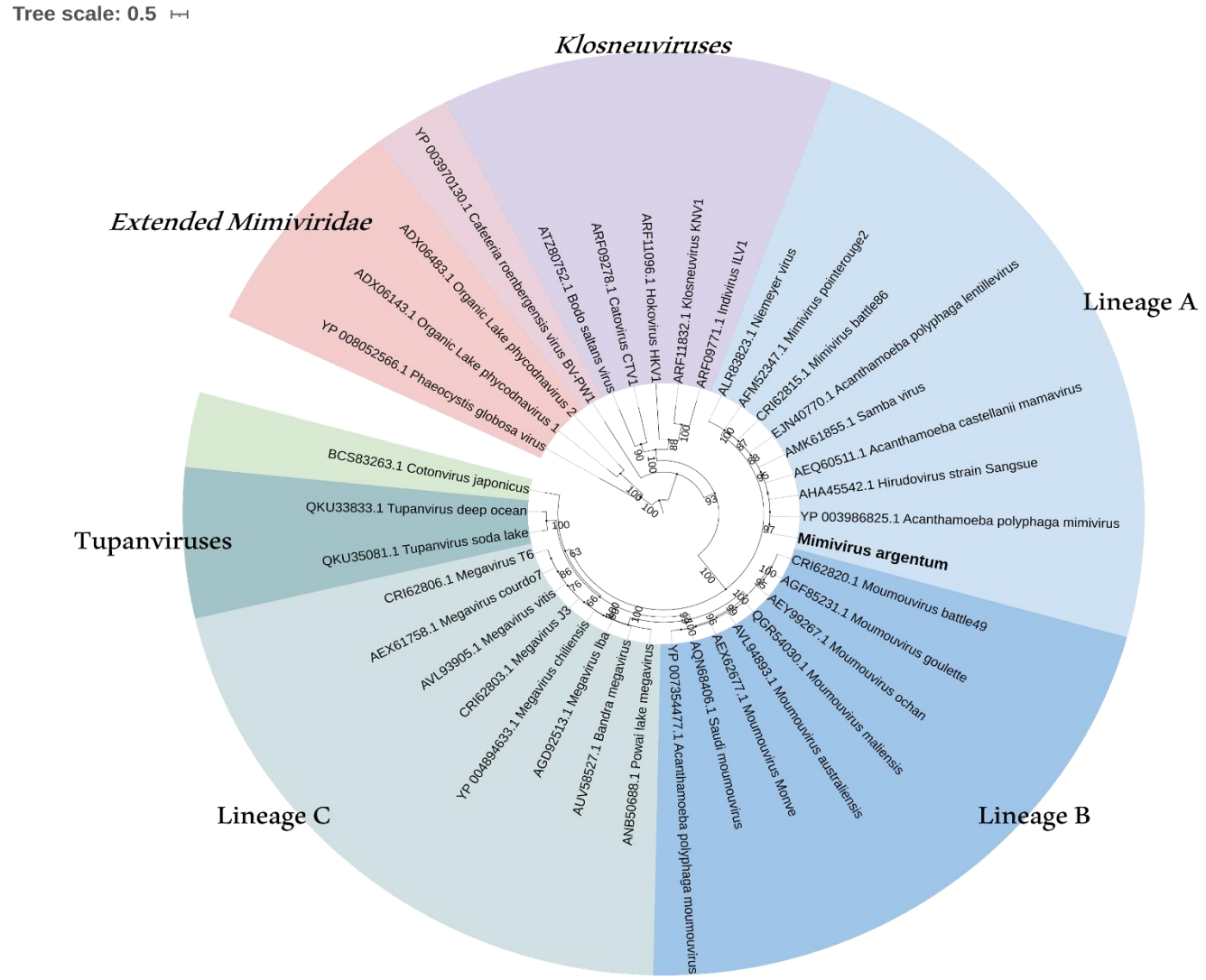
A


Tree scale: 0.5
Klosneuviruses
Extended Mimiviridae
Lineage A
Tupanviruses
Lineage C
Lineage B
YP 003970130.1 Cafeteria roenbergensis virus BV-PW1
ADX06483.1 Organic Lake phycodnavirus 2
ADX06143.1 Organic Lake phycodnavirus 1
YP 008052566.1 Phaeocystis globosa virus
ATZ80752.1 Bodo saltans virus
ARF09278.1 Catovirus CTV1
ARF11096.1 Hokovirus HKV1
ARF11832.1 Klosneuvirus KNV1
ARF09771.1 Indivirus ILV1
ALR83823.1 Niemeyer virus
AFM52347.1 Mimivirus pointerouge2
CRI62815.1 Mimivirus bat1e86
EJN40770.1 Acanthamoeba polyphaga lentillevirus
AMK61855.1 Samba virus
AEQ60511.1 Acanthamoeba castellanii mamavirus
AHA45542.1 Hirudovirus strain Sangsue
YP 003986825.1 Acanthamoeba polyphaga mimivirus
Mimivirus argentum
CRI62820.1 Moumouvirus bat1e49
AGF85231.1 Moumouvirus goulette
AEY99267.1 Moumouvirus ochan
QGR54030.1 Moumouvirus mallensis
AVL94893.1 Moumouvirus australiensis
AEX62677.1 Moumouvirus Monve
AQN68406.1 Saudi moumouvirus
YP 007354477.1 Acanthamoeba polyphaga moumouvirus
ANB50688.1 Powai lake megavirus
AUV58527.1 Bandra megavirus
AGD92513.1 Megavirus lba
YP 004894633.1 Megavirus chiliensis
CRI62803.1 Megavirus J3
AVL93905.1 Megavirus vitis
AEX61758.1 Megavirus courdo7
CRI62806.1 Megavirus T6
QKU35081.1 Tupanvirus soda lake
QKU33833.1 Tupanvirus deep ocean
BCS83263.1 Cotonvirus japonicus

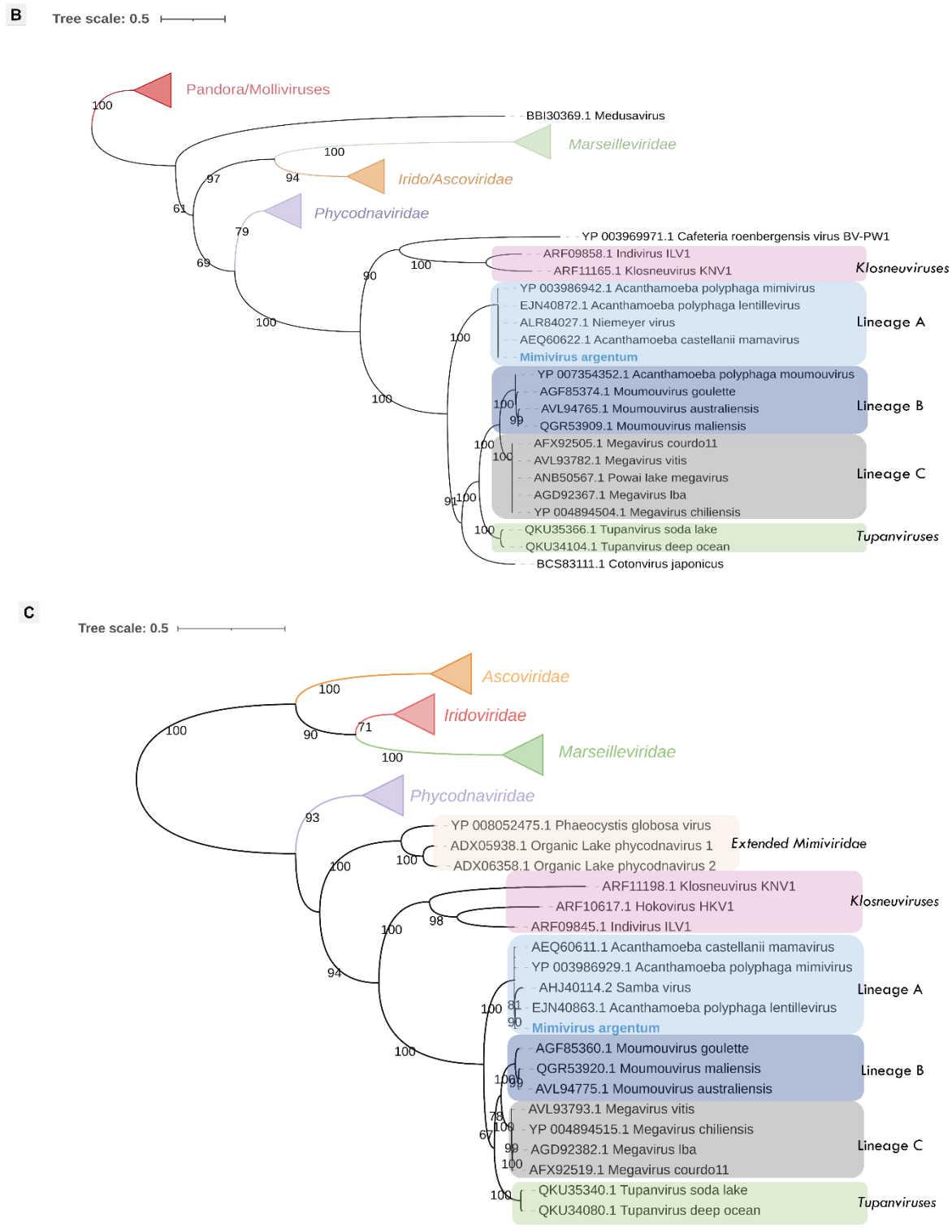


Figure S1. Phylogeny of mimivirus argentum using different data sets. **(A)** *Mimiviridae* maximum-likelihood tree based on DNA polymerase B (ORF678) amino acid sequence. **(B-C)** Maximum-likelihood trees based on VV A32 virion packaging ATPase (ORF539) and major capsid protein (ORF552), respectively. The best-fit models chosen with ModelFinder (implemented in IQtree) for trees A, B and C were LG+F+I+G4, LG+F+I+G4, and VT+F+I+G4 respectively. Mimivirus argentum sequences are labeled with bold and/or blue font. Scale bars represent the number of amino acid substitutions per site.

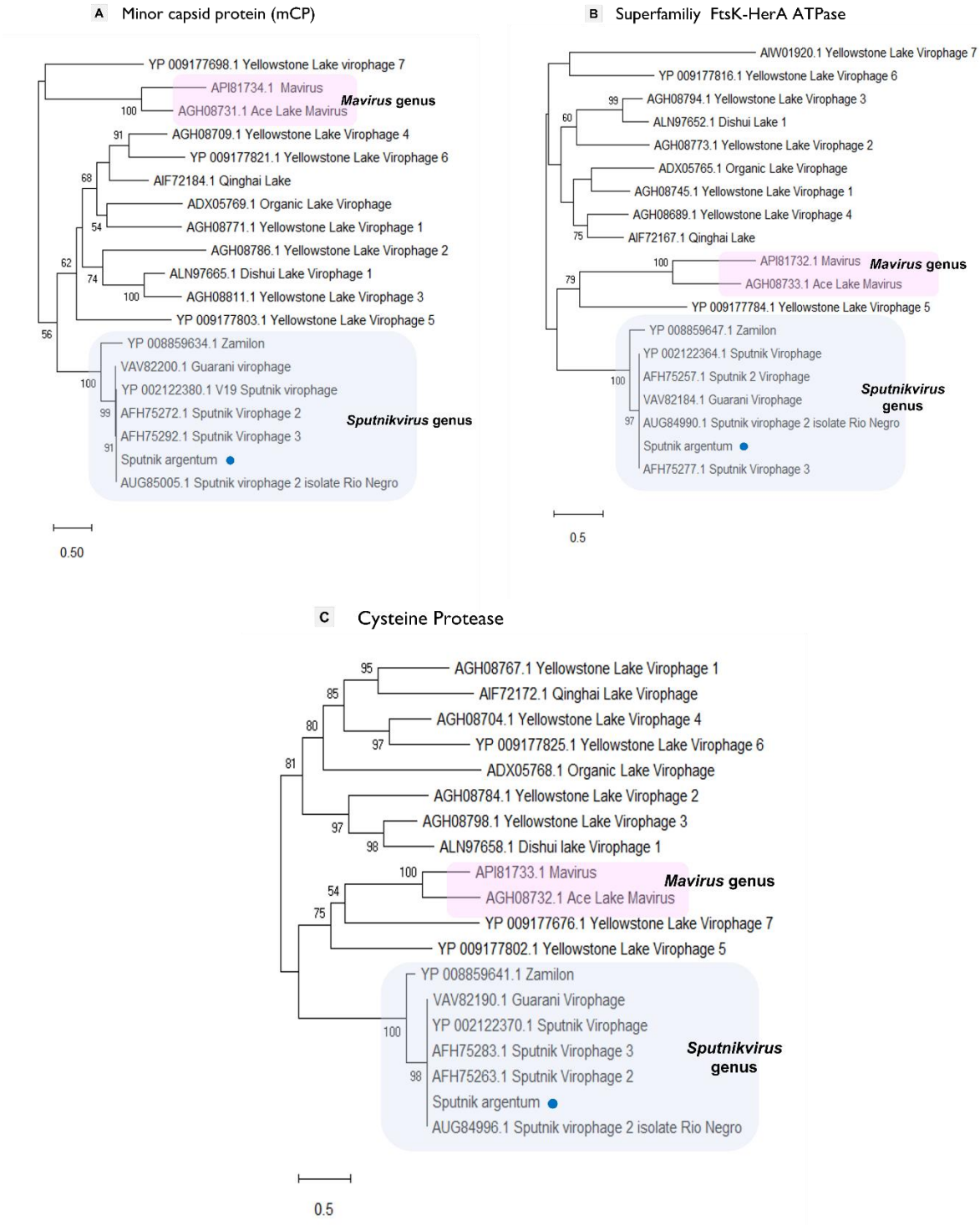


Figure S2. Phylogeny of three virophage core proteins. **(A-C)** Maximum-likelihood trees based on minor capsid "penton" protein (mCP) (ORF18), superfamily FtsK-HerA ATPase (ORF3) and putative cysteine protease (ORF9) amino acid sequences. The best-fit models chosen with ModelFinder (implemented in IQtree) for trees A, B and C were T+F+I+G4; rtREV+F+I+G4, and LG+F+R3, respectively. Sputnik argentum sequences are indicated with the blue circle. Scale bars represent the number of amino acid substitutions per site.

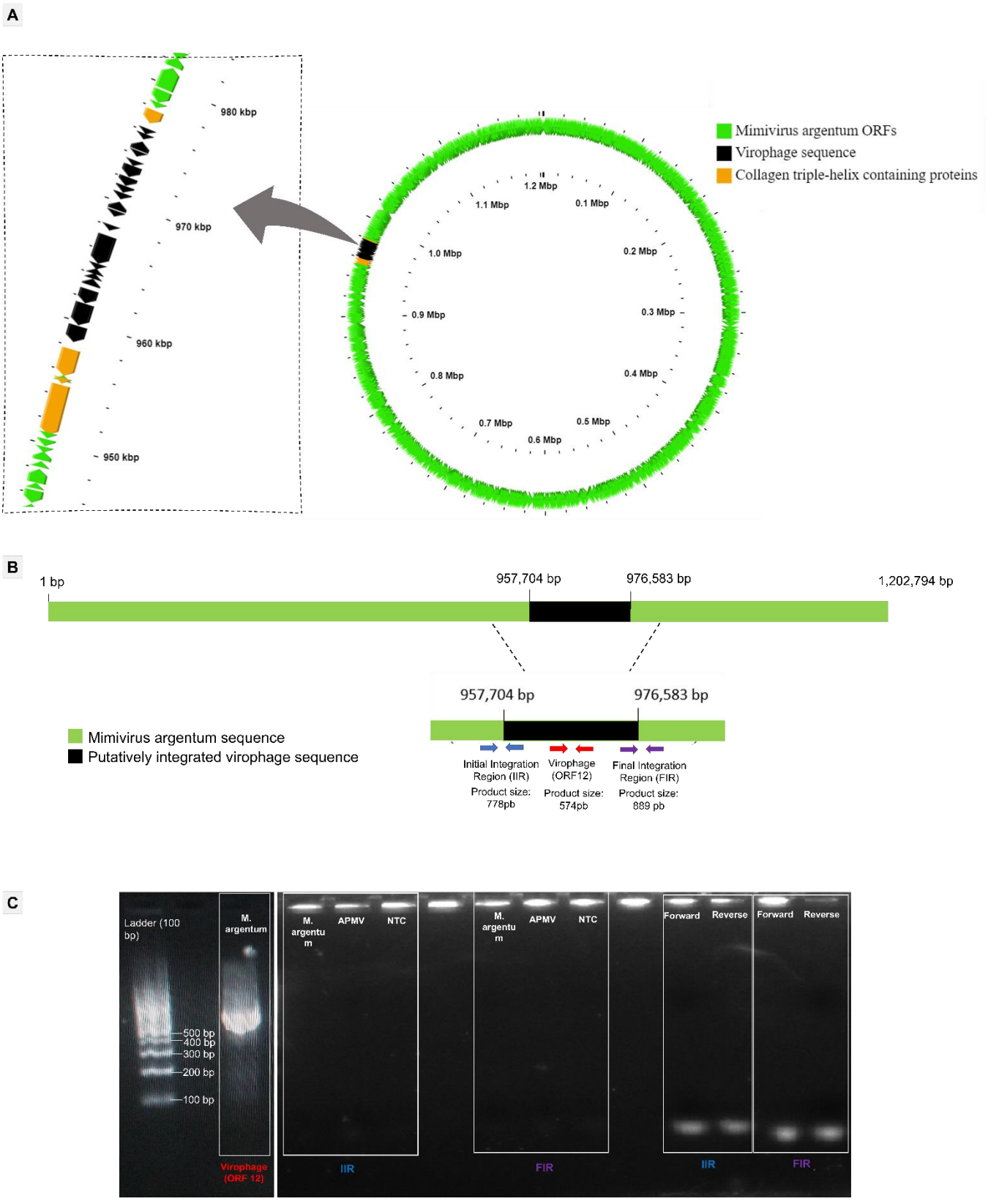


Figure S3. Detection of a provirophage associated with the mimivirus argentum genome only during bioinformatics analysis. **(A)** Representative map of the provirophage putative integration site in mimivirus argentum genome. **(B)** Scheme illustrating the mimivirus argentum genome region used as target in PCR analysis. The green block represents the mimivirus argentum genome, whereas the black block represents the putatively integrated virophage sequence. Arrows represent the targets used for primer design. Blue arrows indicate primers for the Initial Integration Region (IIR), purple arrows indicate primers for the Final

Integration Region (FIR) and red arrows indicate primers targeting ORF12 encoded by sputnik argentum. Right-directed arrows indicate forward primers and the left-directed ones indicate reverse primers. **(C)** Agarose electrophoresis gel after the PCR reaction using DNA extracted from the mimivirus-virophage containing sample. It is possible to observe an amplification with the primer targeted to the virophage's ORF12 and no amplifications with primers targeting both ends of the putative virophage integration site (IIR and FIR), suggesting the absence of a provirophage. The sample with diluted primer pairs (IIR and FIR) used in the PCR reaction was added to the gel as a reaction control, as observed in the amplifications at the bottom right. DNA extracted from Acanthamoeba polyphaga mimivirus (APMV) purified particles (known negative sample for provirophage) was used as the negative control. NTC: no template control.

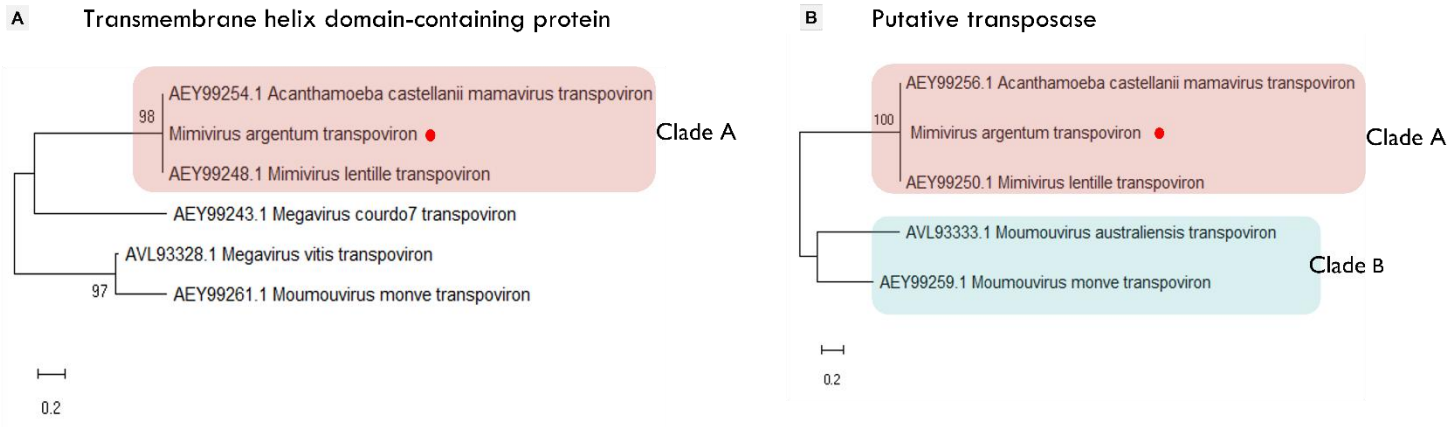


Figure S4. Phylogenetic trees based on two transpoviron proteins. **(A-B)** Maximum-likelihood based on mimivirus argentum-associated transpoviron transmembrane helix domain-containing protein (ORF2) and putative transposase (ORF4), respectively. The best-fit models chosen with ModelFinder (implemented in IQtree) for trees A and B were VT+F and Blosum62+F, respectively. Mimivirus argentum transpoviron sequences are labeled with a red circle. Scale bars represent the number of amino acid substitutions per site.