

Refining a giant virus lineage: a novel order unifying *Mamonoviridae* and “Manesviridae,” unveiled by the discovery of furtivovirus

Jiwan Bae,¹ Masaharu Takemura¹

AUTHOR AFFILIATION See affiliation list on p. 13.

ABSTRACT The evolutionary origins and taxonomic framework of giant viruses related to the family *Mamonoviridae* and its relative group, including clandestinovirus, remain unclassified due to gaps in genome size and host range between these two groups. This study aimed to address this gap by integrating our newly isolated virus with publicly available metagenome-assembled genomes (MAGs) to construct a more robust phylogenetic framework. Here, we report the isolation and characterization of a new giant virus, furtivovirus, using the unicellular amoeba *Vermamoeba vermiformis* as a host. Furtivovirus has a genome of approximately 560 kbp and shares key features with its closest relative, clandestinovirus. Ultrastructural analysis revealed a unique host-nucleus-dependent replication strategy characterized by the breakdown of the nuclear membrane and the packaging of nascent virions directly within the nucleoplasm, distinguishing it from canonical cytoplasmic virion factories. Comprehensive phylogenetic and comparative genomic analyses of shared orthologous groups and nucleocyto-virus marker proteins revealed that furtivovirus, clandestinovirus, ushikuvirus, and usurpativirus form a distinct monophyletic clade, for which we propose a new family, “Manesviridae.” Further analysis using amino acid-based similarity metrics of *Nucleocyto-viricota* viral genomes, including established MAGs, demonstrated that this new family is robustly placed as a sister group to the family *Mamonoviridae*. This study elucidated the evolutionary relationships between viruses with large and small genomes that possess similar virion sizes within this lineage. Based on this cumulative evidence, we propose the establishment of a new order to unify these two families, thereby expanding their diversity and clarifying the evolutionary history of this branch within *Nucleocyto-viricota*.

IMPORTANCE Giant viruses challenge our traditional understanding of viral evolution, raising the question of how a single related group can diverge to infect different hosts while evolving into vastly different genome sizes and replication strategies. The family *Mamonoviridae* and its relatives epitomize this evolutionary divergence: one group possesses massive genomes, whereas the other has genomes that are less than half their size. The discovery of furtivovirus and its unique nucleoplasm-dependent replication cycle provides a critical biological context for this genomic disparity. Through deep comparative genomic analysis, we demonstrated that these seemingly disparate lineages share a cohesive evolutionary origin that is distinct from other established orders. This finding highlights the complexity of genome evolution, demonstrating that giant viruses can expand their overall genome size to adapt to uncertain environments while reducing their core essential genes, thereby providing new insights into the evolutionary pressures that shape the diversity of the virosphere.

KEYWORDS virus isolation, phylogenomics, phylum *Nucleocyto-viricota*, classification, family *Mamonoviridae*, clandestinovirus, furtivovirus, nucleus

Editor Kristin N. Parent, Michigan State University, East Lansing, Michigan, USA

Address correspondence to Masaharu Takemura, giantvirus@rs.tus.ac.jp.

The authors declare no conflict of interest.

See the funding table on p. 14.

Received 26 November 2025

Accepted 15 April 2026

Published 14 May 2026

Copyright © 2026 Bae and Takemura. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Since the first isolation of *Acanthamoeba polyphaga* mimivirus in 2003 (1, 2), numerous amoeba-infecting viruses have been discovered over the past 20 years, most of which are referred to as nucleocytoplasmic large DNA viruses (NCLDV), currently classified under the phylum *Nucleocytoviricota* (3). Initially proposed in 2001, this group was described as comprising large eukaryotic DNA viruses, including the following four viral families: *Poxviridae*, *Asfarviridae*, *Iridoviridae*, and *Phycodnaviridae* (4). Based on current knowledge, nucleocytoviruses are double-stranded DNA viruses that infect eukaryotes, with genome sizes ranging from approximately 45 kilobases (kb) to more than 2.7 megabases, encoding fewer than 100 to over 1,000 genes, and virion sizes ranging from 80 nm to 1.5 μ m (5–7). Furthermore, some amoeba-infecting viruses have been reported to exceed 2 μ m in individual particle size (8). Despite this immense diversity, NCLDVs possess a set of characteristic core genes. The identification of these core genes, such as nucleocytoplasmic virus orthologous groups (NCVOGs) and giant virus orthologous groups (GVOGs), alongside the continuous integration of publicly available metagenome-assembled genomes (MAGs), is crucial for comprehensively charting their evolutionary history (9–11). As of 2025, after more than two decades of research, this group has expanded into a diverse phylum organized into 3 classes, 5 orders, 1 suborder, 15 families, 9 subfamilies, 60 genera, and 134 species. Despite these taxonomic advancements, significant gaps remain, particularly regarding the family *Mamonoviridae* and its close, yet unclassified, relatives. *Acanthamoeba castellanii* medusavirus J1, isolated from a hot spring in Japan in 2019 (12), and medusavirus steno T3, isolated from a freshwater environment in Japan in 2021 (13), were classified into the family *Mamonoviridae* in 2023 (14). Following this classification, medusavirus euryale F10 was isolated from freshwater in South Korea (15), demonstrating the diversity of the genus *Medusavirus*. Medusaviruses exhibit unique features: their genomic replication occurs strictly within the host cell nucleus, and they encode a full set of histone genes, similar to those of eukaryotic organisms (12). Meanwhile, clandestinovirus (CLV), which infects *Vermamoeba vermiformis*, shares a close phylogenetic relationship with medusaviruses, along with the unique feature of encoding a full set of histone genes like medusaviruses, yet it remains unclassified at the family level due to profound differences in genome size and host range (14, 16). Other recent vermamoeba-infecting isolates, such as ushikuvirus (USKV) and usurpativirus (USPV), also fall into this unclassified clade (17–19). Consequently, the family *Mamonoviridae* is currently not assigned to any order in the International Committee on Taxonomy of Viruses (ICTV) classification (14), leaving the systematic position of this entire viral lineage unresolved.

To bridge this taxonomic and evolutionary gap, this study aimed to characterize a newly isolated GV and integrate its genome with established reference sequences and published MAGs to reconstruct a robust phylogenetic framework. Here, we report the isolation of a furtivovirus (FTV) and detail its unique nucleoplasm-dependent replication cycle. By leveraging comparative genomics and orthologous gene networks across isolates and environmental MAGs, we clarified the complex evolutionary relationships of these viruses. Therefore, we propose the establishment of the new family “*Manesviridae*” to encompass FTV and its relatives, and the creation of a novel order unifying “*Manesviridae*” and *Mamonoviridae*.

RESULTS AND DISCUSSION

Isolation and ultrastructural analysis of furtivovirus

Screening tests for GVs infecting *V. vermiformis* resulted in the observation of cytopathic effects (CPEs) following inoculation with a water sample from the Inasegawa River in Kamakura City, Kanagawa Prefecture, Japan (Fig. 1a). At this time, a conventional polymerase chain reaction (PCR) screening was performed on five wells exhibiting CPE using primers targeting the family B DNA polymerase (polB) gene of faustoviruses. This led to the identification of two faustoviruses (faustovirus sigan and faustovirus higan, unpublished) and a newly isolated virus whose particle morphology was similar to that of CLV (16). To characterize the new virus beyond its CPE, we conducted a

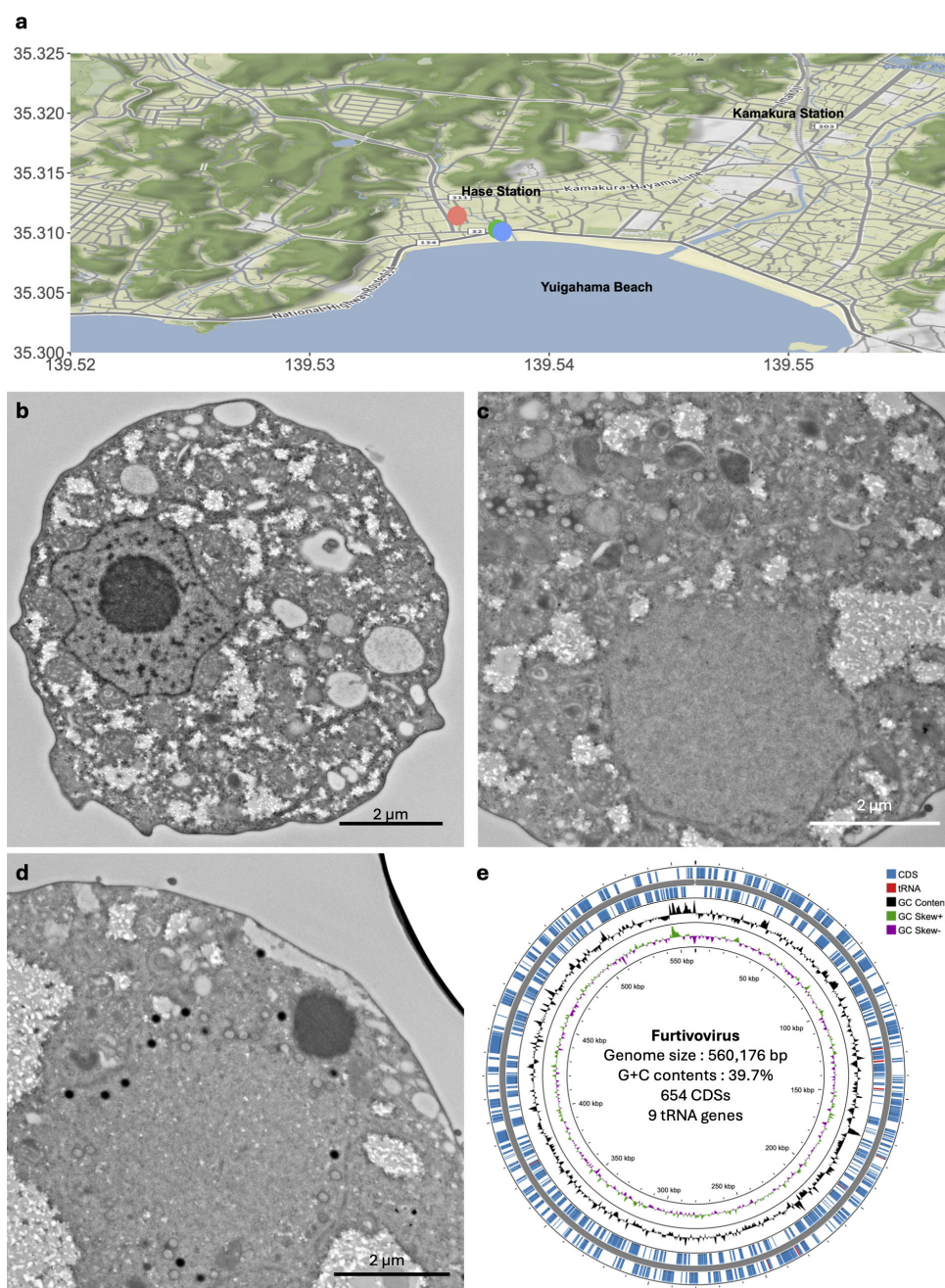


FIG 1 Furtivovirus isolation sites, infected cells, intracellular particle image, and its genome. (a) Kamakura city map showing the collection sites. The red dot represents the locations where a freshwater sample containing furtivovirus was collected. The green and blue dots represent the locations where freshwater samples were collected. A map of the collection area was generated using the “ggmap” package in R via the Stadia Maps API. (b) Cell under FTV-inoculated conditions showing an intact nuclear membrane and visible heterochromatin without visible virions (possible eclipse phase). (c) Infected cell showing empty capsids in the cytoplasm and a vanished host nuclear membrane. (d) Infected cells showing capsids located in the nucleoplasm, including DNA-filled particles. Cells shown in panels b, c, and d were harvested at a multiplicity of infection (MOI) of 10 at 8 h post-infection (hpi). (e) The circular map represents the complete genome and features of furtivovirus.

morphological analysis. We performed conventional transmission electron microscopy (c-TEM) to observe the morphology of intracellular viral particles. The virion had a particle size of approximately 200 nm in diameter and was icosahedral, similar to CLV, USKV, and USPV (16, 17, 19).

Interestingly, our ultrastructural analysis of the infection cycle revealed distinct morphological stages that deviate from canonical giant virus replication (Fig. 1b through d). Under FTV-inoculated conditions, some host cells exhibited an intact nucleus with a clear nuclear membrane and visible heterochromatin, lacking any discernible viral particles in the cytoplasm, potentially representing the eclipse phase (Fig. 1b). However, in actively infected cells, we observed the emergence of empty capsid particles in the cytoplasm, concomitant with the breakdown and vanishing of the host nuclear membrane (Fig. 1c). While the disruption of the nuclear membrane upon virion factory formation has been previously noted in USKV (17), FTV exhibits a far more extraordinary assembly mechanism. As the infection progressed, we observed that capsids were located within the nucleoplasm (somewhat skewed toward the boundary with the cytoplasm, but entirely within the nucleoplasm without overlapping the boundary), including electron-dense, DNA-filled particles (Fig. 1d). Notably, these DNA-filled structures were not observed in the cytoplasm. This sequential observation strongly supports the hypothesis that nascent empty capsids migrate into the area of the disrupted nucleus and are packaged with viral DNA directly within the nucleoplasm. Evidence of virions migrating into the nucleus during the replication cycle of CLV has been previously reported, during which the nuclear membrane was observed to take on an abnormal appearance (16); however, this mode of genome packaging has not been explicitly observed in viruses that share the same host but assemble primarily in independent, well-defined cytoplasmic virion factories (tupanvirus, faustoviruses, and naiavirus) (8, 20, 21) or morphologically similar viruses (CLV and USKV). This finding suggests a profound host-nucleus-dependent replication and assembly strategy for FTV. Despite this unique morphology, it was difficult to fully classify the virus based solely on ultrastructural characteristics; therefore, we conducted a comprehensive whole-genome analysis.

Whole-genome analysis and tRNA repertoire

Whole-genome analysis revealed that this new virus has a linear double-stranded DNA genome of 560,176 base pairs and contains 656 predicted coding sequences (Fig. 1e). Interestingly, FTV encodes a relatively high number of transfer RNAs (nine tRNAs) compared to its close relatives (Table 1). To understand the evolutionary rationale behind this tRNA expansion, we analyzed the anticodons encoded by FTV, CLV, USKV, and USPV in the context of the host's (*V. vermiformis*) codon usage. According to recent host codon usage indices (22), these viral tRNAs do not correspond to extremely rare or unusually frequent host codons, nor do they appear to resolve specific codon mismatches. Instead, they correspond to codons with standard baseline usage frequencies in the host. This suggests that the retention of these tRNAs is not a strict compensatory adaptation for translation bottlenecks, but rather a clade-specific genomic signature whose precise functional benefit during the nucleoplasm-dependent replication cycle warrants further study.

In addition to these genomic features, the GC content of the group of FTV, CLV, USKV, and USPV provides valuable insights when viewed through the lens of recent taxonomic discussions. The GC content of FTV, CLV, USKV, and USPV ranged from 39.72% to 47.88%. Notably, this range closely aligns with the GC content of their shared host, *V. vermiformis* CDC-19 (41.94% [GCA_045837725.1]). Conversely, members of the sister family *Mamonoviridae* (medusaviruses) exhibit a significantly higher GC content ($\approx 60\%$), perfectly mirroring the GC content of their respective host, *Acanthamoeba castellanii* Neff (58.42% [GCA_046055985.1]). Such host-virus sequence composition correlations have been increasingly documented across the *Nucleocytoviricota* landscape, emphasizing that GC content often reflects specific host adaptations and biological cycles rather than strict phylogenetic constraints (23). For instance, giant viruses within the proposed family "Pandoraviridae" also exhibit a profound host-nucleus dependency during their replication cycle and possess a notably high GC content (60%–64%). As highlighted in genomic analyses of other giant viral clades (24), GC content can be highly variable

TABLE 1 Genomic characteristics and tRNA repertoires of furtivovirus and its relatives

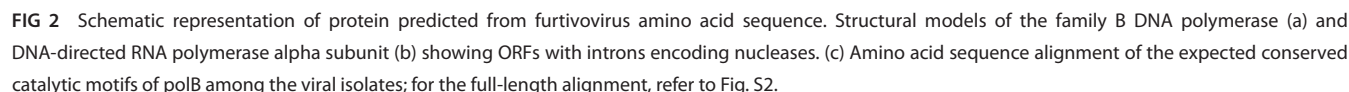
	Furtivovirus	Clandestinovirus	Ushikuvirus	Usurpativirus
Genome size (bp)	560,176	581,987	666,605	669,751
GC content	39.72	43.56	47.87	47.88
Number of CDSs	654	617	784	758
Encoded tRNAs (anticodon)	Leu(aag)	Ser(gct)	Val(cac)	His(gtg)
	Gln(ttg)		His(gtg)	Val(cac)
	Gln(ctg)			
	Tyr(gta)			
	Phe(gaa)			
	Gly(tcc)			
	Cys(gca)			
	Trp(cca)			
	Ser(tga)			

and is not always a reliable, standalone taxonomic marker at the family level. Therefore, the stark divergence in GC content between *Mamonoviridae* and the group of FTV, CLV, USKV, and USPV likely reflects prolonged evolutionary adaptation to their distinct host environments. This host-driven mutational pressure further justifies their separation into two distinct families while still supporting their unification under a higher taxonomic order based on shared core replication mechanisms.

Shifting our focus from these broad genomic signatures to the specific genes driving these mechanisms, we find that the replication and transcription machinery of FTV exhibits notable characteristics. This new viral genome does not encode DNA topoisomerase II, like medusaviruses (12) and other relatives, but it encodes polB (open reading frames [ORF] 129, 131) and DNA-directed RNA polymerase alpha subunit (RNAPL) (ORF 479, 481), with introns similar to those of CLV (16) (Fig. 2a and b). However, unlike CLV, FTV encodes additional genes, including HNH endonuclease and GIY-YIG nuclease, on the complementary strand (Fig. 2a and b). Furthermore, unlike CLV, it encodes five HNH family homing endonucleases and three LAGLIDADG endonucleases, one of which (an HNH family homing endonuclease) showed no homology with CLV protein genes. These endonucleases may have impacted the introns present in the FTV genome (25).

Structural analysis of the family B DNA polymerases

To evaluate the impact of these genomic differences on the core replication machinery, we modeled and compared the 3D structures of the polB using amino acid sequences and structural prediction algorithms AlphaFold3 (Fig. S1 and S2). In this process, we found significant structural differences in regions other than the protein splicing site across the family *Mamonoviridae* and its relatives. When the expected conserved catalytic motifs were mapped onto the primary sequences and tertiary structures (26–28), we observed that the unique secondary structural loops identified in FTV, CLV, USKV, and USPV did not intersect with the core enzymatic domains (Fig. S1 and S2). This suggests that these loops are structural additions rather than modifications of the catalytic activity. Notably, however, distinct changes in the conserved sequences were observed in CLV and FTV, with the changes being more severe in CLV. CLV lacks the YGDTDS motif in Region I and polymerase motif A (SLYPS) in Region II. In polymerase motif B in Region III, the lysine (K) residue in K...X3NS(V)YG was substituted to leucine (L), and several amino acid residues were added to the FTV gene. This motif was completely absent in CLV (Fig. 2c; Fig. S2). These severe structural mutations in the primary replicative enzyme strongly suggest that these genes are unlikely to encode functional enzymes in FTV and CLV. Coupled with the absence of topoisomerase II and the observation that viral particles migrate through the cytoplasm and enter the nucleus during CLV infection (16), this structural evidence corroborates our ultrastructural findings (Fig. 1b through d). Thus, it is highly probable that FTV and its close relatives possess impaired



However, while our *in silico* structural predictions and *in vitro* ultrastructural observations provide compelling evidence for this host dependency, definitive experimental validation remains necessary. Future studies employing reverse genetics and targeted gene-editing techniques, such as knockout or precise mutation of the viral polB gene, will be crucial. Demonstrating whether FTV can successfully complete its replication cycle in the complete absence of its viral polB would explicitly determine whether this gene has degraded into a non-functional pseudogene, or whether the structurally altered protein still retains an essential, albeit non-catalytic, accessory role during nucleoplasmic genome packaging.

Comparative genome analysis of FTV and its relatives

To firmly define the taxonomic boundaries among the seven isolated viruses (three medusaviruses, FTV, CLV, USKV, and USPV), we performed an extensive orthologous group (OG) analysis. First, we visualized the protein-sharing patterns using a bipartite network (Fig. 3a). This network was visibly divided into two clusters: the medusavirus group (family *Mamonoviridae*) and the CLV, FTV, USKV, and USPV group. Quantification revealed that only 29 universal core proteins were shared across all seven isolates (Fig. 3b). To contextualize this, the suborder *Ocovirineae* within the order *Pimascovirales* possesses 56 universal core genes (*Pimascovirales* universal core genes = 4), and the family *Mimiviridae* within the order *Imitervirales* shares 87 universal core genes (*Imitervirales* universal core genes = 4) in the present analysis. The extremely limited core gene set observed among our seven isolates indicates a deep evolutionary divergence that extends beyond the family level. Analysis using the Jaccard index of shared OGs clearly separated the *Mamonoviridae* group from the FTV/relatives group and confirmed the formation of two distinct sub-clusters (FTV-CLV and USKV-USPV) within the latter (Fig. 3c). To further calibrate these findings and extrapolate the taxonomic ranks of our isolates, we evaluated the Jaccard index patterns of the established order *Imitervirales* as a comparative benchmark (Fig. S3). Within *Imitervirales*, when excluding *chrysochromulina parva* virus BQ2—an extreme outlier with uniquely low gene-sharing indices across the entire order—the minimum intra-family Jaccard value observed within the family *Allomimiviridae* is 7 (Fig. S3). Furthermore, within the densely populated family *Mimiviridae*, the boundaries demarcating subfamilies typically span Jaccard values of 15–32, with the lowest internal intra-subfamily minimum reaching 11 (Fig. S3). By applying these quantitative benchmarks to seven isolates, the profound separation between the *Mamonoviridae* group and the FTV/relatives group strongly supports their classification into separate, distinct families. Subsequently, the gene-sharing indices between the two subclusters (FTV-CLV and USKV-USPV) aligned seamlessly with the established

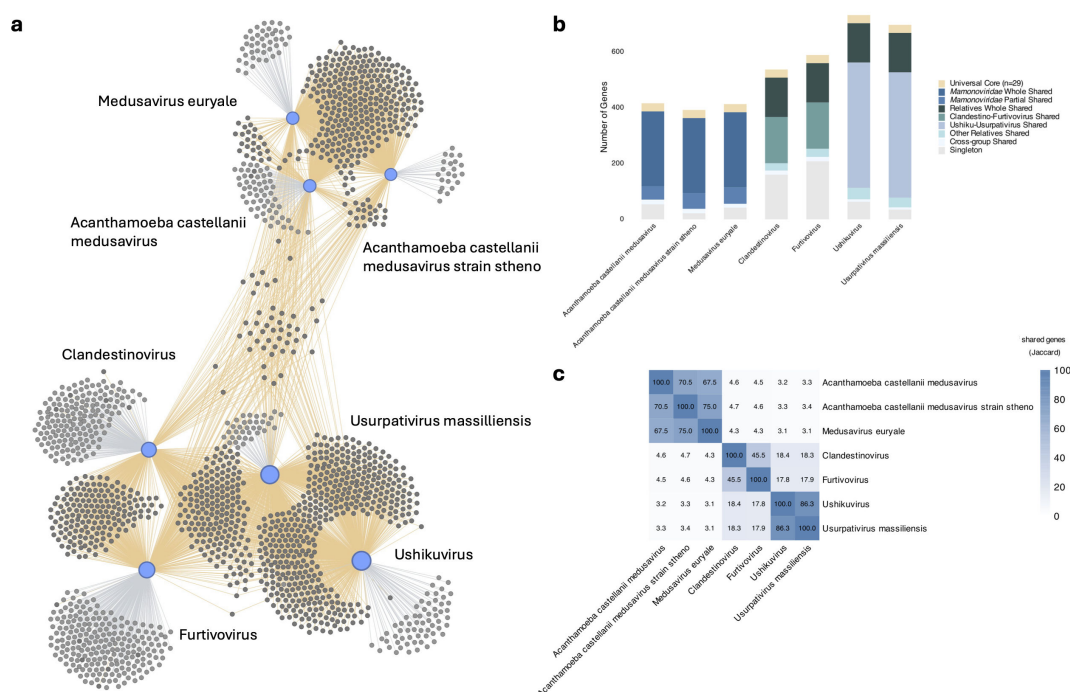


FIG 3 Gene sharing networks and orthologous relationships. (a) Bipartite network visualization of shared proteins among the seven isolates. Each viral genome is represented by a blue node, and the proteins are represented by gray nodes. Singletons are connected to the genome node by gray lines, while shared proteins are connected by wheat-colored lines. (b) Quantification showing the presence of only 29 universal core genes across the group. (c) Jaccard index matrix of shared orthogroups demonstrating distinct clustering between medusaviruses and the FTV/CLV/USKV/USPV group.

parameters for subfamily-level divergence of Mimiviridae (Fig. S3). Finally, the subtle yet persistent differences observed between FTV and CLV within their subclusters (Jaccard values of 45.5) suggest a separation of these two viruses at the genus level (Fig. 3c). Together, these metrics provide a highly resolved hierarchical framework for the taxonomy of this entire lineage.

Because simple ortholog counts can underestimate genetic distance, we evaluated the nucleotide-level identity of the shared genes. We calculated the identity matrix of the shared OGs using a customized pipeline and blastn (Fig. 4a and b). To complement this, we performed a whole-genome average nucleotide identity (wANI) analysis, including non-coding regions (Fig. 4c). The trends of these three metrics were consistent, clearly separating the *Mammonoviridae* and FTV/relatives groups at the family level (Fig. 4). Therefore, we name the FTV/relatives group “Manesviridae.” Furthermore, these metrics showed wANI values below 70% between the clusters within the newly proposed “Manesviridae” family (Fig. 4c), supporting its subdivision into at least three distinct genera: CLV, FTV, and USKV-USPV.

Phylogeny and the challenging of the “Pandoravirales” hypothesis

To position FTV within the broader evolutionary landscape of the phylum *Nucleocy-toviricota*, we conducted a comprehensive phylogenetic analysis using seven concatenated core marker genes extracted from 372 viral genomes (10, 29). Crucially, this data set incorporated published high-quality MAGs (29–33) to test deep evolutionary links (Fig. 5). The resulting unrooted phylogenetic tree robustly placed FTV within a monophyletic clade alongside CLV, USKV, and USPV. Furthermore, an environmental MAG ERX682739.15.fa.dc (ultrafast bootstrap $\geq 99\%$, SH-aLRT $\geq 89.2\%$) and ERX682739.15.fa.dc (ultrafast bootstrap $\geq 73\%$, SH-aLRT $\geq 34.3\%$) clustered dependably within this lineage. This entire assemblage constitutes the newly proposed family “Manesviridae,” which forms an unequivocal sister group to the family *Mammonoviridae* (Fig. 5).

In early taxonomic frameworks, it was hypothesized that medusavirus might fall under the proposed order “Pandoravirales” (10). To evaluate the validity of this

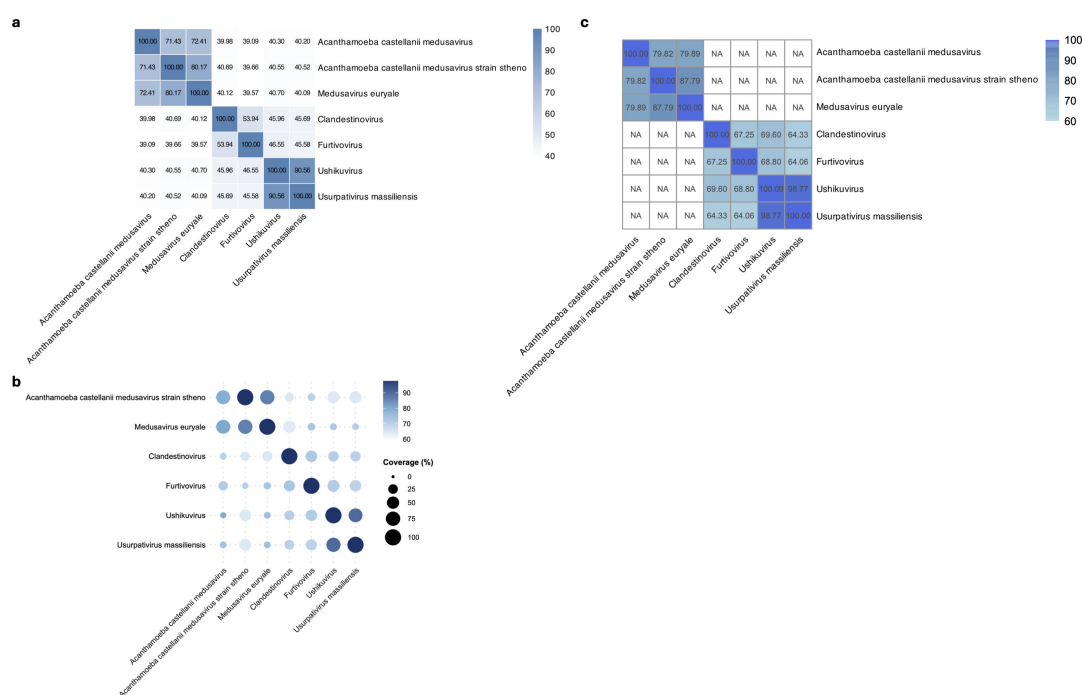


FIG 4 Nucleotide-level comparisons between *Mammonoviridae* and the proposed “Manesviridae.” (a) Gene-based average nucleotide identity (gANI) of shared orthologous genes calculated via a customized pipeline. (b) gANI of shared orthologs utilizing blastn. (c) wANI heatmap.

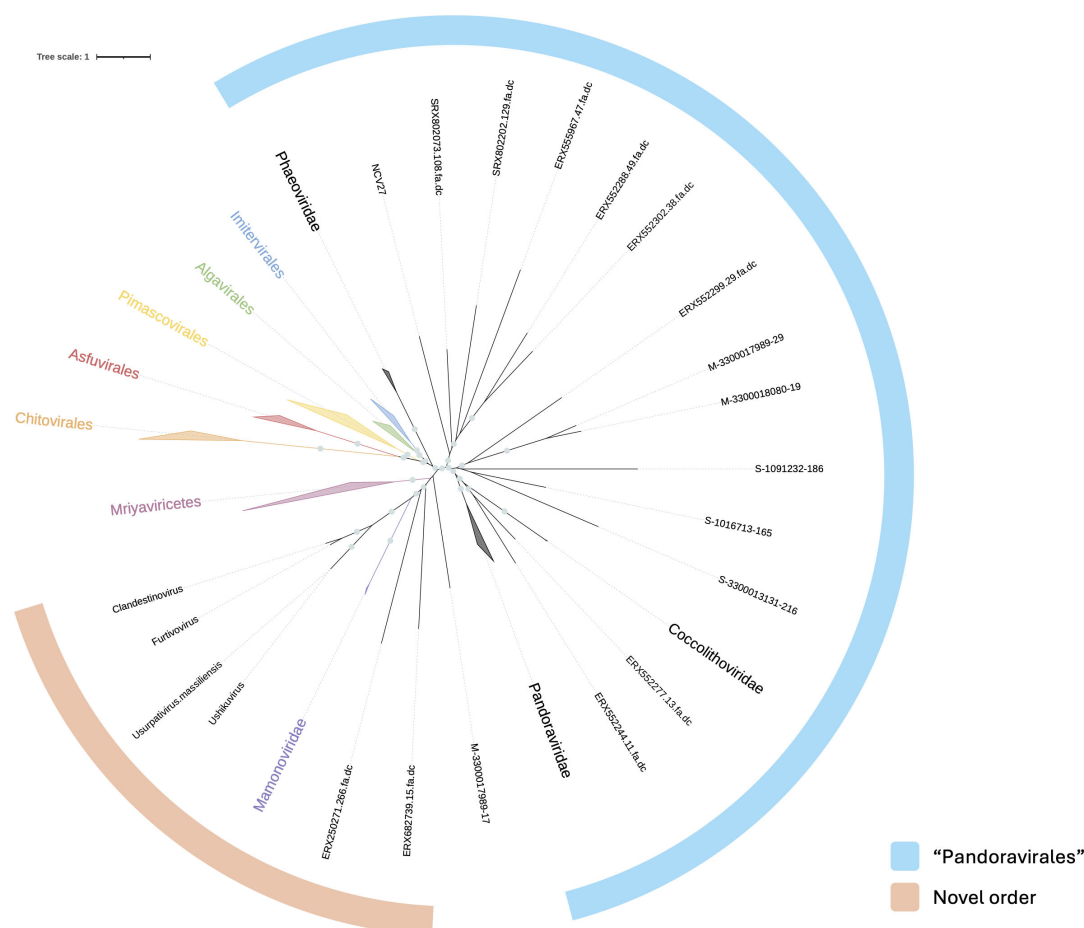


FIG 5 Unrooted phylogenetic tree of 372 *Nucleocyotviricota* genomes (including published MAGs) based on the concatenated amino acid sequence alignment of seven marker genes using IQ-TREE. Nodes are colored according to the ICTV-established orders/classes. Bubbles at nodes indicate ultrafast bootstrap $\geq 95\%$ and SH-aLRT $\geq 80\%$.

hypothesis using our expanded data set, we conducted a targeted comparative analysis of universal core genes. The proposed *Mamono*-“Manes”*viridae* superclade (comprising the seven isolates and the associated MAG ERX682739.15.fa.dc) shared 11 universal core genes related to basic transcription, replication, and repair. Building upon our hierarchical Jaccard index framework, we incorporated ERX682739.15.fa.dc into this matrix confirming its distinct placement (Fig. S4). The observed gene-sharing index robustly positions this environmental MAG at the subfamily level within the *Mamono*-“Manes”*viridae* superclade, reinforcing its genetic cohesion as a discrete yet deeply related evolutionary branch. In contrast, the “Pandoravirales” group independently shares only two universal core genes (A32 packaging ATPase and polB). However, when we merged these two massive groups to test their taxonomic unity, the number of shared core genes collapsed to one (A32 packaging ATPase). Most notably, despite both groups encoding the fundamental replication enzyme polB, they are not grouped as orthologs because of severe sequence divergence. This finding suggests significant genetic discontinuity between these lineages. To further quantify this discontinuity, we assessed the nucleotide-level identity of the shared core genes between the proposed superclade and “Pandoravirales” using a gap-sensitive blastn alignment approach, identical to our intra-family evaluations. The resulting identity matrix revealed a stark separation, distinctly clustering the “Pandoravirales” members away from the proposed novel order (Fig. S5). This group-specific gene-sharing dynamic and pronounced sequence distance

unequivocally indicate that merging the *Mamono*-“*Manes*”*viridae* clade into the order “*Pandoravirales*” is taxonomically inappropriate.

To expand our comparative framework beyond the “*Pandoravirales*” hypothesis and validate the evolutionary distances across the entire phylum, we assessed the amino acid similarity of the seven core marker genes across all established orders using a BLOSUM62-based identity metric (Fig. 6). This approach accounts for deep evolutionary conservation, which nucleotide-level comparisons may obscure. The inter-order comparison (Fig. 6a) definitively confirmed that the newly proposed lineage shares a consistently low baseline similarity with all other established orders (including the proposed order “*Pandoravirales*”). Meanwhile, the internal diversity of the nine genomes (seven isolates + two MAGs) comprising our proposed group (Fig. 6b) was consistent with the level of variation accepted within other established orders. The heatmap summarizing these relationships (Fig. 6c) provides strong quantitative evidence for establishing a distinct novel order encompassing the sister families *Mamonoviridae* and “*Manesviridae*.” Additionally, we analyzed relative evolutionary divergence (RED) values; however, this approach did not provide statistically significant support, limiting our ability to convincingly distinguish the proposed *Mamonoviridae* and “*Manesviridae*” based on evolutionary rates. This is likely because conclusively demonstrating a difference in evolutionary rates between the two families requires more genome data than is currently available to achieve sufficient statistical power.

Conclusions

In this study, we present a comprehensive characterization of a new virus, FTV, bridging the evolutionary gap in the taxonomy of the phylum *Nucleocytoviricota*. Our ultrastructural observations clearly demonstrated a nucleoplasm-dependent genome packaging strategy for FTV, deviating from cytoplasmic virion factories and providing context for the loss-of-function structural mutations observed in its primary replication machinery. Furthermore, by integrating genomic data with existing environmental MAGs and applying multi-metric comparative genomics, we academically justified the proposal of the new family, “*Manesviridae*.” Crucially, the core-gene analysis results challenge the previous taxonomic hypothesis that placed this lineage within “*Pandoravirales*.” When these groups are merged and analyzed as a single taxon, the number of shared core genes plummets to one, and despite both groups encoding polB, they fail to group as orthologs because of sequence divergence, strongly suggesting a deep genetic discontinuity between them. Meanwhile, separating “*Manesviridae*” and *Mamonoviridae* into distinct families is firmly justified by the stark contrast in genome size (an approximately twofold gap), the extremely limited shared core genes (only 29), low nucleotide sequence identity comparable to distances between other established viral families, and severe structural divergence of the primary replication machinery. Furthermore, unifying these two families into a new single order is supported by their shared host-nucleus dependency and genetic cohesion within the lineage, which is comparable to other established viral orders. The stark evolutionary contrast observed within this newly proposed order—from the highly reduced, nucleus-dependent genomes of *Mamonoviridae* to the massive genomes of “*Manesviridae*”—provides a compelling real-world example of the dynamic, accordion-like evolution of GVs (34, 35). While some giant virus lineages expand their genomic robustness to adapt to uncertain environments (36), FTV and its relatives have adopted a strategy of extreme reductive evolution, shedding core replicative machinery to strictly parasitize the host nucleoplasm. This study not only resolves taxonomic ambiguity but also presents this clade as a useful model system for exploring the host-nucleus dependency and genome evolution mechanisms of giant viruses.

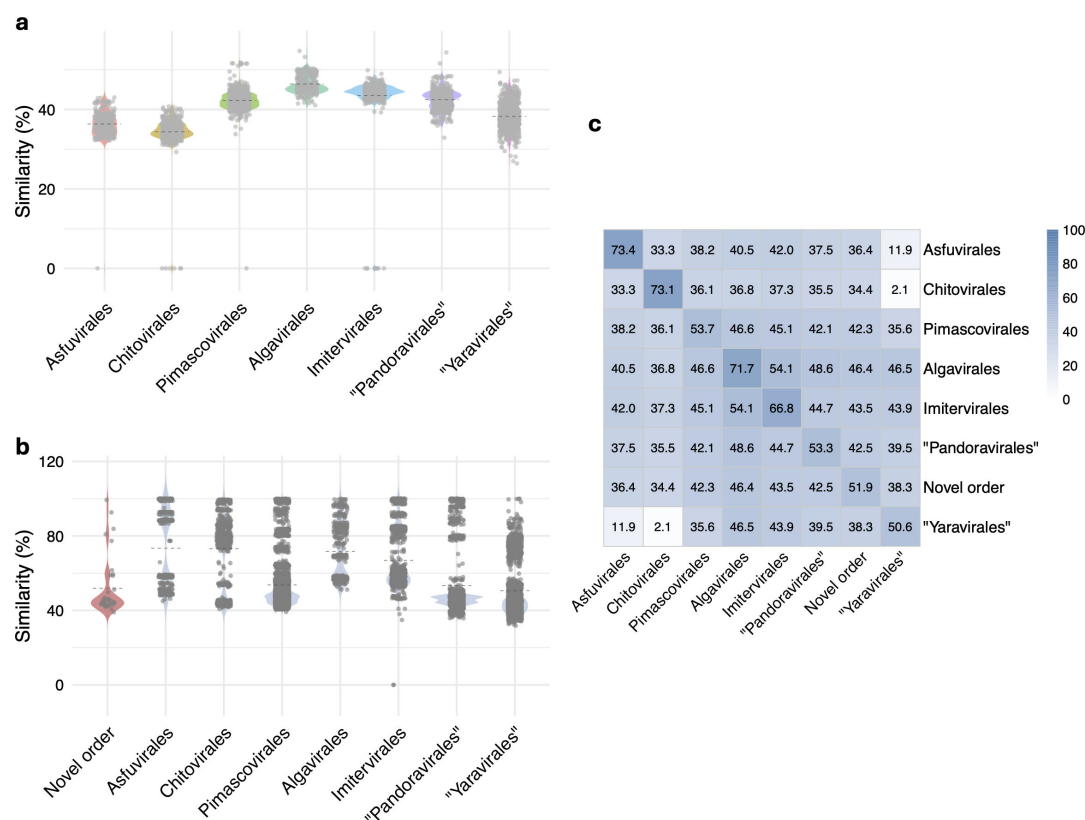


FIG 6 BLOSUM62-based amino acid similarity analysis of the seven marker genes across diverse viral groups. (a) Inter-order relationship demonstrating low similarity between the proposed order and established groups. (b) Intra-order diversity showing that the internal variation of the proposed order is consistent with established orders. (c) Heatmap summarizing the mean similarity values between taxonomic orders.

MATERIALS AND METHODS

Viral isolation and virion production

Fifty milliliters of freshwater samples was collected from three different locations in the Inasegawa River in Kamakura City, Kanagawa Prefecture, Japan. The 1:10 diluted water samples were inoculated and inoculated with *Vermamoeba vermiformis* CDC-19 (ATCC 50237) at 5×10^5 cells per well in 96-well plates containing peptone-yeast extract-glucose medium supplemented with mixed antibiotics, as previously described (17, 37, 38). Plates were incubated at 26°C, and upon observation of CPEs, a conventional PCR screening was performed targeting the polB gene of faustoviruses and ushikuvirus. The supernatants were collected from the corresponding wells. The virions were subsequently cloned and propagated. Viral purification was performed using a 0.45 μ m membrane filter, as previously described (17, 37, 38). From these isolates, we detected viral particles from two faustovirus isolates and a related clandestinovirus (CLV)-like virus, which were identified by whole-genome sequencing, as described below. We named this CLV-like isolate "furtivovirus" (FTV).

c-TEM assay

For ultrastructural analysis, FTV-infected cells were harvested from the culture flasks at a multiplicity of infection (MOI) of 10 at 8 hpi. The harvested cells were centrifuged for 5 min at $500 \times g$. The pellet was resuspended in 1× phosphate-buffered saline (PBS) and centrifuged again for 5 min at $500 \times g$. Samples were chemically fixed five times in 2% glutaraldehyde in PBS for 30 min each and washed five times in PBS for 1 h each. The post-fixed samples were stained with 2% osmium tetroxide (NISSHIN EM Co.) for 1 h and washed twice with PBS. The fixed pellets were dehydrated through a graded

ethanol series (50%, 70%, 80%, 90%, 95%, and 99.5% for 5 min each, followed by three additional treatments with 99.5% for 10 min), followed by three dehydration steps using propylene oxide for 10 min each. The dehydrated samples were infiltrated with a 1:1 mixture of EPON resin. The infiltrated sample was then exchanged twice with fresh resin, and finally, the resin sample was polymerized at 60°C for 48 h. All procedures from fixation to dehydration were performed at 4°C. c-TEM observations were performed as previously described (17, 37, 38).

Genome analysis and tRNA annotation

Genomic DNA (gDNA) was extracted from purified FTV particles using the phenol/chloroform/isoamyl alcohol extraction method (15). The integrity of the gDNA was confirmed using agarose gel electrophoresis. The FTV genome was sequenced using the PACBIO_SMRT Revio sequencing platform, and reads were assembled *de novo* using Flye (version 2.9, <https://github.com/fenderglass/Flye>) by Macrogen Japan (Koto-ku, Tokyo, Japan). ORF prediction was performed using PROKKA (version 1.14.6). For functional annotation, eggNOG 4.5 and InterProScan (version 5.34-73.0) were utilized. Additionally, BLASTp searches against the NCBI non-redundant (nr) database (February 2025, e-value 10^{-5}) were conducted to assess amino acid homology, followed by manual curation of the annotations. Transfer RNAs (tRNAs) were identified using ARAGORN (version 1.2.41) with the following parameters: -t (tRNA type), -i (allow intron), -l (linear topology), and -s (both strands) (39). To evaluate the evolutionary context of the identified tRNAs, the encoded viral anticodons were compared with the host (*V. vermiformis*) codon usage indices, as recently described by Willemssen et al. (22). The genome map was visualized using Proksee (40).

Amino acid sequence and structural comparisons

The complete nucleotide (.fasta) and protein (.faa) sequences for the comparative viruses were obtained from the National Center for Biotechnology Information (NCBI). Tertiary structure predictions for polB and RNAPL proteins across the isolates were generated using AlphaFold3 (41) and visualized using PyMOL (version 3.1.3: <https://pymol.org/support.html>). To identify and map conserved catalytic motifs onto these structures, the primary amino acid sequences of polB from the respective viruses were aligned and visualized using the R package “DECIPHER” (42), referencing established canonical polymerase motifs (26–28).

OG, network, and nucleotide identity analysis

OrthoFinder (version 3.1.2) (43) was used to identify OGs and calculate gene-sharing levels among viral genomes. The overall protein-sharing patterns were visualized as a bipartite network using Gephi (version 0.10.1) (44). To measure the similarity of the shared gene contents, a Jaccard index matrix was computed based on OG clustering. Because simple ortholog counts may underrepresent true genetic distances, we assessed the nucleotide-level identity of shared OGs. Two customized Python pipelines were utilized to comprehensively assess nucleotide-level similarities. First, to evaluate the identity of shared OGs, gene-based ANI coding sequences were extracted and aligned using MAFFT (45), followed by pairwise sequence identity calculations that explicitly excluded alignment gaps. To account for the presence of paralogs within a single OG, an all-vs-all pairwise comparison was conducted between the sequences of the matched genomes to calculate the average orthologous identity. A second comparison was performed using NCBI BLASTn (dc-megablast) to compute the length-weighted ANI and genomic coverage matrices. Additionally, whole-genome ANI (orthoANLu), which computes identities to include non-coding regions, was calculated (46). Matrix and heatmap visualizations were performed using R software (version 4.5.0).

Phylogenomic reconstruction and inter-lineage similarity analysis

To reconstruct the evolutionary history of FTV, a comprehensive data set of 372 viral genomes was compiled, spanning three classes: *Megaviricetes*, *Mriyaviricetes*, and *Pokkesviricetes* (Data S1). Crucially, this data set incorporated published high-quality MAGs to bridge deep evolutionary lineages (29–33). Using the *nclv_markersearch* tool (29), seven GVOGs serving as core marker genes (DEAD/SNF2-like helicase, transcription initiation factor IIB, family B DNA polymerase, RNA polymerase large subunit, packaging ATPase, DNA topoisomerase II, and poxvirus late transcription factor VLTf3) were identified and extracted. A concatenated multiple amino acid sequence alignment was generated using Clustal Omega (version 1.2.4) (47). To determine the best-fit substitution model, ModelFinder (-m TEST), built into IQ-TREE 3 (version 3.0.1) (48) was employed. The unrooted maximum-likelihood phylogenetic tree was constructed using IQ-TREE 3 with the parameters -m Q.PFAM + F + I + R9 B 1000 -alrt 1000, where -B specifies the ultrafast bootstrap (UFboot) approximation (49) and -alrt performs the SH-aLRT test (50). The resulting tree was visualized using iTOL (version 7.5.0) (51). To quantitatively validate the taxonomic distances across deeply divergent lineages independently of ortholog presence/absence, we calculated the amino acid similarity of the seven concatenated marker genes across all orders. This was performed using a custom Python 3 script employing the BLOSUM62 substitution matrix (52) to account for evolutionary conservation and gaps, providing a more robust metric than the simple sequence identity. The resulting similarity matrix was visualized as a heatmap in R. Furthermore, RED (53) values were computed from the rooted phylogenetic tree using the *get_reds* function from the R package “castor” (54) to assess lineage-specific evolutionary rates.

Etymology of taxon nomenclature

FTV was named “furtivovirus” (from the Latin *furtivus*, meaning hidden or stealthy) because of its initial cryptic co-isolation alongside faustoviruses (similar to how CLV was co-isolated with faustovirus ST1 and usurpativiruses with faustovirus LCD7). The proposed family name, *Manesviridae*, is derived from the *Manes*—chthonic deities or spirits of the dead in the ancient Roman religion. This nomenclature adheres to the conventions established for CLV, reflecting the observation that three of the four isolated members of this family (FTV, CLV, and USPV) were discovered as “ghostly” or secondary entities emerging in viral screenings only after the lysis caused by the faster-growing faustoviruses (18–20).

ACKNOWLEDGMENTS

We thank Dr. Kazuyoshi Murata of the National Institutes of Natural Sciences for his helpful discussions. We also thank Macrogen Japan (Koto-ku, Tokyo, Japan) for genome sequencing, *de novo* assembly, and annotation, as well as the Hanaichi UltraStructure Research Institute (Okazaki, Aichi, Japan) for TEM analyses.

This research was supported by JSPS/KAKENHI grant number 20H03078 awarded to M.T. and Joint Research of the Exploratory Research Center on Life and Living Systems (ExCELLS) to M.T. (ExCELLS program No. 22EXC601-4).

J.B. isolated furtivovirus, performed infection experiments, genomic analyses (including annotation, classification, molecular phylogenetic analysis, and phylogenomics), structural prediction of proteins, and TEM analysis. J.B. contributed to writing the initial version of the manuscript. M.T. conceived the study design, acquired funding, and contributed to writing the initial version of the manuscript. Both authors contributed to the final manuscript.

AUTHOR AFFILIATION

¹Department of Mathematics and Science Education, Graduate School of Science, Tokyo University of Science, Shinjuku, Tokyo, Japan

AUTHOR ORCID*s*

Jiwan Bae  <https://orcid.org/0009-0001-5516-7559>

Masaharu Takemura  <http://orcid.org/0000-0002-6261-3297>

FUNDING

Funder	Grant(s)	Author(s)
Japan Society for the Promotion of Science	20H03078	Masaharu Takemura
Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences	22EXC601-4	Masaharu Takemura

AUTHOR CONTRIBUTIONS

Jiwan Bae, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization | Masaharu Takemura, Conceptualization, Project administration, Resources, Supervision

DATA AVAILABILITY

The sequence data for the furtivovirus genome are available in GenBank ([LC890700](https://doi.org/10.1128/JVI.02031-25-s0001)) under BioProject [PRJDB35989](https://doi.org/10.1128/JVI.02031-25-s0001) and [SAMD01617881](https://doi.org/10.1128/JVI.02031-25-s0002). The raw reads were obtained from the Sequence Read Archive ([DRR724694](https://doi.org/10.1128/JVI.02031-25-s0003)).

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Data S1 (JVI02031-25-s0001.xlsx). Virus list and raw data of BLOSUM62-based amino acid similarity analysis.

Data S2 (JVI02031-25-s0002.txt). Phylogenetic tree file data text.

Supplemental figures (JVI02031-25-s0003.docx). Fig. S1 to S5.

REFERENCES

- La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie J-M, Raoult D. 2003. A giant virus in amoebae. *Science* 299:2033. <https://doi.org/10.1126/science.1081867>
- Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie J-M. 2004. The 1.2-megabase genome sequence of mimivirus. *Science* 306:1344–1350. <https://doi.org/10.1126/science.1101485>
- Koonin EV, Dolja V, Krupovic M, Varsani A, Wolf Y, Yutin N, Zerbini F, Kuhn J. 2019. Create a megataxonomic framework, filling all principal taxonomic ranks, for dna viruses encoding vertical jelly roll-type major capsid proteins. ICTV [International Committee for Taxonomy of Viruses] Proposal. <https://doi.org/10.13140/RG.2.2.14886.47684>
- Iyer LM, Aravind L, Koonin EV. 2001. Common origin of four diverse families of large eukaryotic DNA viruses. *J Virol* 75:11720–11734. <https://doi.org/10.1128/JVI.75.23.11720-11734.2001>
- Boratto PVM, Oliveira GP, Machado TB, Andrade A, Baudoin J-P, Klose T, Schulz F, Azza S, Decloquement P, Chabrière E, Colson P, Levasseur A, La Scola B, Abrahão JS. 2020. Yaravirus: a novel 80-nm virus infecting *Acanthamoeba castellanii*. *Proc Natl Acad Sci USA* 117:16579–16586. <https://doi.org/10.1073/pnas.2001637117>
- Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, Koonin EV, Raoult D. 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc Natl Acad Sci USA* 106:21848–21853. <https://doi.org/10.1073/pnas.0911354106>
- Philippe N, Legendre M, Dautre G, Couté Y, Poirot O, Lescot M, Arslan D, Seltzer V, Bertaux L, Bruley C, Garin J, Claverie JM, Abergel C. 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341:281–286. <https://doi.org/10.1126/science.1239181>
- Abrahão J, Silva L, Silva LS, Khalil JYB, Rodrigues R, Arantes T, Assis F, Boratto P, Andrade M, Kroon EG, Ribeiro B, Bergier I, Seligmann H, Ghigo E, Colson P, Levasseur A, Kroemer G, Raoult D, La Scola B. 2018. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat Commun* 9:749. <https://doi.org/10.1038/s41467-018-03168-1>
- Iyer LM, Balaji S, Koonin EV, Aravind L. 2006. Evolutionary genomics of nucleocytoplasmic large DNA viruses. *Virus Res* 117:156–184. <https://doi.org/10.1016/j.virusres.2006.01.009>
- Aylward FO, Moniruzzaman M, Ha AD, Koonin EV. 2021. A phylogenomic framework for charting the diversity and evolution of giant viruses. *PLoS Biol* 19:e3001430. <https://doi.org/10.1371/journal.pbio.3001430>
- Yutin N, Wolf YI, Raoult D, Koonin EV. 2009. Eukaryotic large nucleocytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol J* 6:223. <https://doi.org/10.1186/1743-422X-6-223>
- Yoshikawa K, Blanc-Mathieu R, Song C, Kayama Y, Mochizuki T, Murata K, Ogata H, Takemura M. 2019. Medusavirus, a novel large DNA virus discovered from hot spring water. *J Virol* 93:e02130-18. <https://doi.org/10.1128/JVI.02130-18>
- Yoshida K, Zhang R, Garcia KG, Endo H, Gotoh Y, Hayashi T, Takemura M, Ogata H. 2021. Draft genome sequence of medusavirus stheno, isolated from the Tatakai River of Uji, Japan. *Microbiol Resour Anounc* 10:e01323-20. <https://doi.org/10.1128/MRA.01323-20>

14. Zhang R, Takemura M, Murata K, Ogata H. 2023. “*Mamonoviridae*”, a proposed new family of the phylum *Nucleocytoviricota*. Arch Virol 168:80. <https://doi.org/10.1007/s00705-022-05633-1>
15. Bae J, Takemura M. 2025. Near-complete genome of medusavirus euryale, putative third species of the genus *Medusavirus*, isolated from the TaeHwaGang river, Ulsan, South Korea. Microbiol Resour Announc 14:e0117124. <https://doi.org/10.1128/mra.01171-24>
16. Rolland C, Andreani J, Sahmi-Bounsar D, Krupovic M, La Scola B, Levasseur A. 2021. Clandestinovirus: a giant virus with chromatin proteins and a potential to manipulate the cell cycle of its host *Vermamoeba vermiformis*. Front Microbiol 12:715608. <https://doi.org/10.3389/fmicb.2021.715608>
17. Bae J, Hatori N, Burton-Smith RN, Murata K, Takemura M. 2025. A newly isolated giant virus, ushikuvirus, is closely related to claudestinovirus and shows a unique capsid surface structure and host cell interactions. J Virol 99:e0120625. <https://doi.org/10.1128/jvi.01206-25>
18. Sahmi-Bounsar D, Boratto PV de M, Oliveira GP, Bou Khalil JY, La Scola B, Andreani J. 2021. Single cell micro-aspiration as an alternative strategy to fluorescence-activated cell sorting for giant virus mixture separation. J Vis Exp 152. <https://doi.org/10.3791/60148>
19. Andreani J, Baudoin J-P, Laumier L, Pinault L, Chanteloup A, Colson P, Scola BL. 2025. Usurpatorvirus massiliensis, a new giant virus related to claudestinovirus. Virol J 22:385. <https://doi.org/10.1186/s12985-025-0297-z>
20. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, Rossmann M, Colson P, Raoult D, La Scola B. 2015. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. J Virol 89:6585–6594. <https://doi.org/10.1128/JVI.00115-15>
21. Rodrigues M, Queiroz V, Arantes T, Limborço H, Neiva B, Arias N, Machado T, Barcelos M, Cortines JR, Thiemann OH, Marques RE, Melo-Hanchuk TD, Hurtado Celis EL, Araujo JP Jr, Reis E, Junior Alcantara LC, Santos CBC, Jivaji A, Rodrigues RAL, Aylward FO, Abrahão JS. 2025. Naivirus: an enveloped giant virus with a pleomorphic, flexible tail. Nat Commun 16:8306. <https://doi.org/10.1038/s41467-025-63463-6>
22. Willemsen A, Manzano-Marín A, Horn M. 2025. Novel high-quality amoeba genomes reveal widespread codon usage mismatch between giant viruses and their hosts. Genome Biol Evol 17:evae271. <https://doi.org/10.1093/gbe/evae271>
23. Witt ASA, Carvalho J, Serafim MSM, Arias NEC, Rodrigues RAL, Abrahão JS. 2024. The GC% landscape of the *Nucleocytoviricota*. Braz J Microbiol 55:3373–3387. <https://doi.org/10.1007/s42770-024-01496-7>
24. Mendonça-Santos T, Abrahão J, Del-Bem L-E. 2025. Integrative genomic analyses support the division of the extended *Asfarviridae* clade into multiple viral families. J Virol 99:e0133725. <https://doi.org/10.1128/jvi.01337-25>
25. Stoddard BL. 2005. Homing endonuclease structure and function. Q Rev Biophys 38:49–95. <https://doi.org/10.1017/S0033583505004063>
26. Hopfner KP, Eichinger A, Engl RA, Laue F, Ankenbauer W, Huber R, Angerer B. 1999. Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. Proc Natl Acad Sci USA 96:3600–3605. <https://doi.org/10.1073/pnas.96.7.3600>
27. Wang J, Sattar AK, Wang CC, Karam JD, Konigsberg WH, Steitz TA. 1997. Crystal structure of a pol α family replication DNA polymerase from bacteriophage RB69. Cell 89:1087–1099. [https://doi.org/10.1016/S0092-8674\(00\)80296-2](https://doi.org/10.1016/S0092-8674(00)80296-2)
28. Wang TS, Wong SW, Korn D. 1989. Human DNA polymerase α: predicted functional domains and relationships with viral DNA polymerases. FASEB J 3:14–21. <https://doi.org/10.1096/fasebj.3.1.2642867>
29. Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO. 2020. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. Nat Commun 11:1710. <https://doi.org/10.1038/s41467-020-15507-2>
30. Schulz F, Roux S, Paez-Espino D, Jungbluth S, Walsh DA, Denev VJ, McMahon KD, Konstantinidis KT, Elie-Fadrosh EA, Kyrpides NC, Woyke T. 2020. Giant virus diversity and host interactions through global metagenomics. Nature 578:432–436. <https://doi.org/10.1038/s41586-020-1957-x>
31. Yutin N, Mutz P, Krupovic M, Koonin EV. 2024. Mriyaviruses: small relatives of giant viruses. mBio 15:e0103524. <https://doi.org/10.1128/mbio.01035-24>
32. Piedade GJ, Schön ME, Lood C, Fofanov MV, Wesdorp EM, Biggs TEG, Wu L, Bolhuis H, Fischer MG, Yutin N, Dutilh BE, Brussaard CPD. 2024. Seasonal dynamics and diversity of Antarctic marine viruses reveal a novel viral seascape. Nat Commun 15:9192. <https://doi.org/10.1038/s41467-024-53317-y>
33. Liu W, Nagasaka K, Wu J, Ban H, Mimick E, Meng L, Neches RY, Moniruzzaman M, Yoshida T, Nishimura Y, Endo H, Okazaki Y, Ogata H. 2025. Giant viruses specific to deep oceans show persistent presence and activity. mSystems 10:e0093225. <https://doi.org/10.1128/mSystems.00932-25>
34. Koonin EV, Yutin N. 2019. Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism. Adv Virus Res 103:167–202. <https://doi.org/10.1016/bs.aivir.2018.09.002>
35. Filée J. 2013. Route of NCLDV evolution: the genomic accord. Curr Opin Virol 3:595–599. <https://doi.org/10.1016/j.coviro.2013.07.003>
36. Bisio H, Legendre M, Giry C, Philippe N, Alempic JM, Jeudy S, Abergel C. 2023. Evolution of giant pandoravirus revealed by CRISPR/Cas9. Nat Commun 14:428. <https://doi.org/10.1038/s41467-023-36145-4>
37. Takemura M. 2016. Morphological and taxonomic properties of *Tokyovirus*, the first *Marseilleviridae* member isolated from Japan. Microbes Environ 31:442–448. <https://doi.org/10.1264/jmsme2.ME16107>
38. Takahashi H, Fukaya S, Song C, Murata K, Takemura M. 2021. Morphological and taxonomic properties of the newly isolated *Cotonvirus japonicus*, a new lineage of the subfamily *Megavirinae*. J Virol 95:e0091921. <https://doi.org/10.1128/JVI.00919-21>
39. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. <https://doi.org/10.1093/nar/gkh152>
40. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C-Y, Graham M, Van Domselaar G, Stothard P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res 51:W484–W492. <https://doi.org/10.1093/nar/gkad326>
41. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630:493–500. <https://doi.org/10.1038/s41586-024-07487-w>
42. S. Wright E. 2024. Fast and flexible search for homologous biological sequences with DECIPHER v3. R J 16:191–200. <https://doi.org/10.32614/RJ-2024-022>
43. Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol 20:238. <https://doi.org/10.1186/s13059-019-1832-y>
44. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. ICWSM 3:361–362. <https://doi.org/10.1609/icwsm.v3i1.13937>
45. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>
46. Yoon SH, Ha SM, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>
47. Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci 27:135–145. <https://doi.org/10.1002/pro.3290>
48. Wong T, Ly-Trong N, Ren H, Baños H, Roger A, Susko E, Bielow C, De Maio N, Goldman N, Hahn M, Huttley G, Lanfear R, Minh BQ. 2025. IQ-TREE 3: phylogenomic inference software using complex evolutionary models. Life Sciences. <https://doi.org/10.32942/X2P62N>
49. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. <https://doi.org/10.1093/molbev/msx281>
50. Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol 60:685–699. <https://doi.org/10.1093/sysbio/syr041>
51. Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 52:W78–W82. <https://doi.org/10.1093/nar/gkae268>
52. Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919. <https://doi.org/10.1073/pnas.89.22.10915>
53. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004. <https://doi.org/10.1038/nbt.4229>

54. Louca S, Doebeli M. 2018. Efficient comparative phylogenetics on large trees. *Bioinformatics* 34:1053–1055. <https://doi.org/10.1093/bioinformatics/btx701>