (**Tables 2–4**).

Amino acid sequence alignment of G proteins indicates that caligrhaviruses have a unique arrangement of conserved cysteine residues in the ectodomain featuring all 12 VSIV G protein disulphide bridge-forming cysteine residues (CI**–**CXII) and two additional cysteine residues in the N terminal and C terminal regions that are likely to form an additional, unique disulphide bridge (**Figure 3**). Shared and unique patterns of disulphide bridge formation occur commonly for viruses assigned to individual genera of *Rhabdoviridae*.

**Genus description.**

Caligrhaviruses infect sea lice, i.e., crustaceans in the family Caligidae (Copepoda: Siphonostomatoida), marine [ectoparasites](https://en.wikipedia.org/wiki/Ectoparasite) that feed on the blood and tissues of marine fish. These viruses form a monophyletic clade based on Maximum Likelihood trees inferred from alignments of complete rhabdovirus L gene sequences.

**Species demarcation criteria.**

Viruses assigned to different species within the genus *Caligrhavirus* have several of the following characteristics: A) minimum amino acid sequence divergence of 15% in N proteins; B) minimum sequence divergence of 20% in the L proteins; C) minimum amino acid sequence divergence of 20% in G proteins; D) significant differences in genome organization as evidenced by numbers and locations of ORFs; E) can be distinguished in virus neutralisation tests; and F) occupy different ecological niches as evidenced by differences in hosts and/or vectors.

All proposed members of the new genus meet demarcation criteria A, B, and C. CRogRV also meets demarcation criteria D and F.

**Derivation of the genus name.**

*Caligrhavirus* is derived from Caligidae, the family of copepods that includes sea lice, and rhabdovirus.

**Type species.**

*Lepeophtheirus caligrhavirus* is designated as the type species of the genus as the virus assigned to it (Lepeophtheirus salmonis rhabdovirus 127) is the most thoroughly described of the three currently known caligrhaviruses.

|  |
| --- |
| **Figures**  **Figure 1.** Caligrhavirus genome organisations (shown in positive polarity). Arrows indicate the locations of long open reading frames (ORFs), each of which is located within a single transcriptional unit bounded by conserved transcription initiation and transcription termination/polyadenylation sequences. N, P, M, G, and L represent ORFs encoding the canonical rhabdovirus structural protein genes. An additional ORF (shaded grey) in CRogRV encodes a putative protein of unknown function.  **Figure 2.** The evolutionary history was inferred from a Clustal W alignment of complete L protein sequences of 126 rhabdoviruses currently assigned to species and three proposed caligrhaviruses. Phylogenetically informative sites were selected from the alignment using GBLOCKS resulting in 560 positions in the final dataset. The tree was inferred in MEGA by using the Maximum Likelihood method based on the Whelan And Goldman + Freq. model. The tree with the highest log likelihood (-60612.9864) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. 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. Bootstrap values (1,000 iterations) are shown for each node. [Note that viruses currently assigned to the genus *Nucleorhabdovirus* are not monophyletic.]  LSalRV\_9\_G M--LTRTARTTSTSLFIKLSILCVLTDHITSGD----------KNGTTDEK--KSHFVYP  LSalRV\_127\_G M--VQLVLQSFWTALILNLSIP------PSSGDIQDNTIEWRHPVFSVPDKGVEVNLMIH  CRogRV\_G MKSVFFFFVILFNTMDADLSIPLGNTNNLPGLTTPLEPIVEGRSRTSRKEKDFGEMTVLP  \* : .:: .\*\*\* .. : :\* :  **a CI CII**  LSalRV\_9\_G SRIVGSLLPCLPEDLVCPPTVKDYATVGLTPTKKKLLMPAGGKTEHVGGVLCHLVKKRTT  LSalRV\_127\_G GPIKYPWKTCSRESFVCPPTVHDGIEEGLIKEILDLEVPDNLDKAVIDGYLCHKIRRRTI  CRogRV\_G LRRLEPWRECAIADLQCPPRYDFGEKIGSLITTEKLWPVRGLS-VLQEGYFCTKTSRDRT  . \* .: \*\*\* . \* .\* . . \* :\* :  **CIII CIV CV**  LSalRV\_9\_G CWTSLWGSNDISQQEYRTPVVLDRCRLAVNNYLRGEHENVEFPESECSWMSTIDMDITGA  LSalRV\_127\_G CETSFFGHQTITYHRYRLVITKDNCREVMREYEEGSYEPSGQPDPECAWMSTDVSDVDIM  CRogRV\_G CSTSFFGSEDLSGSEEYLYPNDSDCLKEVKSLESGRYSPPVWPEHTCAWMATRTVTLVQY  \* \*\*::\* : :: . . \* :.. \* :. \*: \*:\*\*:\* :  **CVI CVII CVIII**  LSalRV\_9\_G IITPHETFFDPYKTSVYDKHLIR-SCRNRVCETVRRDLVWFALEEFPLPSDLFQKQDCII  LSalRV\_127\_G EITPHPVFFNPFDSSMEDAVIKG-FCKGEFCQTHREGTLWIRKSSS-KPALDYKKVSCNF  CRogRV\_G QLNLHNVLWEDIGGTYHDAKLKGGKCSTRICSTNNPGVLWIRGKKANHLLRPEDRLPCKI  :. \* .::: : \* : \* ..\*.\* . . :\*: .. .: \* :  **CIX CX**  LSalRV\_9\_G YSSD--PDNEATLIKCEGYPYLTISNQSCQINYGGRTGVATPHHFAIFGNIPGHDSLPP-  LSalRV\_127\_G FRGEEITDHLKWLIKCPRFPYFKVNKRTCRTILDGKSALGNPDGLIIITDLGKKGPLPT-  CRogRV\_G YENK---TSSILQVHCDYHHPLHFKVGACKFSYRGESGIRSEEGVGLAWDLKKGSKIARY  : .. ::\* . : .. :\*: \*.:.: . . . : :: . :.  **CXI CXII**  LSalRV\_9\_G ----CSDN-VIIGVTGPMEKSIERRGEYADMNLRERCLDAIDRITNENTVTLRTLGHFMP  LSalRV\_127\_G ----CSKK-VKFGLLKSSGQIQSGLEEIKDDFLYERCIDSLSRIVHQSSVNFRDLGFFYP  CRogRV\_G VGPECDKKKTPIYQWSANSKFRYDAATKADDDLHARCLDALVRIREHKDISQWDLGYFYP  \*.: . : . : \* \* \*\*:\*:: \*\* ... :. \*\*.\* \*  **b**  LSalRV\_9\_G RSPGRHPVYILINST--LMCGSAKYKEYTGSLDNDNLWTLLKNSLWVHWADDNHLSYNGV  LSalRV\_127\_G RSPGLHPLYMLKDKK--IWCNRALFEEKKVTLDDKSLEKTFVHYSWMRWKNTSLYQSANG  CRogRV\_G SSPGPFPAYRLNKSSKVLECSKYLFTLKEVKSGRSLVDYLPEESIIPDPRTGEKLGVSGL  \*\*\* .\* \* \* ... : \*. : . . . : . . .  LSalRV\_9\_G LRYGKESQP-----EIHIPHLNDIRDSLVALHTEELELIPTNRVVFIPDKEV-PLNASIP  LSalRV\_127\_G VTWNKETNSSDSGTKYHIPHLMEMSASLAITHTRGHSVAYPRTFHIRQNMDD-PLTDVDT  CRogRV\_G IVKND--------SSVYVPHQRSNQMISWEHHLGPKDKVEISRMSLTPNVDDNTLLFSED  : .. . ::\*\* . \* . . : : : .\*  LSalRV\_9\_G TQRKTTKIGGDLGTSFNAIGSWFQTNMSLLITLGTGMALLGLGYVLLMIVLKCIKGLKKP  LSalRV\_127\_G HVQKRHPLGENFAKWWDNLGEVVKGVISTITTLIVILVSLFIVWVLVFCLKKIGGGCREN  CRogRV\_G PQNHATENGTSQATFLSHLGFGIEHLFYTLLGLS-IAGLVGVCVVKVCLENAFKCGCKYC  .: \* . .. . :\* .: : : \* : : \* : \* :  LSalRV\_9\_G PKSDSINLNLQPFQPSGKDRAEH  LSalRV\_127\_G VNELSKPNNEFGMDSWGSQLP--  CRogRV\_G SRSDLDDYSH-------------  .. .  **Figure 3.** Clustal X alignment of caligrhavirus G protein sequences illustrating conservation of cysteine residues in the ectodomain. Residues CI to CXII align with 12 cysteine residues in the G protein of vesicular stomatitis Indiana virus which are known to form six disulphide bridges. Two additional cysteine residues (a and b) are likely to form an additional disulphide bridge in caligrhavirus G proteins. Predicted N-terminal signal domains and near C-terminal transmembrane domains are shaded in grey. Other fully conserved (\*), strongly conserved (:) and weakly conserved (.) amino acids are indicated below the alignment. |

**Tables**

**Table 1.** Percentage nucleotide sequence identities (p-distance) of a MUSCLE alignment of caligrhavirus genomes.

|  |  |  |  |
| --- | --- | --- | --- |
|  | CRogRV | LSalRV-9 | LSalRV-127 |
| CRosRV |  |  |  |
| LSalRV-9 | 29.7 |  |  |
| LSalRV-127 | 26.8 | 33.1 |  |

**Table 2.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of caligrhavirus N proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | CRogRV | LSalRV-9 | LSalRV-127 |
| CRosRV |  |  |  |
| LSalRV-9 | 24.9 |  |  |
| LSalRV-127 | 28.1 | 27.8 |  |

**Table 3.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of caligrhavirus G proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | CRogRV | LSalRV-9 | LSalRV-127 |
| CRosRV |  |  |  |
| LSalRV-9 | 21.1 |  |  |
| LSalRV-127 | 21.7 | 27.2 |  |

**Table 4.** Percentage amino acid sequence identities (p-distance) of a CLUSTAL W alignment of caligrhavirus L proteins.

|  |  |  |  |
| --- | --- | --- | --- |
|  | CRogRV | LSalRV-9 | LSalRV-127 |
| CRosRV |  |  |  |
| LSalRV-9 | 44.9 |  |  |
| LSalRV-127 | 44.9 | 49.8 |  |

| **References:** |
| --- |
| 1. **Økland AL, Nylund A, Øvergård AC, Blindheim S, Watanabe K, Grotmol S, Arnesen C-E, Plarre H.** 2014. Genomic characterization and phylogenetic position of two new species in *Rhabdoviridae* infecting the parasitic copepod, salmon louse (*Lepeophtheirus salmonis*). PLoS One **9:**e112517.  2. **Økland AL, Skoge RH, Nylund A.** 2018. The complete genome sequence of CrRV-Ch01, a new member of the family *Rhabdoviridae* in the parasitic copepod *Caligus rogercresseyi* present on farmed Atlantic salmon (*Salmo salar*) in Chile. Archives of Virology doi:10.1007/s00705-018-3768-z. |