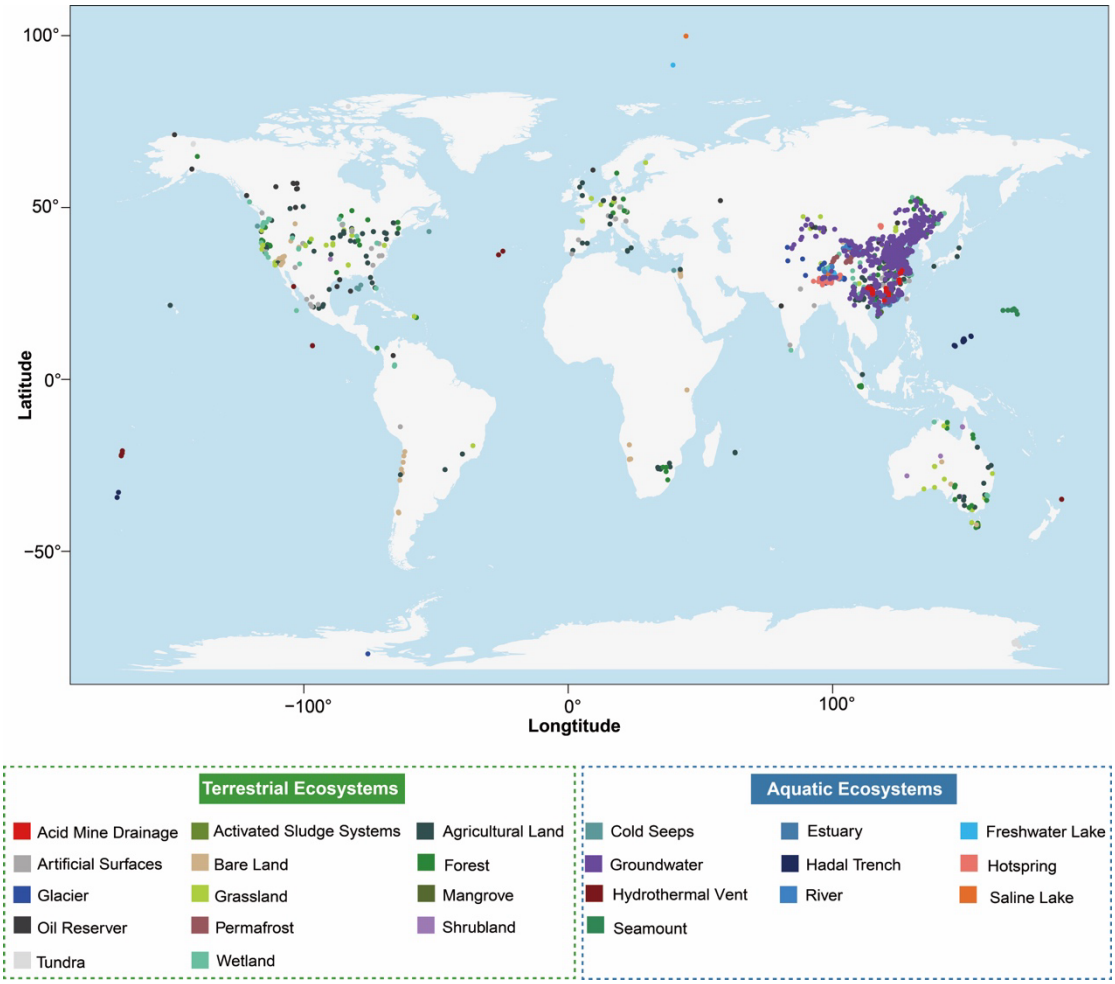


Figure S1: Global distribution of publicly available sampling sites, classified by habitat. Map of global sampling sites derived from publicly available datasets, with habitats color-coded and broadly partitioned into terrestrial and aquatic ecosystems.

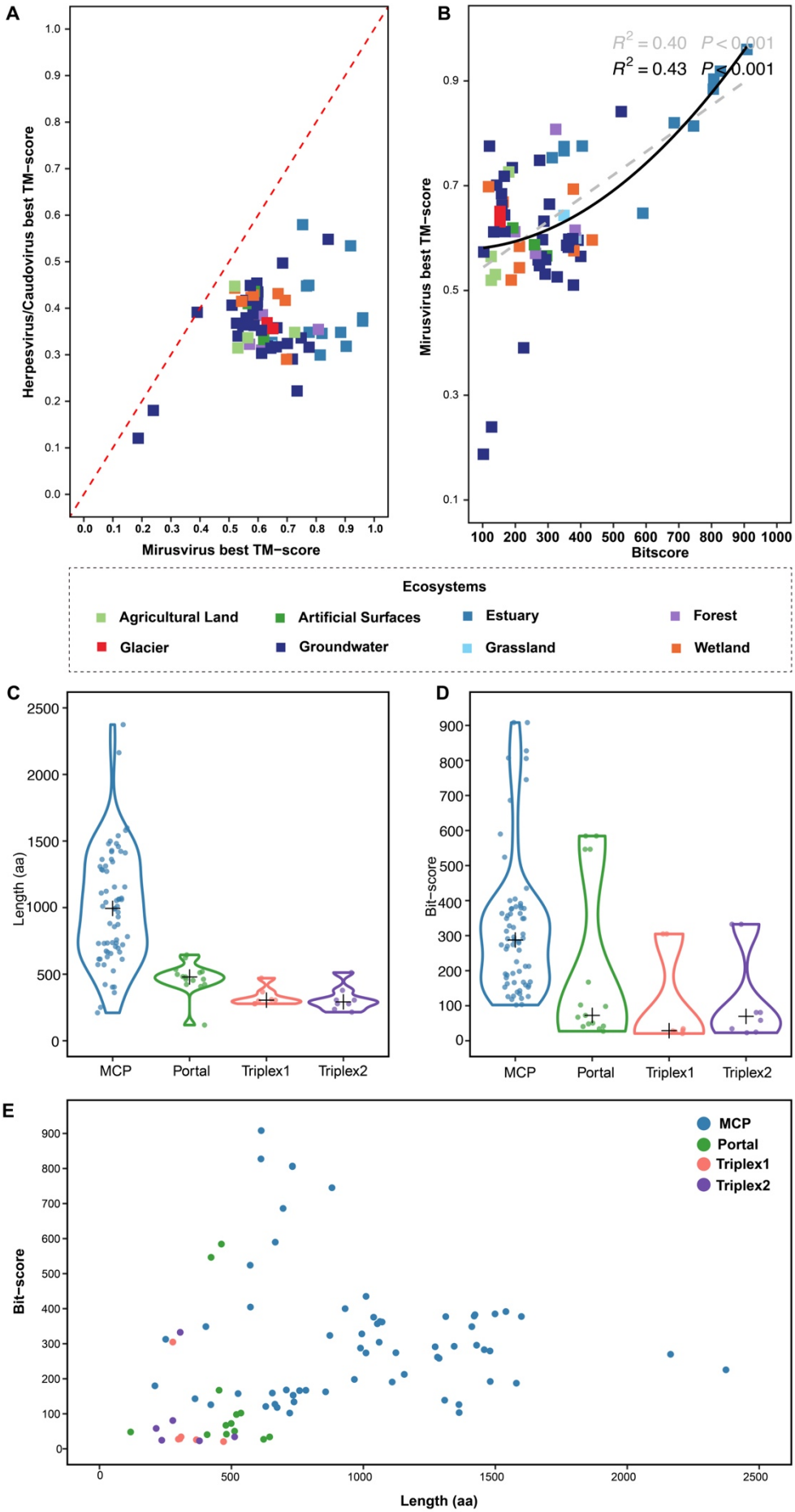

A
Herpesvirus/Caudovirus best TM−score
Mirusvirus best TM−score
B
$R^2 = 0.40$ $P < 0.001$
$R^2 = 0.43$ $P < 0.001$
Mirusvirus best TM−score
Bitscore
Ecosystems
Agricultural Land
Artificial Surfaces
Estuary
Forest
Glacier
Groundwater
Grassland
Wetland
C
Length (aa)
MCP
Portal
Triplex1
Triplex2
D
Bit−score
MCP
Portal
Triplex1
Triplex2
E
Bit−score
Length (aa)
MCP
Portal
Triplex1
Triplex2

Figure S2: Statistical analysis of four mirusvirus marker genes. **(A)** Highest TM-score of the MCP homologs compare with reference mirusviruses and herpesviruses/caudovirus. The slope of the red dashed line is 1. **(B)** Correlation between bitscore and TM score aligning to reference mirusvirus sequences. **(A-B)** Each dot represents an MCP homolog, with the color spectrum indicating their origins from different ecosystems. **(C)** Length distributions of the four mirusvirus marker genes. The black cross in each violin plot indicated the median value. **(D)** Distribution of bit scores for the four mirusvirus marker genes. The black cross in each violin plot indicated the median value. **(E)** Relationships between length and bit score of the four mirusvirus marker genes.

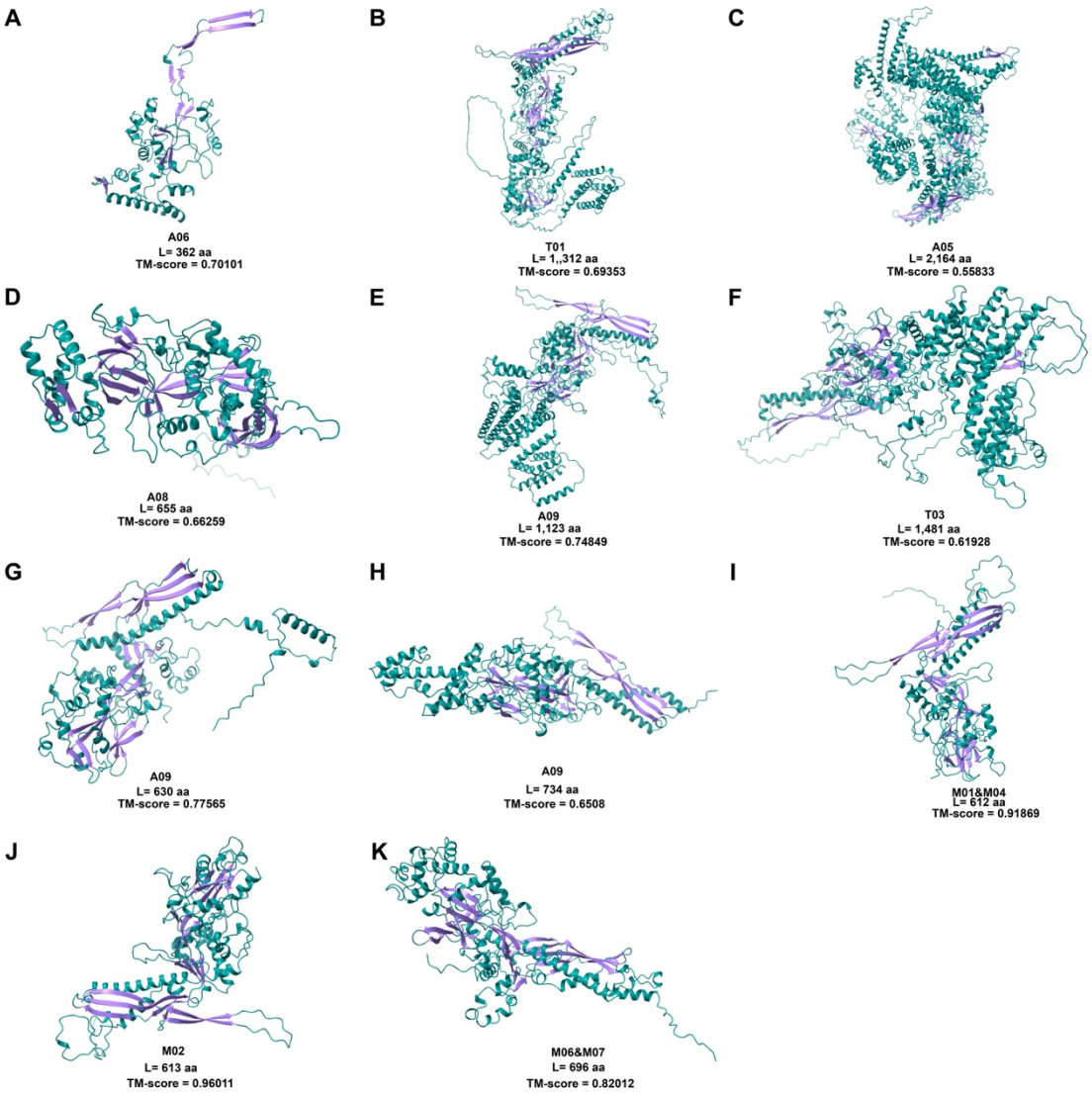


Figure S3: 3D structures of the representative mirusvirus MCPs in each lineage.

(**A-K**) Predicted 3D structures and corresponding TM-scores of representative mirusvirus MCPs from the lineages identified in this study. β-sheets are shown in purple.

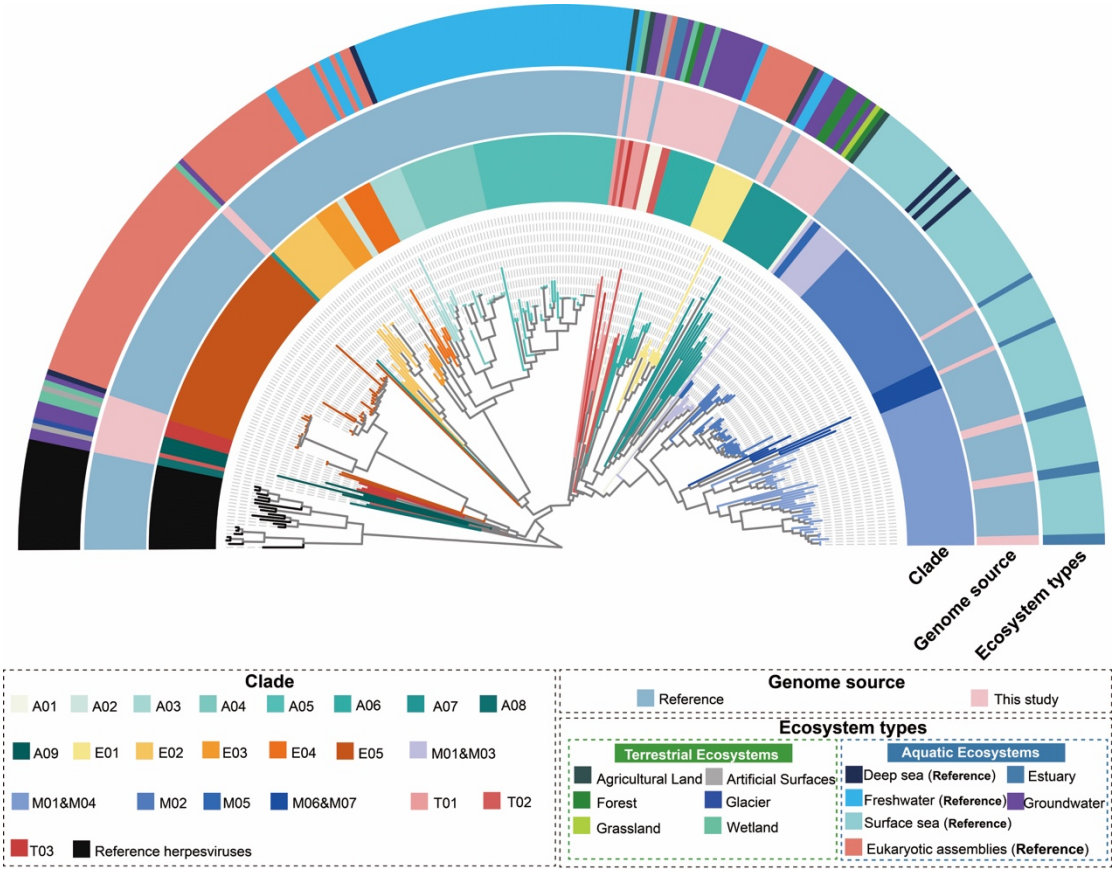


Figure S4: Maximum-likelihood phylogenetic tree of mirusvirus MCP protein sequences constructed using the Q.PFAM+F+I+R7 substitution model. The tree is midpoint-rooted, and 22 reference herpesvirus MCP protein sequences (obtained from NCBI) were used as the outgroup. Tree branches and the innermost half-circle labeled "Clade" are colored according to clade classification. The outer half-circles labeled "Genome source" and "Ecosystem types" are colored to represent genome source and ecosystem type classifications, respectively.

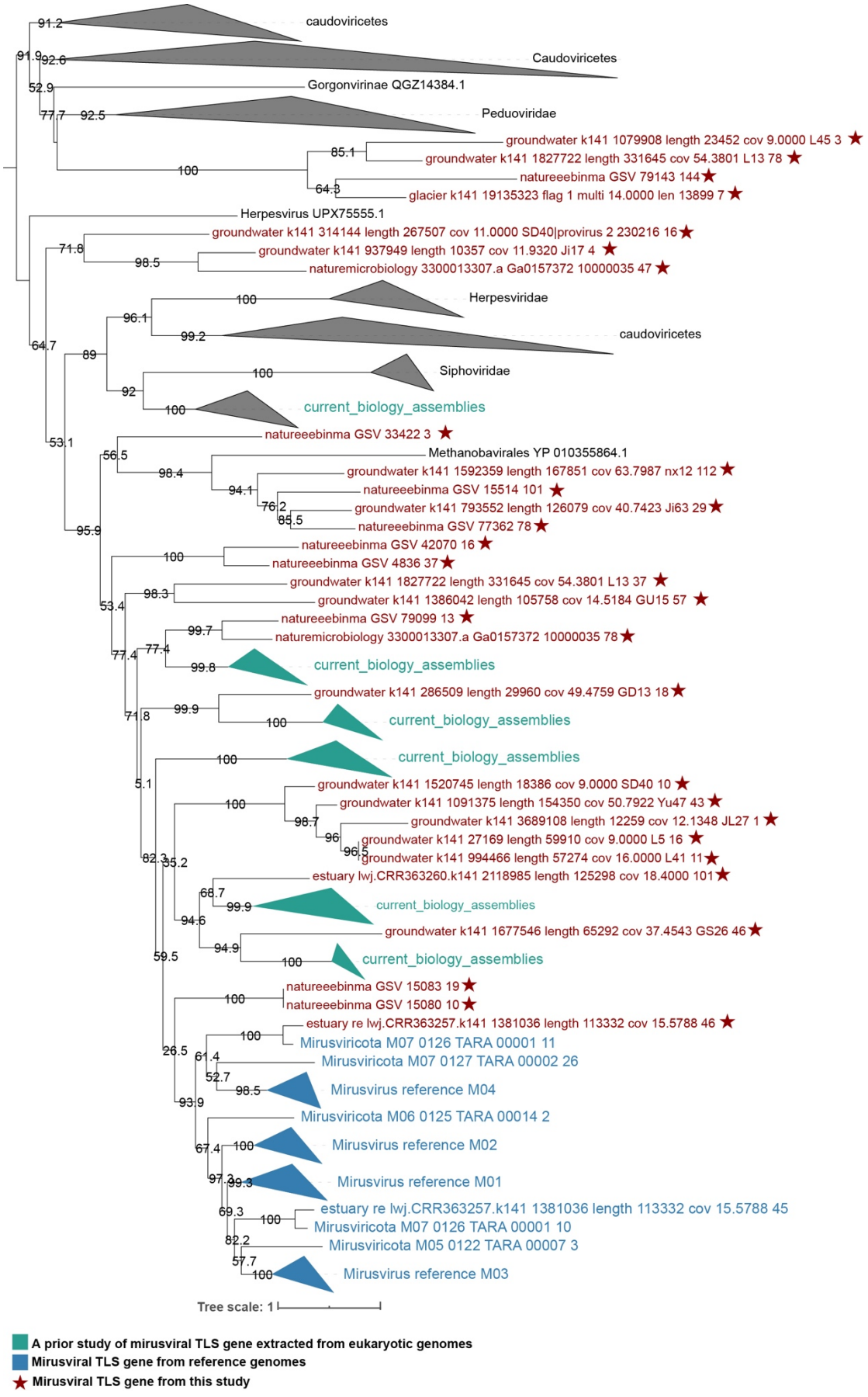
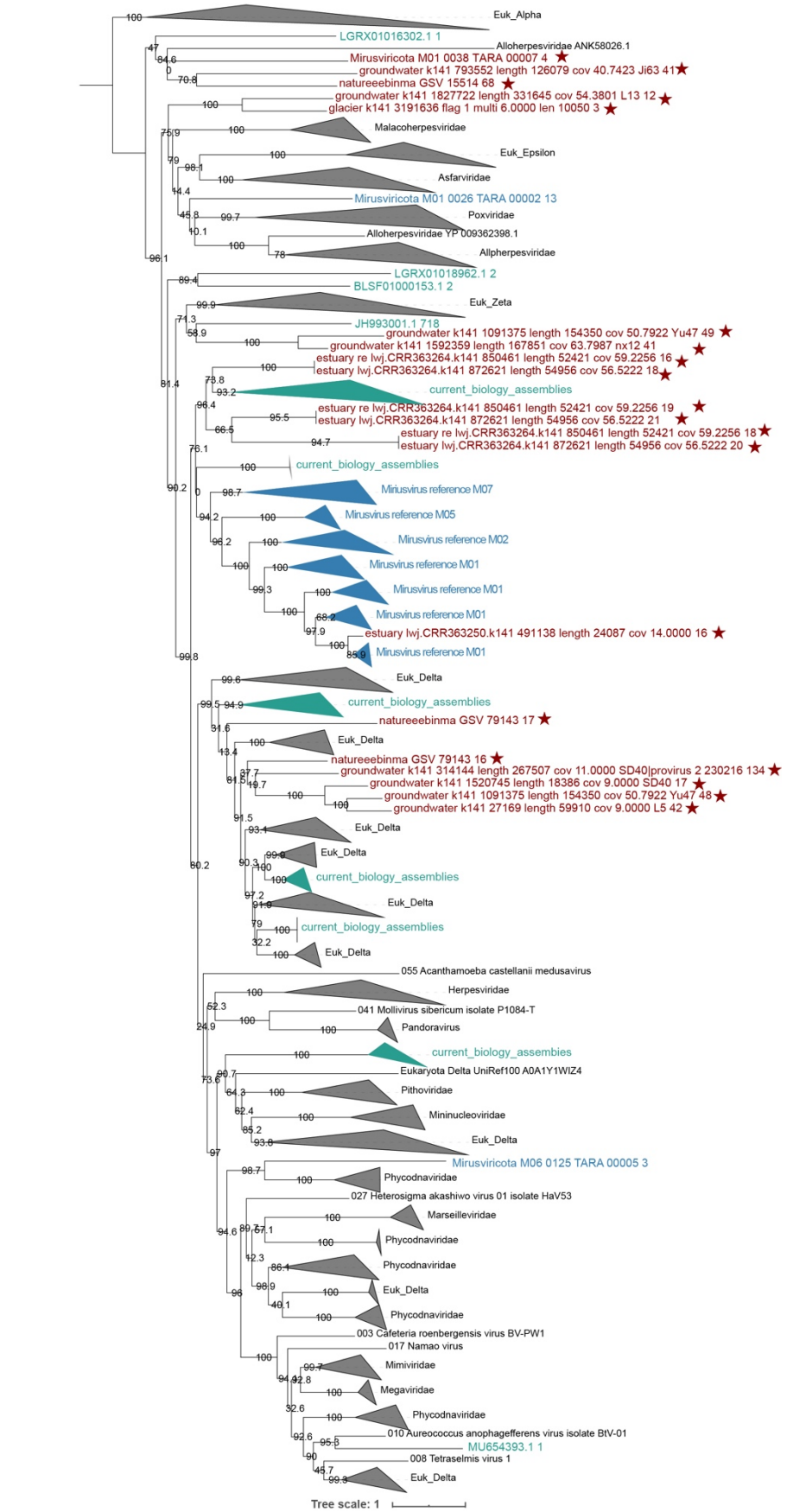


Figure S5: Maximum-likelihood phylogenetic tree of terminase large subunit

(TerL) from all representative duplodnaviruses. The tree is rooted in the midpoint. The TerL homologs identified in our mirusvirus genome is indicated by red stars. Reference homologs from marine mirusvirus genomes are shown in blue, and those previously curated from the eukaryotic genome survey are shown in teal green. Ultrafast bootstrap support values are indicated at the branch nodes. Tree scale bars represent the number of substitutions per site.

A



A prior study of mirusviral DNApolB gene extracted from eukaryotic genomes
Mirusviral DNApolB gene from reference genomes
★ Mirusviral DNApolB gene from this study

Figure S6: Maximum-likelihood phylogenetic tree of DNApolB. The tree is rooted with the Euk_Alpha sequences. The DNApolB homologs identified in our mirusvirus genomes are indicated by red stars. Reference homologs from marine mirusvirus genomes are shown in blue, and those previously curated from the eukaryotic genome survey are shown in teal green. Branch support was assessed using ultrafast bootstrap approximation, and the scale bars indicate evolutionary distances measured as substitutions per site.

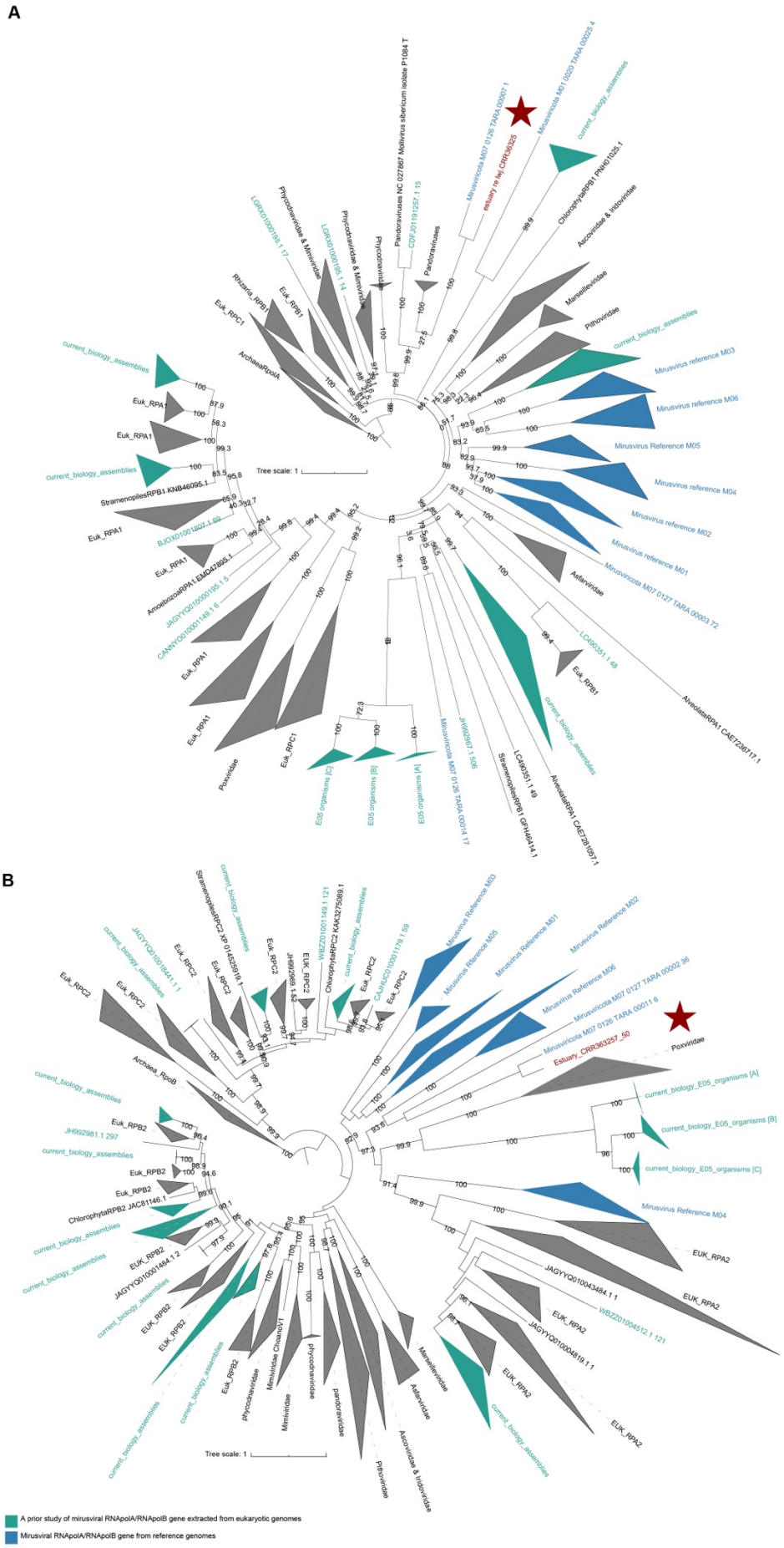

A
Mirusviricota M07 0126 TARA 00007 1
estuary_re fwl_CRR363525
Mirusviricota M01 0020 TARA 00025 4
current_biology_assemblies
ChlorophytaRPB1 PNH01035.1
Ascoviridae & Iridoviridae
Marseilleviridae
Pithoviridae
current_biology_assemblies
Mirusvirus reference M03
Mirusvirus reference M06
Mirusvirus Reference M05
Mirusvirus reference M04
Mirusvirus reference M02
Mirusvirus reference M01
Mirusviricota M07 0127 TARA 00003 72
Asfarviridae
LCA90351.1 48
Euk_RPB1
AlveolataRPA1 CAE72687.1
current_biology_assemblies
LCA90351.1 49
StramenopilesRPB1 GFH46414.1
AlveolataRPA1 CAE72810971
Mirusviricota M07 0126 TARA 00014 17
JH992967.1 506
E05 organisms [A]
E05 organisms [B]
E05 organisms [C]
Euk_RPC1
Poxviridae
Euk_RPA1
Euk_RPA1
CANNYO010001195.1 5
AmoebozoaRPA1 EMD47865.1
Euk_RPA1
BJOX01001802.1 59
StramenopilesRPB1 KNB46095.1
current_biology_assemblies
Euk_RPA1
Euk_RPA1
current_biology_assemblies
Tree scale: 1
ArchaeaRpoA
Euk_RPC1
Rhizaria_RPB1
Euk_RPB1
LGRX01000195.1 17
Phycodnaviridae & Mimiviridae
LGRX01000195.1 14
Phycodnaviridae & Mimiviridae
Phycodnaviridae
Pandoraviruses NC 027867 Mollivirus sibericum isolate P1084 T
CDF.01191257.1 15
Pandoraviruses
B
StramenopilesRPC2 XP 014325919.1
current_biology_assemblies
JAGYYQ01016441.1 1
current_biology_assemblies
Euk_RPC2
Euk_RPC2
Euk_RPC2
Euk_RPC2
JH992968.1 62
Euk_RPC2
WBZZ01001149.1 121
ChlorophytaRPC2 KAK3275089.1
current_biology_assemblies
Euk_RPC2
CAJHJD010001179.1 59
Euk_RPC2
Euk_RPC2
Euk_RPC2
Archaea_RpoB
Mirusvirus Reference M03
Mirusvirus Reference M05
Mirusvirus Reference M01
Mirusvirus Reference M02
Mirusvirus Reference M06
Mirusviricota M07 0127 TARA 00002 36
Mirusviricota M07 0126 TARA 00011 6
Estuary_CRR363257_50
Poxviridae
current_biology_assemblies
Euk_RPB2
JH992981.1 297
current_biology_assemblies
Euk_RPB2
Euk_RPB2
ChlorophytaRPB2 JAC81146.1
current_biology_assemblies
current_biology_assemblies
EUK_RPB2
JAGYYQ01000484.1 2
current_biology_assemblies
EUK_RPB2
EUK_RPB2
current_biology_assemblies
current_biology_assemblies
Euk_RPB2
phycodnaviridae
Mimiviridae Choanovirus
Mimiviridae
phycodnaviridae
pandoraviridae
Marseilleviridae
Asfarviridae
Ascoviridae & Iridoviridae
Pithoviridae
current_biology_assemblies
current_biology_assemblies
current_biology_E05_organisms [A]
current_biology_E05_organisms [B]
current_biology_E05_organisms [C]
Mirusvirus Reference M04
EUK_RPA2
EUK_RPA2
JAGYYQ01004484.1 1
WBZZ01004512.1 121
EUK_RPA2
JAGYYQ01004818.1 1
EUK_RPA2
EUK_RPA2
current_biology_assemblies
Tree scale: 1
A prior study of mirusviral RNApolA/RNApolB gene extracted from eukaryotic genomes
Mirusviral RNApolA/RNApolB gene from reference genomes
Mirusviral RNApolA/RNApolB gene from this study

Figure S7: Maximum-likelihood phylogenetic tree of RNApolA/RNApolB. (**A**) Phylogenetic tree of RNApolA. The tree is rooted with the archaeal RNApolA sequences. The RNApolA homolog identified in our mirusvirus genome is indicated by a red star. Reference homologs from marine mirusvirus genomes are shown in blue, and those previously curated from the eukaryotic genome survey are shown in teal green. (**B**) Phylogenetic tree of RNApolB. The tree is rooted with the archaeal RNApolB sequences. The RNApolB homolog identified in our mirusvirus genome is indicated by a red star. Reference homologs from marine mirusvirus genomes are shown in blue, and those previously curated from the eukaryotic genome survey are shown in teal green. Support values were calculated using ultrafast bootstrap replicates and are shown at each node. Scale bars denote the expected number of substitutions per site.

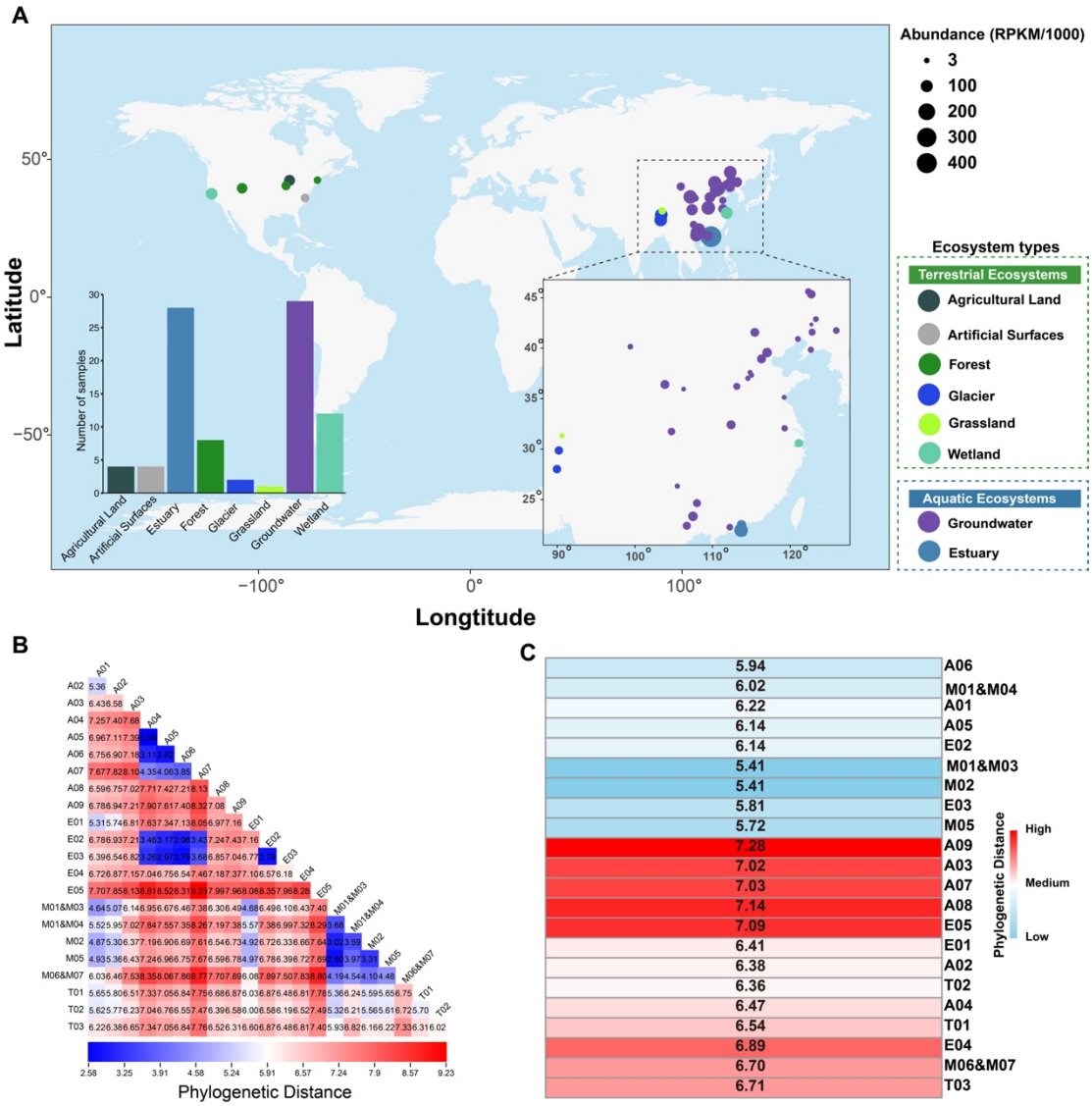


Figure S8: Global distributions and phylogenetic distance of mirusviruses. (**A**) Geographic distribution and abundance of all mirusviruses. The size of bubble represented the abundance value. (**B**) Pairwise phylogenetic distances between mirusvirus lineages. To further explore the evolutionary relationships among mirusvirus clades, we conducted pairwise comparisons of phylogenetic distance between clades. (**C**) Phylogenetic distance of each mirusvirus lineage from the midpoint of the phylogenetic tree.

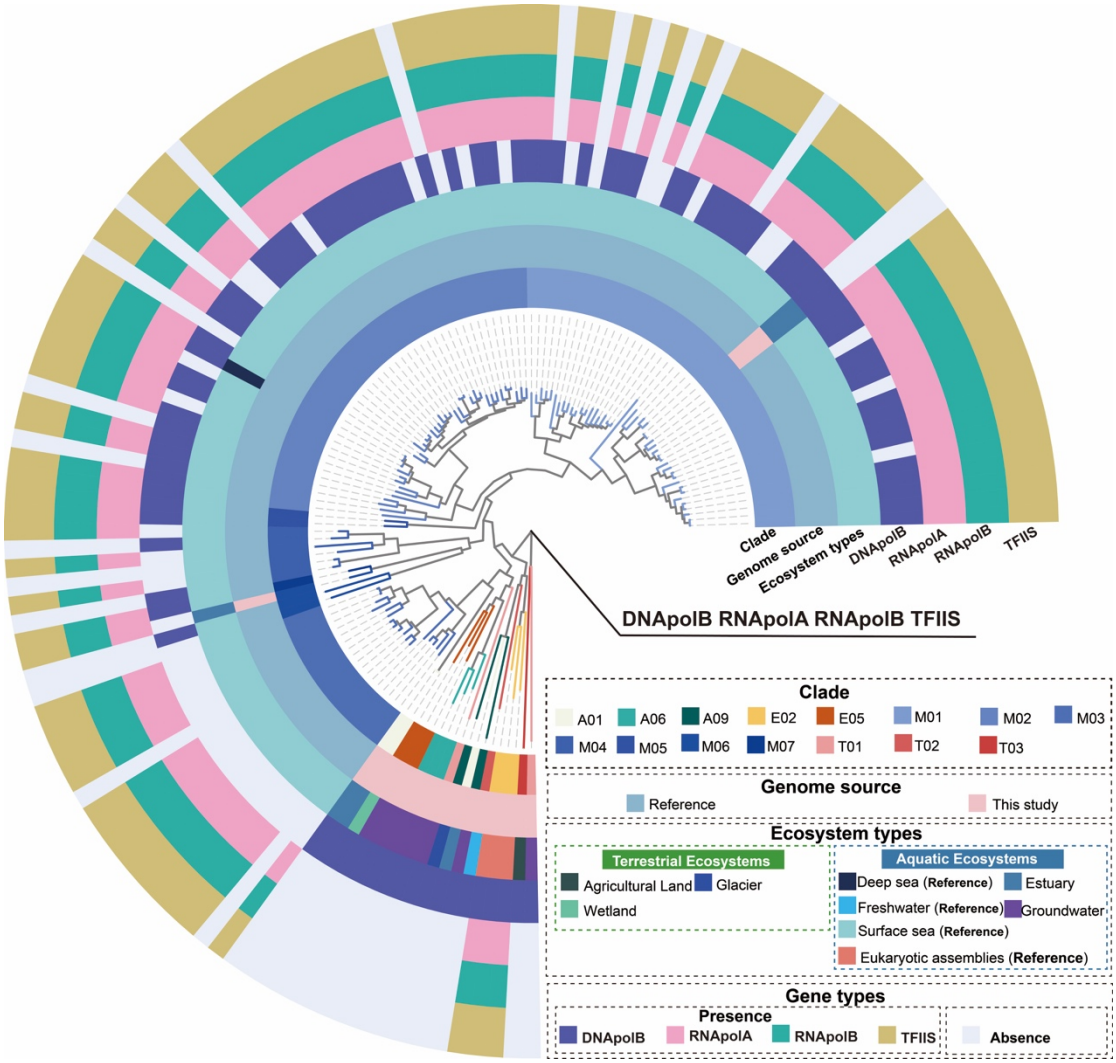


Figure S9: Maximum-likelihood phylogenetic tree of mirusvirus genomes inferred from a concatenated alignment of amino acid sequences of DNApolB, RNApolA, RNApolB, and TFIIS genes. Phylogenetic reconstruction was performed with IQ-TREE using the LG+I+G substitution model and 1000 ultrafast bootstrap replicates, and the tree was midpoint-rooted. The tree was annotated with concentric rings of complementary information and visualized with the ggtreeExtra package.

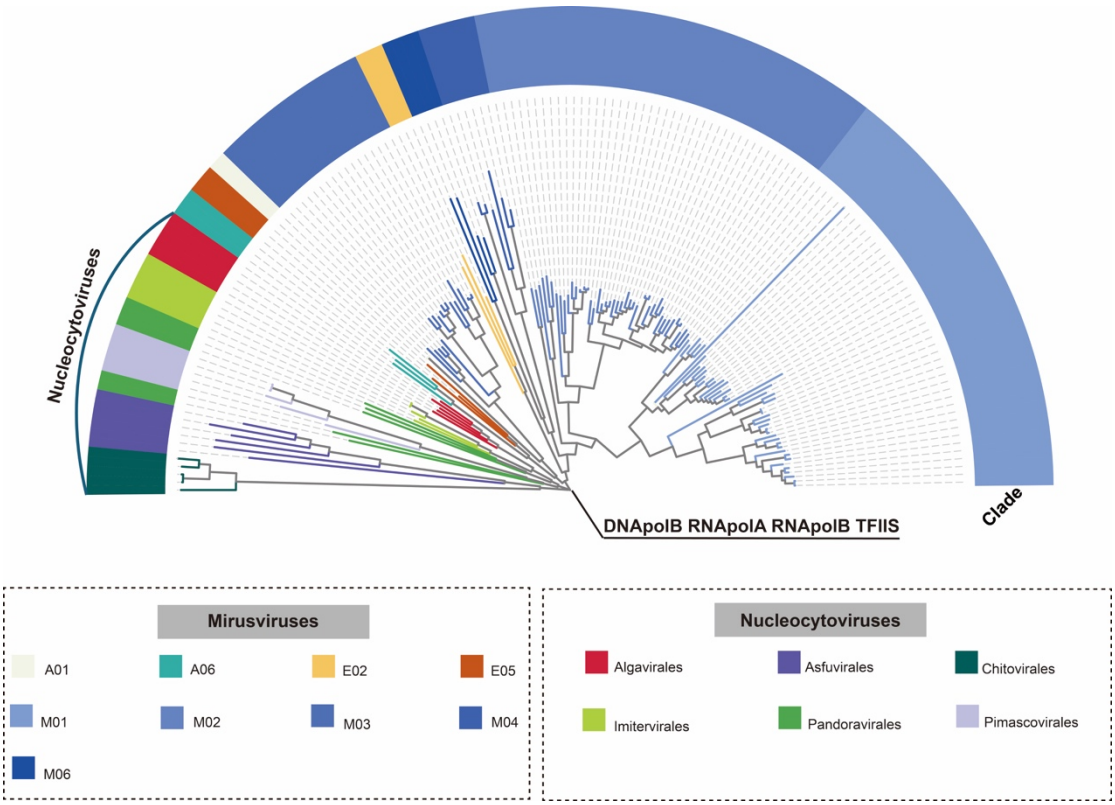


Figure S10: Maximum-likelihood phylogenetic tree built from mirusvirus and nucleocytovirus genomes on the basis of a concatenation of manually curated DNApolB, RNApolA, RNApolB, and TFIIS genes. Phylogenetic reconstruction was performed with IQ-TREE using the Q.PFAM+F+I+R8 substitution model and 1000 ultrafast bootstrap replicates, and the tree was midpoint-rooted.

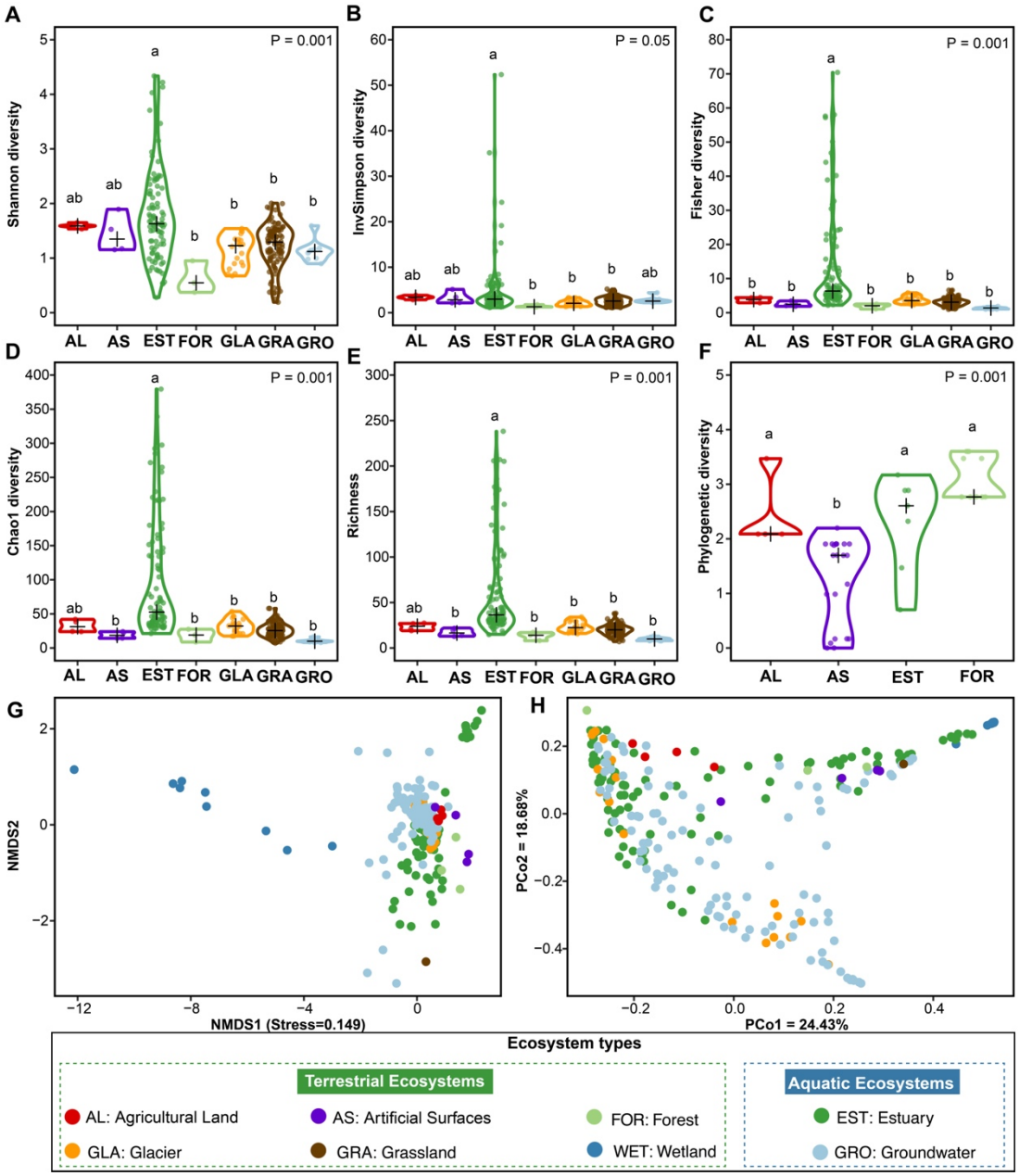


Figure S11: The diversity of the mirusviruses in different ecosystems. (**A-E**) Alpha diversity (Shannon, InvSimpson, Fisher, Chao1, and Richness index) of mirusvirus communities among different ecosystems. (**F**) Phylogenetic diversity of mirusviruses among different ecosystems. Statistical significance was assessed using one-way ANOVA followed by least significant difference (LSD) post hoc tests. Distinct lowercase letters denote significant differences at $\alpha = 0.05$. (**G-H**) Mirusviruses non-metric multidimensional scaling (NMDS) analysis and principal coordinates analysis (PCoA) based on the Bray-Curtis distance among different ecosystems.

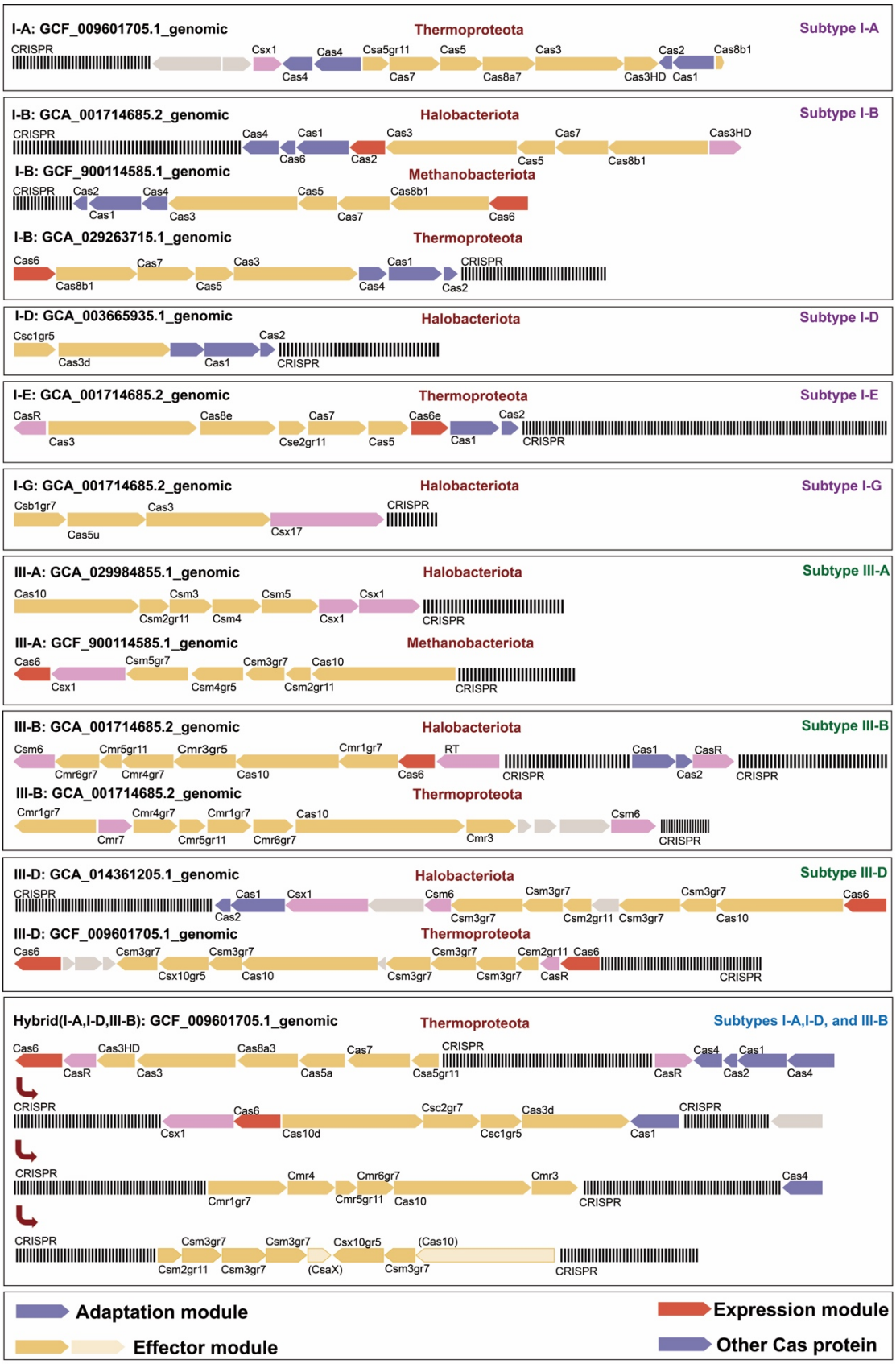


Figure S12: Genetic arrangements of CRISPR–Cas systems in Archaeal hosts of mirusviruses. The genetic arrangements of each CRISPR–Cas subtype identified in

archaeal hosts of mirusviruses are illustrated. CRISPR-Cas subtypes belonging to class 1 are highlighted in purple, while those in class 3 are highlighted in green, as indicated in the top-right corner of each frame. Specifically, we found type I CRISPR–Cas systems in 5 archaeal genomes from three phyla: Halobacteriota (subtype I-B, I-D, and I-G), Methanobacteriota (subtype I-B), and Thermoproteota (subtype I-A, I-B, and I-E). Similarly, Type III CRISPR–Cas systems were detected in 6 archaeal genomes from three phyla: Halobacteriota (subtype III-A, III-B, and III-D), Methanobacteriota (subtype III-A), and Thermoproteota (subtype III-B, and III-D).