Template for Taxonomic Proposal to the ICTV Executive Committee
To create a new Family

Code†	**2003.094F.01**	To create a new family* of ssRNA picorna-like viruses infecting red algae
Code†	**2003.095F.01**	To name the new family* ***Marnaviridae***
Code†	**2003.096F.01**	To create a new genus in the family created in 2003.094F.01*
Code†	**2003.097F.01**	To designate the following genera as part of the new family*: ***Marnavirus***
Code†	**2003.098F.01**	To designate the following virus as type species in the genus*: ***Heterosigma akashiwo* RNA Virus (HaRNAV)**

† Assigned by ICTV officers
° Leave blank is not appropriate
* repeat these lines and the corresponding arguments for each genus created in the family

Author(s) with email address(es) of the Taxonomic Proposal

Curtis A. Suttle, Department of Earth and Ocean Sciences, University of British Columbia
email: suttle@eos.ubc.ca

Andrew S. Lang, Department of Microbiology and Immunology, University of British Columbia
email: aslang@interchange.ubc.ca

Alexander I. Culley, Department of Botany, University of British Columbia, 1461-6270
email: culley@interchange.ubc.ca

New Taxonomic Order

Order
Family ***Marnaviridae***
Genus ***Marnavirus***
Type Species ***Heterosigma akashiwo* RNA Virus**
List of Species in the genus ***Heterosigma akashiwo* RNA Virus**
List of Tentative Species in the Genus
List of Unassigned Species in the Family

Argumentation to create a new family:

HaRNAV is a single-stranded, positive sense, RNA virus that infects several strains of the red-tide forming Raphidophyte *Heterosigma akashiwo* from coastal British Columbia waters. Our analysis of the HaRNAV genome establishes that it is clearly in the picorna-like superfamily based on the following criteria HaRNAV particles are icosahedral with a diameter of approximately 25 nm (Tai, 2003), a size and structure consistent with picorna-like viruses (reviewed in Liljas, 2002).The genome is 8587 nts in length, plus a poly(A) tail The determined genome sequence contains 1 large open reading frame (orf) on the positive strand that is 7743 bases long and predicted to encode a protein of 2581 amino acid residues. The 5' and 3' untranslated regions (UTRs) are 483 and 361 nucleotides long, respectively, accounting for a total of 9.8 % of the genome. There is a potential stem-loop structure (with 14 of 15 bases capable of hybridizing and a loop of 4 bases) starting at the fifth nucleotide The secondary structure near the 5' end is likely functionally important for the virus for replication of the RNA as seen in other picorna-like viruses (e.g. poliovirus (Andino, 1993). Secondary structures close to the start of the polyprotein are likely functionally important as part of an internal ribosome entry site (IRES) for translation of the polyprotein as in other picorna-like viruses (reviewed in Martinez-Salas, 2001). There is a notable pyrimidine-rich stretch of sequence wherein 22 of 29 bases are pyrimidines that ends 8 bases upstream of the predicted start codon of the large orf, and such sequences are conserved in picorna-like viruses and important as part of the IRES (Pestova, 1991). We performed N-terminal sequence analysis of protein bands from purified virus particles. These sequences were found in the amino acid sequence predicted for the large orf in the genome sequence, which therefore encodes the viral structural proteins Analsysis of the predicted polyprotein sequence by BLAST (Altschul et al., 1997) revealed the presence of conserved domains defined for picorna-like helicase and RNA-dependent RNA polymerase (RDRP) proteins, and capsid proteins (Figure 1).

The structure of the viral genome and the patterns of sequence relationships of HaRNAV proteins to other known viral picorna-like proteins clearly shows that it does not belong with any of the currently established picorna-like families. The HaRNAV genome structure is most like the potyviruses [e.g. tobacco etch virus; (Allison, 1986) in that the non-structural protein domains are located at the N-terminus and the structural proteins are at the C-terminus of a single large polyprotein encoded on a monopartite genome. However, potyvirus capsids are filamentous and NCBI database searches with putative HaRNAV open reading frames, as well as phylogenetic analyses demonstrated no significant homology with this family. We constructed phylogenetic trees to evaluate the evolutionary relationship of the HaRNAV RNA-dependent RNA polymerase (RdRp) to RdRp sequences from other families within the picorna-like superfamily. The alignments were done with residues 1362-1619 of the HaRNAV polyprotein that represent the conserved regions I-VIII (Koonin and Dolja, 1993) and the corresponding regions from the other viruses included. The trees we constructed using the maximum likelihood and neighbor-joining methods (Figure 2) support previously established family classifications within the picorna-like superfamily. However, these analyses do not place the HaRNAV sequence within any of these established families. This is not surprising that HaRNAV is in a new family as it is the first picorna-like virus that has been described that infects a protist.

Origin of the proposed family/genus name

Marna is a sigla derived from ma<u>re</u> (mare: the sea (*latin*)), referring to the environment from which the virus was isolated and <u>RNA</u>, referring to the genetic material of the virus.

References

Allison, R.F., Johnson, R.E., and Dougherty, W.G. (1986). The nucleotide sequence of the coding region of Tobacco etch virus genomic RNA: Evidence for the synthesis of a single polyprotein. *Virology*. **154**, 9-20.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc. Acid. Res.* **25**, 3389-3402.

Andino, R., Rieckhof, G.E., Achacoso, P.L., Baltimore, D.A.. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J*. **12**, 3587-3598.

Koonin, E.V., and Dolja, V.V. (1993). Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28**, 375-430.

Liljas, L., Tate, J., Lin, T., Christian, P., and Johnson, J.E. (2002). Evolutionary and taxonomic implications of conserved structural motifs between picornaviruses and insect picorna-like viruses. *Arch. Virol.* **147**, 59-84.

Martinez-Salas, E., Ramos, R., Lafuente, E., and Lopex de Quinto, S. (2001). Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *J. Gen. VIrol.* **82**, 973-984.

Pestova, T.V., Hellen, C.U.T., and Wimmer, E. (1991). Translation of Poliovirus RNA: Role of essential *cis*-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57-kilodalton protein. *J. Virol*. **65,** 6194-6204.

Tai, V., Lawrence, J.E., Lang, A.S., Chan, A.M., Culley, A.I., and Suttle, C.A. (2003). Characterization of HaRNAV, a single-stranded RNA virus causing lysis of *Heterosigma akashiwo* (Raphidophyceae). *J. Phycol.* In press.

Annexes:

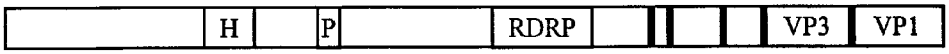

Figure 1. Representation of the HaRNAV polyprotein. The location of conserved picornalike protein domains are indicated within the polyprotein box: H, helicase; P, protease; RDRP, RNA-dependent RNA polymerase; VP3 and VP1, structural proteins. The location of the N-termini found by sequencing the HaRNAV structural proteins are shown by black lines in the box.

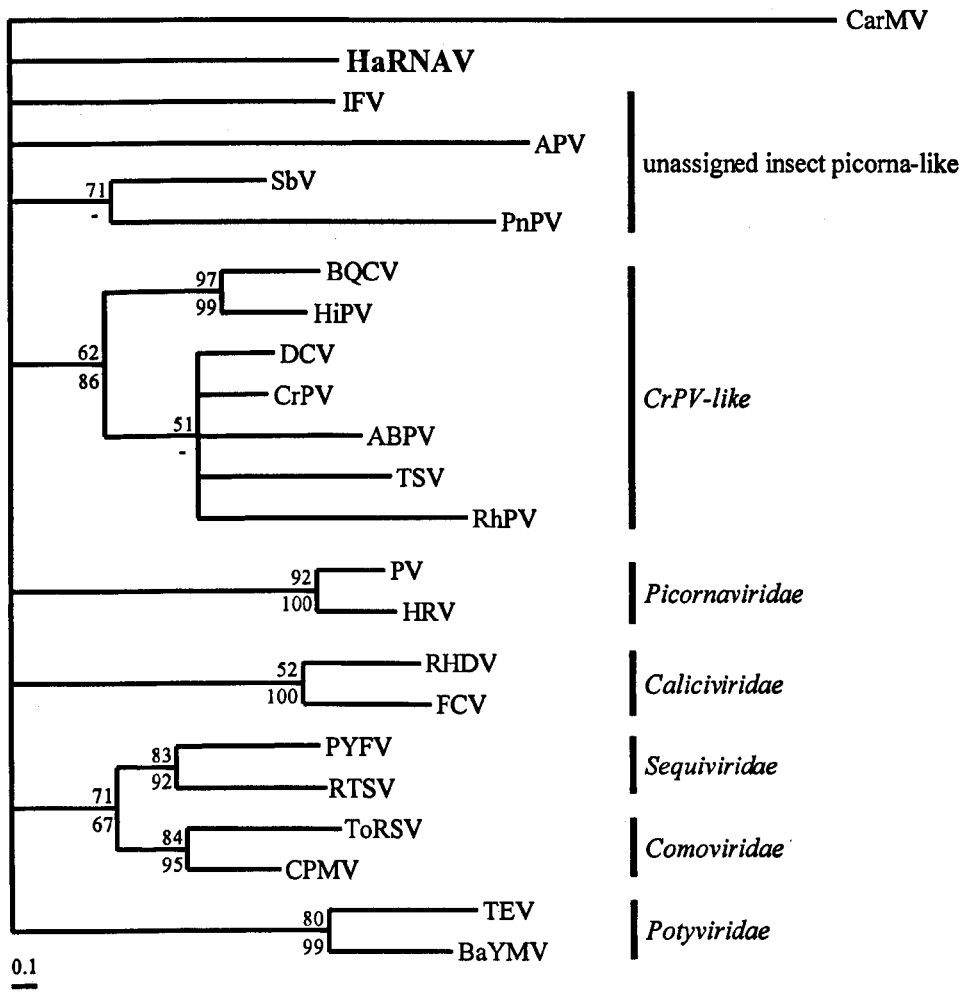

Figure 2. Phylogenetic analysis of picornalike RNA dependent RNA polymerase domain protein sequences. The tree is based on maximum likelihood distances and the sequence from the carnation mottle virus (CarMV) was used as an outgroup. Support values based on 10,000 puzzling steps are shown above the branches. Bootstrap values (based on 1,000 replicates) for branches that are supported by > 50 % by neighbor-joining analysis are labeled below the branches (a dash indicates there was no corresponding branch in the neighbor-joining tree). The scale bar represents the expected number of changes per residue position.