

**Data S1. Motifs identified in the RdRp domains of novel RNA virus phyla inferred in this work.** The conserved motifs G [GxS], F [long motif enriched in basic amino acids], A [DxxxxD], B [(S/T)Gxxx(T/G)xxxN], C [(S/G)DD], D[GxxxK], and E [FL] were manually inspected in multiple sequence alignments of the domain protein sequences, and their identity was supported by HHPred search. Sections of the alignments not containing motifs are not shown. Note that motif A of “*Taraviricota*” is DxxxxE instead of the canonical DxxxxD. Motif G was not found in most of the cases.

