

# Naiavirus: an enveloped giant virus with a pleomorphic, flexible tail

Received: 28 April 2025

Accepted: 20 August 2025

Published online: 17 September 2025

 Check for updates

Matheus Rodrigues<sup>1,10</sup>, Victória Queiroz<sup>1,2,10</sup>, Thalita Arantes<sup>1</sup>, Henrique Limborço<sup>1</sup>, Bruna Neiva<sup>1</sup>, Nidia Arias<sup>1</sup>, Talita Machado<sup>1</sup>, Matheus Barcelos<sup>1</sup>, Juliana R. Cortines<sup>3</sup>, Otavio Henrique Thiemann<sup>1</sup>, Rafael Elias Marques<sup>1</sup>, Talita Diniz Melo-Hanchuk<sup>1</sup>, Eliana Leonor Hurtado Celis<sup>1</sup>, João Pessoa Araujo Jr<sup>1</sup>, Erik Reis<sup>1</sup>, Luiz Carlos Junior Alcantara<sup>7</sup>, Cezar Batista Cunha Santos<sup>2</sup>, Abdeali Jivaji<sup>8</sup>, Rodrigo A. L. Rodrigues<sup>1</sup>, Frank O. Aylward<sup>1</sup> & Jônatas Santos Abrahão<sup>1</sup>

Numerous studies have shown that viruses are present in a variety of environments on Earth, acting as drivers of biogeochemical cycles and powerful selective forces. Among them, giant viruses of amoebae have garnered attention from the scientific community due to their large particles and extensive genomes. Here, we describe the discovery of one of the largest tailed viruses in the known virosphere (averaging 1350 nm), named Naiavirus. This virus, isolated from a swamp biome in Brazil, has particles with a never-before-seen morphology and composition, and represents the first giant amoeba virus with an external envelope covering the capsid and extending over a flexible tail region. The Naiavirus genome, with nearly 1 million base pairs, reveals a unique set of genes, and does not resemble any other virus previously isolated so far.

Giant viruses of amoebae have garnered attention from the scientific community due to their large particles and extensive genomes, capable of encoding a significant array of unknown proteins, or proteins previously only observed in cellular organisms. Most known giant viruses infect free-living amoebae, particularly *Acanthamoeba castellanii*, which is the primary model for their study and isolation in laboratory settings<sup>1</sup>. Among amoeba-infecting nucleocytopiruses, the first to be isolated and characterized was mimivirus, in 2003, with impressive 750 nm particles and a 1.2 Mb genome<sup>2</sup>. In the following

years, several giant viruses of amoebae, all belonging to the phylum *Nucleocytopiricota*, were isolated, revealing significant structural and genomic diversity<sup>3</sup>. Despite this diversity, the known giant viruses of amoeba share some structural features, such as non-enveloped particles ranging from 0.1 to up to a micrometer or more, and the presence of an internal lipid membrane beneath the capsid, which is essential for the uncoating of genome and the proteins involved in the early stages of infection. The known giant viruses have double-stranded DNA genomes and share a series of genes encoding proteins that belong to

<sup>1</sup>Universidade Federal de Minas Gerais, Departamento de Microbiologia / Centro de Microscopia, Av. Antônio Carlos 6627, Belo Horizonte, Brazil. <sup>2</sup>Sexto Distrito Naval da Marinha do Brasil, Ladario, Mato Grosso do Sul, Brazil. <sup>3</sup>Universidade Federal do Rio de Janeiro, Departamento de Virologia, Instituto de Microbiologia Paulo de Góes, Rio de Janeiro, Brazil. <sup>4</sup>Universidade de São Paulo, Physics Institute of São Carlos, Trabalhador São Carlense Av., São Carlos, Brazil. <sup>5</sup>Brazilian Biosciences National Laboratory (LNBio), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, Brazil. <sup>6</sup>São Paulo State University (UNESP), Department of Genetic, Microbiology, and Immunology, Bioscience Institute and Institute of Biotechnology (IBTEC), Botucatu, Brazil. <sup>7</sup>Instituto René Rachou, Fundação Oswaldo Cruz and Expanded Navigation for Intensive and Optimized Surveillance (NAVIO) Network, Belo Horizonte, Brazil. <sup>8</sup>Department of Biological Sciences, Virginia Tech, 926 West Campus Drive, Blacksburg, VA, USA. <sup>9</sup>Center for Emerging, Zoonotic, and Arthropod-Borne Pathogens (CeZAP), Virginia Tech, Blacksburg, VA, USA. <sup>10</sup>These authors contributed equally: Matheus Rodrigues, Victória Queiroz.

 e-mail: [faylward@vt.edu](mailto:faylward@vt.edu); [jonatas.abrahao@gmail.com](mailto:jonatas.abrahao@gmail.com)

the structural and replicative modules of these viruses<sup>4</sup>. Metagenomic studies have also shown a great abundance and diversity of giant viruses worldwide, reinforcing their presence in oceans, sediments, rivers, hot springs and permafrost samples and expanding the repertoire of genes identified in giant viruses<sup>5,6</sup>.

Here we report the discovery of Naiavirus, the largest tailed virus identified to date in the known virosphere, measuring ~1350 nm (average). Isolated from a Brazilian swamp ecosystem, this virus exhibits an unprecedented particle structure and composition. It is the first giant virus infecting amoebas to feature an external envelope that encases the capsid and extends along a flexible tail. Naiavirus also possesses a genome of nearly one million base pairs, encoding a unique repertoire of genes that shows no similarity to any previously isolated virus.

## Results

### Naiavirus discovery, particle and cycle

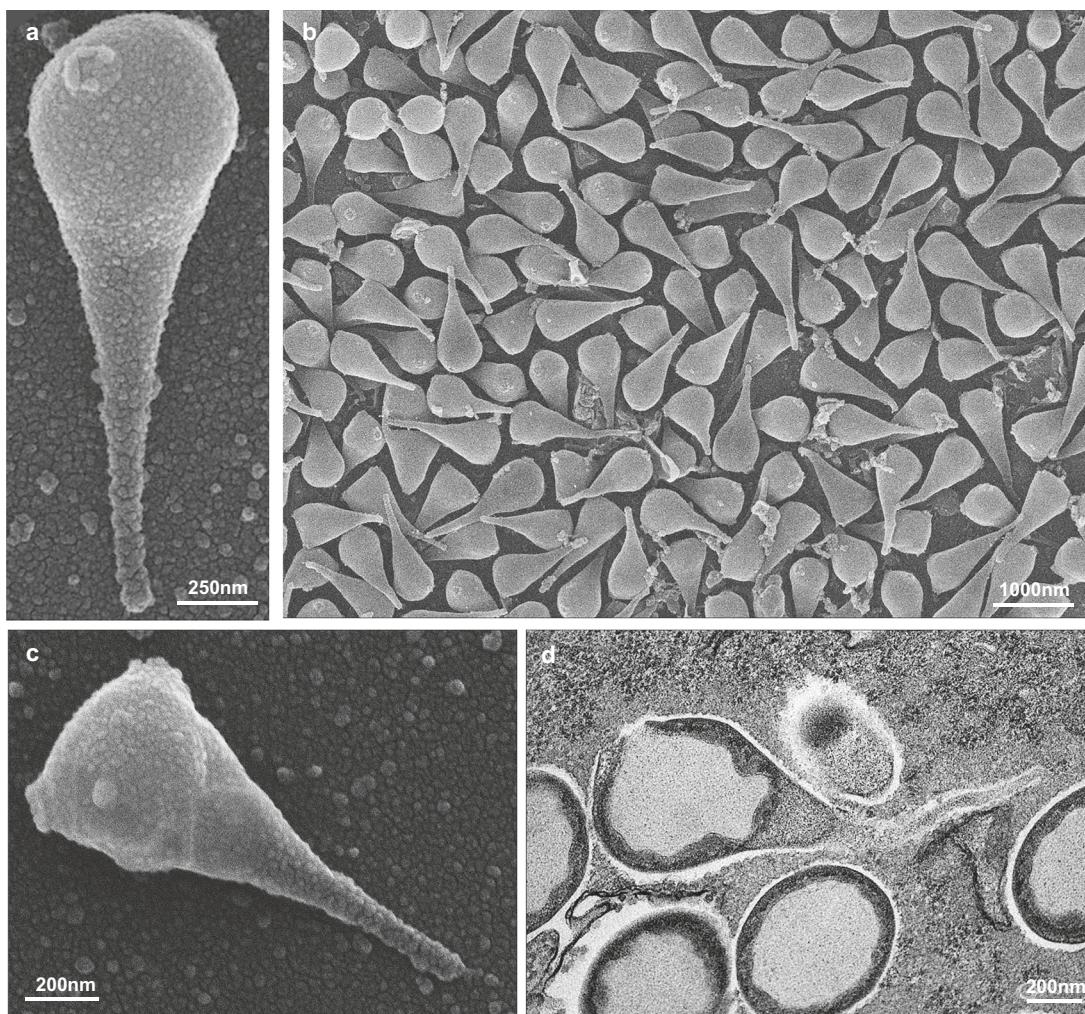
We surveyed samples collected from the southern region of the Paraguay River in the state of Mato Grosso do Sul, in the Brazilian Pantanal biome. This region is widely recognized as one of the largest hotspots for biological diversity in the world. The prior isolation of the first Tupanvirus from a sample collected in this biome motivated our further exploration in the Pantanal to identify additional giant viruses<sup>7</sup>. To this end, water samples were collected at 44 locations along the river. In the laboratory, the samples were processed and then inoculated in *Vermamoeba vermiformis*.

After unsuccessful prospecting in 439 samples, we observed cytopathic effects in *V. vermiformis* (ATCC50237) inoculated with a water sample collected in the Porto Murtinho region. The amoebae exhibited rounding and lysis, and in the supernatant we observed regular elements, possibly giant virus particles or cells of an amoeba parasitic microorganism (Supplementary Fig. 1a, b). Once isolated, this agent was inoculated again simultaneously into *V. vermiformis*, *A. castellanii* (ATCC30010) and *A. polyphaga* (UFMG isolate), and after 48 h, we observed that all tested amoebae exhibited cytopathic effects. This lysate was subjected to crystal violet staining and then observed under an optical microscope. The cells of *V. vermiformis* were lysed, releasing a massive amount of a strongly stained entity that appeared to have a tail (Supplementary Fig. 1c). We then produced specific antibodies in mice against this entity and applied them in immunofluorescence (IF) assays. IF clearly confirmed the presence of a tail in the structure of the entity (Supplementary Fig. 1d).

Scanning electron microscopy (SEM) of the lysates revealed an unexpected and completely novel virion morphology (Fig. 1a–c). With a structure consisting of an enlarged region, here referred to as the head, and a narrowed region, here referred to as the tail, the virion was pleomorphic and asymmetric, reaching up to 1824 nm in length (mean 1350 nm) (Supplementary Fig. 2). Curiously, the entire external surface of the particles appeared to be made of the same membranous material, similar to the plasma membrane of amoebae. It is important to mention that what we refer to as a tail does not correspond to a viral tail structure *sensu strictu*, but rather to a tapered extension of the envelope structure originating from the region referred to as the head. After analyzing hundreds of particles, we identified two axes: one in the shape of a drop (head region with an average width of 540 nm) (Fig. 1a and Supplementary Fig. 2) and another in the shape of a triangle (head region with an average width of 570 nm) (Fig. 1c and Supplementary Fig. 2). At the upper end of the latter axis, at least one, and more frequently two, ostioles were present, asymmetrically arranged in an antipodal distribution. These ostioles were more clearly observed in the particle positioned along the drop-shaped axis and appeared as approximately circular protrusions on the particle surface, about 100–125 nm in size (Figs. 1d and 2a–c). In some particles, the ostioles resembled a flower of the Asteraceae family.

Transmission electron microscopy (TEM) revealed even more impressive details of this entity. In the head region, an oval structure, similar to the capsid of some giant viruses like Orpheovirus, occupied the upper region where the two ostioles were located (Fig. 1d and 2a). A total of five structural layers were observed in the head region: 1) an external envelope covering the entire particle; 2) the outer layer of the oval capsid, thicker (averaging 25 nm) and electron-dense; 3) a layer below that, with internal undulations; 4) an internal membrane, as seen in nucleocytopiruses infecting amoebae; and 5) a broad electron-lucent region, where the genome is located (Fig. 2d and Supplementary Fig. 1d). In the ostiole region, there was a discontinuity in the external envelope and the layers composing the capsid wall (Fig. 2b, c). Interestingly, the internal content of the capsid is delimited in the ostiole region by the internal membrane and three additional layers of what appears to be membranous material, a structure never before described to our knowledge (Fig. 2b, c). Below the oval capsid, still within the head region, a roughly conical-shaped area (averaging 300 nm in length and 200 nm in width) was observed also covered by the external envelope (Fig. 1d). For a more detailed view of the outer membrane, new preparations were made for TEM, but using Spurr resin embedding with lower viscosity, which revealed more clearly the continuity of the envelope from the head region to the tip of the tail (Supplementary Fig. 3). In addition, longitudinal electron tomography analyses confirmed that the external envelope covers the entire viral particle from the apical region of the head to the tip of the tail (Fig. 2g, and Supplementary Movie 1). The oval-shaped capsid is clearly visible beneath the envelope, in the head region, with the aforementioned asymmetry (Fig. 2g). Particle analyses through tomography in the transverse plane revealed characteristics already described regarding the elements that make up the capsid and the tail region (Fig. 2h; and Supplementary Movies 1 and 2). At this point, we suspected that it was truly an unprecedented giant virus, with an average size of 1,350 nm, making it the largest tailed virus ever described in the virosphere and also the largest virus with an external envelope described to date. We named this entity Naiavirus, in honor of an indigenous Brazilian people, the Tupi-Guaranis. In their mythology, Naia was a beautiful indigenous woman who fell in love with and died from love for the Moon.

We then turned our attention to the tail region of Naiavirus. Unlike the head region, which showed less variation in size, the tail, with its membranous appearance, appeared quite flexible and variable in size in both SEM and TEM analyses (Fig. 3a–g and Supplementary Fig. 4). Notably, this structure appeared capable of folding, stretching, and shrinking in response to interactions with the adjacent substrate, a rare feature in the known virosphere (Fig. 3a–g). In TEM images, the tail appeared filled with an electron-lucent material and was delimited by a cylindrical membranous structure (Figs. 1d and 2c–f). Analyses of the replication cycle of Naiavirus in *V. vermiformis* and *A. castellanii* revealed that, most of the time, the first interaction between the amoebae and viral particles occurred via the tail. In several SEM images, the tail was observed attaching to the acanthapodia and pseudopodia of the amoebae, possibly stimulating their phagocytosis (Fig. 3a, b, d; Supplementary Fig. 5a and b). Once inside the phagosome, Naiavirus particles fused their internal membrane with the phagosomal membrane at the region of at least one of the ostioles, releasing the virion's internal content into the amoeba's cytoplasm (Supplementary Fig. 5c–g). TEM analyses revealed that the Naiavirus cycle occurs in the cytoplasm, with the cell's nucleus remaining intact even at late stages (Fig. 3j and Supplementary Fig. 6g and h), raising questions about whether the nucleus might play a role in the late stages of the cycle, when morphogenesis takes place. An electron-lucent viral factory occupying about 1/4 to 1/2 of the amoeba's cytoplasm was observed (Supplementary Fig. 6g and h). The morphogenesis of Naiavirus particles occurs from crescent-like structures, a feature described for several giant viruses of amoebas (Fig. 3h; and Supplementary Fig. 6a). These crescents, measuring



**Fig. 1 | Overview of the Naivavirus particles.** **a–c** Scanning electron microscopy (SEM) of purified Naivavirus particles. In **a**, the particle in its drop-like axis. In **(c)** the particle in its triangular-like axis. **d** Transmission electron microscopy (TEM) of

Naivavirus particles within the cytoplasm of *Acanthamoeba castellanii*. SEM and TEM experiments were repeated independently three times, with similar results.

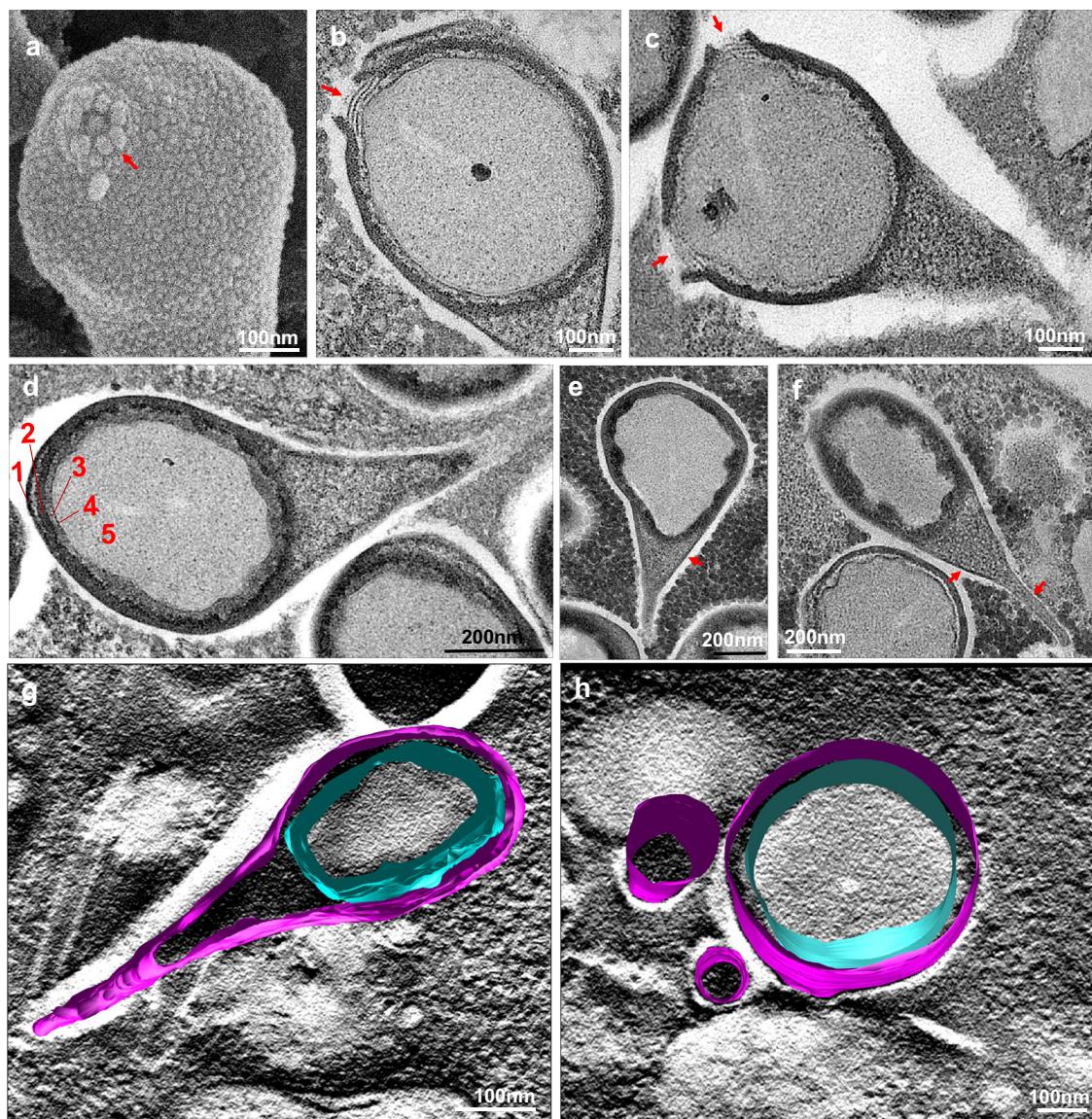
hundreds of nanometers, extend in various axes, encompassing the internal contents of the particle (Supplementary Fig. 6a–h). During their formation, at least two layers of membranous material are observed (Fig. 3h). Notably, the tail structure and the outer envelope appear to be incorporated into the particles within the viral factory (Fig. 1d; Fig. 3i, j; Supplementary Fig. 6d–h). At the viral factory's periphery, the particles seem to undergo maturation stages, with the capsid appearing thicker and its outer layer more defined (Supplementary Fig. 6e–h). No evidence of budding or exocytosis was observed. SEM and TEM analyses strongly suggest that all morphogenesis occurs within the cell's cytoplasm, and the particle release mechanism is via cell lysis induction (Fig. 3k, l; and Supplementary Fig. 7). It is important to mention that the complete replication cycle lasts ~24 h. In amoeba cultures infected synchronically, the viral factory becomes visible from 12 h; mature particles appear from 14 h; and cell lysis occurs 24 h onward.

As mentioned, Naivavirus has an external envelope that covers the entire particle. In the virosphere, envelopes are typically lipidic in nature, making them susceptible to detergents. Treatment of Naivavirus particles with Tween 20 at concentrations of 0.1% and 1% did not cause a significant reduction in viral titers. While the untreated particles induced infection with average titers of  $10^{7.5}$ , particles treated with 0.1% and 1% Tween induced infections with average titers of  $10^{7.4}$  and  $10^{7.2}$ , respectively. However, when treated with a 10%

concentration, complete inactivation of the particles was observed. SEM analysis revealed damage to the viral envelope (Supplementary Fig. 8). These results suggest that the envelope of Naivavirus may have a composition different from that of regular lipids, possibly consisting of a lipid resistant to the detergent tested at lower concentrations. Although the nature of the envelope still requires further characterization, analyses using TEM, SEM, IF, and tomography indicate that the particle is completely covered by this mantle. Therefore, Naivavirus represents the largest enveloped virus in the known virosphere.

#### Naivavirus genome and evolutionary history

We then sequenced the Naivavirus genome using the Illumina MiSeq platform to provide insight into its evolution and genetic repertoire (see Methods). The DNA of purified virus was sequenced, and the assembly revealed a circular genome, 922 kbp in size with 33.6% G + C content and an average coverage of 791x. A total of 867 genes were predicted, ranging from 90 to 4323 bp. We performed homology searches using BLASTp against the NCBI nr database, revealing that a quarter of Naivavirus proteins had no detectable matches (i.e., ORFan genes). Furthermore, nearly 50% of the identified genes correspond to proteins with unidentified functions that were previously detected in other organisms, especially in nucleocytoplasma metagenomic sequences. To improve the annotation of Naivavirus proteins, we additionally



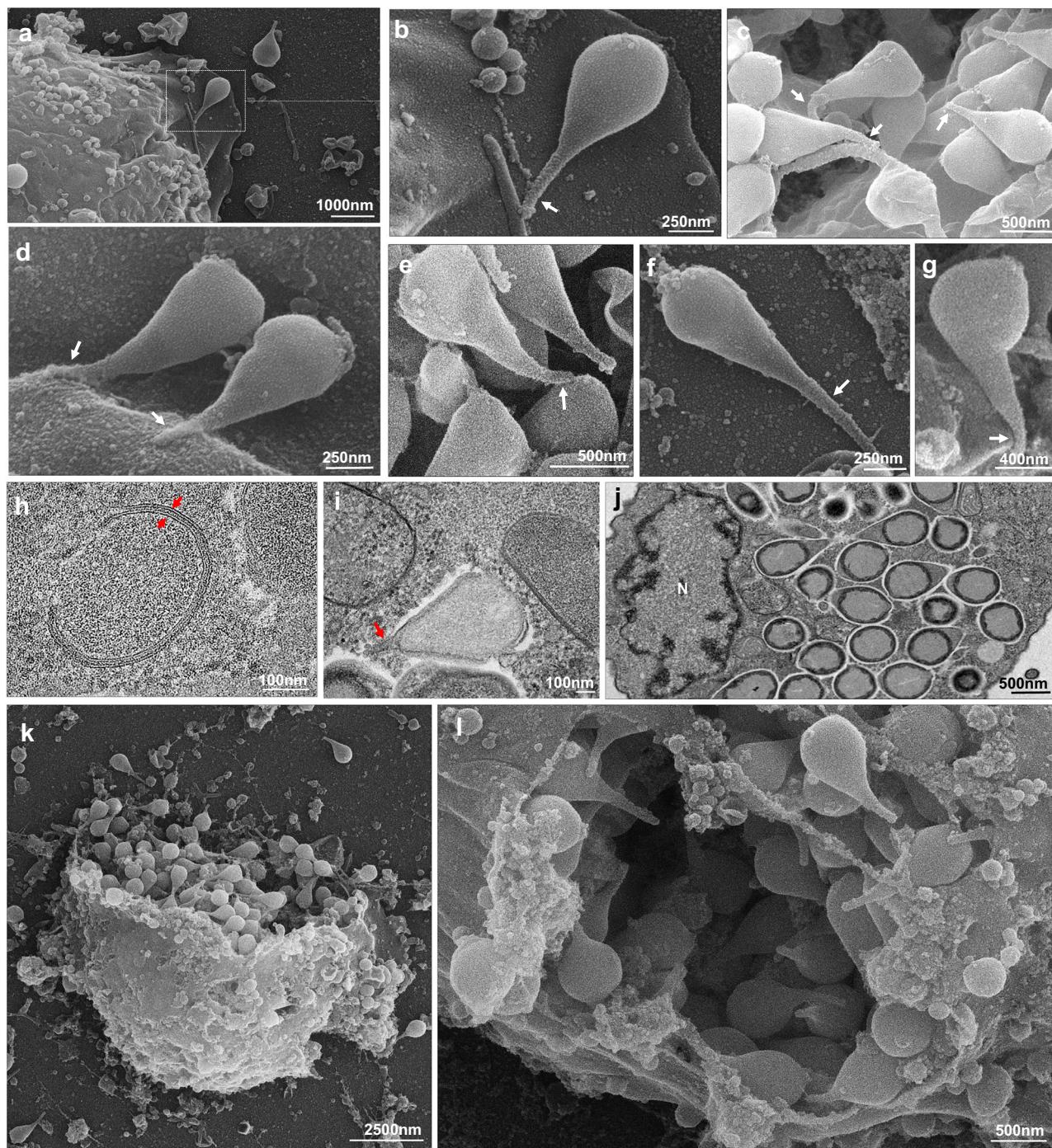
**Fig. 2 | Special features of Naiaivirus particles.** **a–c** Ostiole of Naiaivirus particles. In **a** SEM of the ostiole, highlighting the external aspect of this structure (arrow). In **b, c** the internal aspect of the ostiole, observed by TEM, showing that the ostiole is composed of three membranous structures (arrows), positioned over the inner membrane of the capsid. **d** Layers identified in Naiaivirus particles: 1) the external envelope covering the entire particle; 2) the outer layer of the oval capsid, thicker (25 nm) and electron-dense; 3) a layer below that, with internal undulations; 4) an internal membrane, as seen in nucleocytoporiviruses infecting amoeba; and 5) a broad

electron-lucent region, where the genome is located. **e, f** Arrows highlighting the external envelope in the tail region. **g, h** Electron tomography of Naiaivirus particles, highlighting the external envelope (magenta) and the capsid (light blue). In **g** the particle in a longitudinal section. In **h** three Naiaivirus particles in a cross-section, with the ones on the left in the tail region and the one on the right in the head region. SEM and TEM experiments were repeated independently three times, with similar results.

employed InterProScan, HHpred, and EggNOG-mapper searches<sup>9–12</sup>. Nevertheless, 127 proteins remained classified as ORFans.

Analysis of these annotation results revealed that the Naiaivirus genome encodes a balanced distribution of predicted functional categories. Among other functional categories widely present in the Naiaivirus genome, genes associated with protein translation are notable, including 5 aminoacyl tRNA synthetases (Glycine (ID\_189), Methionine (ID\_222), Tryptophan (ID\_457), Tyrosine (ID\_657), Aspartate/Asparagine (ID\_694), a peptidyl-hydrolase (ID\_318) and 14 translation factors, including initiation, elongation, and termination. Additionally, 6 tRNAs were predicted in Naiaivirus genome [tRNA-Ile (tat), tRNA-Ile (aat), tRNA-Leu (taa), tRNA-Leu (caa), tRNA-Cys (aca), and tRNA-Cys (gca)]. There is no clear correlation between the presence of the viral tRNA genes and the codon usage observed for Naiaivirus. Nevertheless, Naiaivirus encodes tRNAs associated with the most commonly used

codons for leucine and isoleucine, which may be advantageous for the viral translation process during infection (Suppl. Table 1). Genes involved in carbohydrate metabolism are also widely represented, particularly ten glycosyltransferase homologs, belonging to the families 2, 25, 45, 85 and others. Among the genes related to lipids, notable mentions include 7-dehydrocholesterol reductase (ID\_7), three class-3 lipases (IDs\_38, 176, 660), glycerol-3-phosphate cytidylyltransferase (ID\_392), fatty acid hydroxylase (ID\_407), cyclopropane-fatty-acyl-phospholipid synthase (ID\_444) and two patatin phospholipases (IDs\_783, 533). Host-virus and signal transduction genes were also found, including several kinases, phosphatases, and an endonuclease. A total of 47 genes related to transcription and RNA processing were found, including all those necessary for transcription, such as the canonical multimeric RNA polymerase subunits (IDs\_5, 393, 396, 799), RNase III (IDs\_61 and 469), RNA ligase (IDs\_68, 353) Cap-specific mRNA



**Fig. 3 | Naiavirus cycle.** **a–g** SEM of Naiavirus particles emphasizing the flexibility and mobility of the tail region (arrows). The tail of the particles adheres to the adjacent substrate, adapting to the surface according to its topography. **h–j** Details of the cycle observed by TEM. **h** Crescents observed in the viral factory. The arrows show two juxtaposed structures at the beginning of crescent formation. **i** The tail

region being formed along with the rest of the particle, in the viral factory (arrow). **j** Nuclear region (N) present at the end of the cycle. On the right, newly formed Naiavirus particle in the cytoplasm. **k, l** SEM of cells infected by Naiavirus after lysis. SEM and TEM experiments were repeated independently three times, with similar results.

(nucleoside-2-O-)methyltransferase (IDs\_102 and 103) and others. An impressive number of 52 genes related to DNA replication, recombination and repair were found, including DNA mismatch repair ATPase (ID\_57), AP (apurinic) endonuclease family 2 (ID\_84), NAD-dependent DNA ligase (ID\_92), D5 helicase-primase (IDs\_120), VV D6-like helicase (ID\_159), DNA primase (ID\_326), DNA topoisomerase IIA (ID\_398), DNA repair exonuclease (ID\_430) and DNA polymerase family B (ID\_582).

In addition, a long list of proteins with unexpected functions was identified in Naiavirus. A phosphoglycerate mutase (ID\_648), an

enzyme involved in glycolysis and gluconeogenesis, was predicted and is essential for metabolizing glucose and 2,3-phosphoglycerate. Curiously, a staygreen protein (ID\_776), typically found in plants and involved in regulating chlorophyll senescence, was also identified. Two paralogs of thaumatin family protein (IDs\_365, 693) were predicted, which have also been previously identified in plants and are involved in defense functions. Bacterial-related proteins were also identified: an exospore glycoprotein BclB-related protein (IDs\_29 and 792), typically associated with the formation of the exospore layer of bacterial

spores; and two paralogs of hemolysin III (IDs\_209 and 852), a virulence factor found in some bacteria that causes hemolysis. Additionally, an aquaporin (ID\_689), a mitochondrial carrier protein (ID\_365), and a tumor suppressor protein, the cyclin-dependent kinase inhibitor 3 (ID\_181), were predicted. Finally, HHpred predicted that ID\_802 bears remote homology to a 39S ribosomal protein L38 as well as a Phosphatidylethanolamine-binding protein.

Considering that the particles of Naiavirus do not resemble any virus previously described, a central question that has intrigued us since the discovery of this entity was its phylogenetic position in the virosphere. Thus, with the genome annotated, we searched for genes commonly shared among giant viruses. Orthologs of hallmark and core genes from nucleocytopiruses were consistently found in the Naiavirus genome, including the previously mentioned DNA polymerase family B (ID\_582), DNA topoisomerase IIA (ID\_398), VLTf3 late transcription factor (ID\_540), D5 helicase-primase (ID\_120), and a major capsid protein (ID\_581) with a double-jelly roll domain. A phylogenetic tree of DNA polymerase family B was then constructed, as this gene has been used in the literature as a phylogenetic marker for nucleocytopiruses (Supplementary Fig. 9). The topology of the tree suggests that Naiavirus forms a deeply-branching lineage within the known representatives of the *Pimascovirales* order. Importantly, none of the pimascoviruses discovered through viral isolation, such as Pithovirus, Cedratvirus, and Orpheovirus, cluster near Naiavirus. Interestingly, one metagenome-derived genome that exhibits phylogenetic affinity for Naiavirus is the recently described Hydrivirus, which was assembled from a Siberian permafrost sample dated to 42,000 years ago<sup>13</sup>. Since it was detected through metagenomics, the morphological characteristics of the Hydrivirus particle, as well as data related to its host and life cycle, have remained mysterious.

To examine the prevalence of other viruses with phylogenetic affinity for Naiavirus, we surveyed >15,000 publicly available metagenomes and identified several PolB sequences that fall within the new Naiavirus clade. We then performed metagenome binning to recover 11 novel metagenome-assembled genomes (MAGs) from this lineage. The other metagenomic sequences that cluster near Naiavirus were detected in various habitats, such as permafrost/cryo soil, acid mine sediment, bioreactor wastewater, water samples from lakes, etc., and several locations, including Russia, China, the USA, Canada, the Czech Republic, Sweden, Austria, and Greenland, with detections dating back to 2011 (Fig. 4c). This indicates that Nai-like viruses are present in different regions of the globe but, curiously, have never been isolated before. To assess the consistency of the results obtained from the previous phylogenetic analysis, a new tree was constructed, this time utilizing seven concatenated genes from both isolated viruses and MAGs (Fig. 4a). As previously observed, the concatenated tree revealed that Naiavirus falls within the order *Pimascovirales*, forming a distinct clade from the already known viruses in this group. Both Naiavirus, as well as Orpheovirus and Hydrivirus, are distributed among MAGs that also harbor several core genes of nucleocytopiruses. This finding indicates that this large branch of pimascoviruses remains poorly characterized, and that entirely new viruses belonging to this group are likely to be isolated in the coming years.

We also investigated the evolutionary history of Naiavirus through a global analysis of the sharing of predicted proteins among the pimascoviruses. In this analysis, not only Naiavirus but also all groups of pimascoviruses that have been isolated (cedrat-, pitho-, Orpheovirus) and metagenomic assembled genomes (MAGs) – including Hydrivirus – were considered. The result of this analysis is a network showing the protein sharing among 29 viruses (isolated or MAGs) (Fig. 4b). The largest cluster observed for Naiavirus consists of unique predicted proteins not shared with any other virus. However, several small clusters were observed between Naiavirus and various MAGs, with one large exclusive cluster between Naiavirus and MAG-Hydrivirus. This result suggests that Naiavirus has a distinct

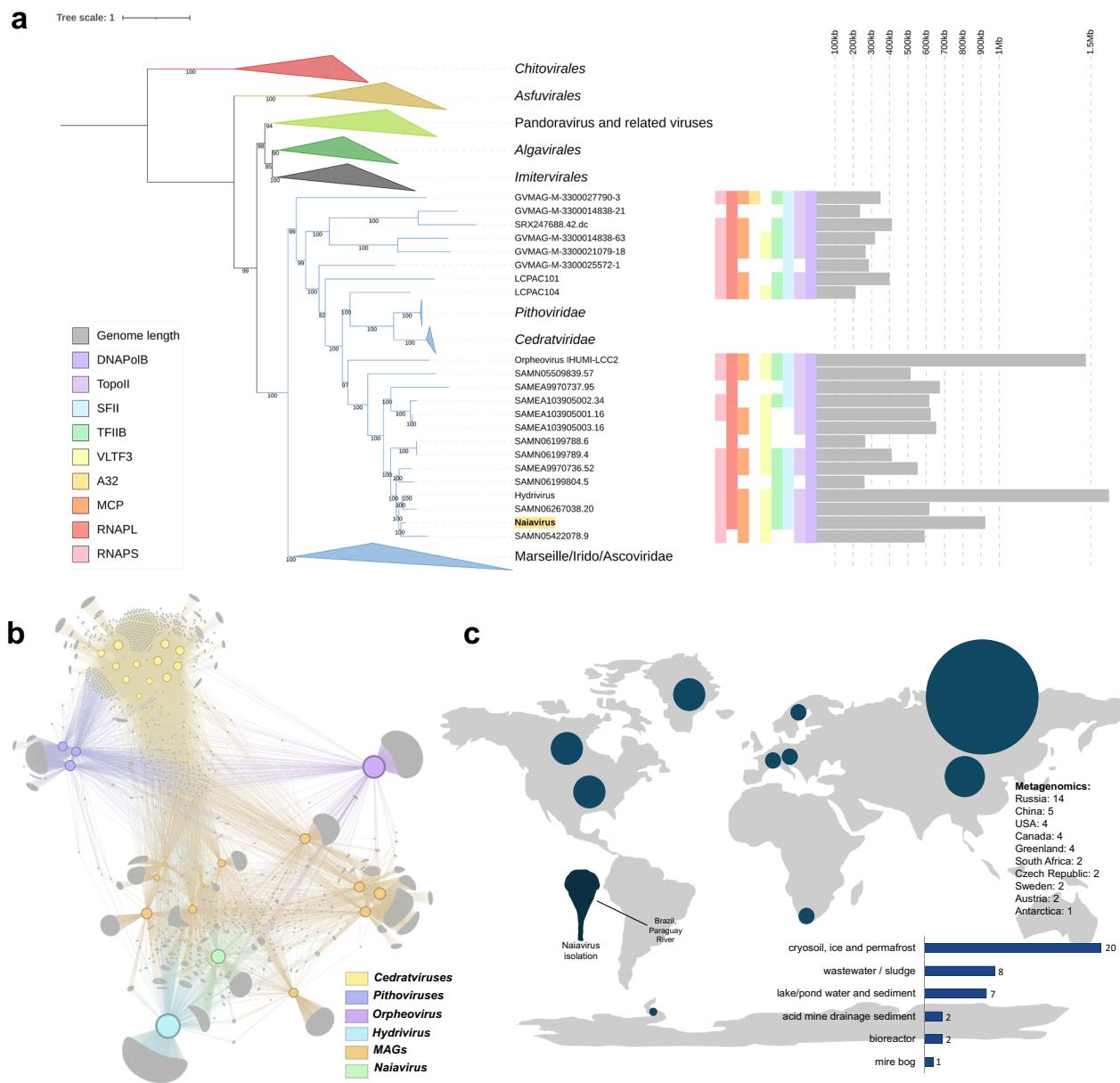
evolutionary history, with unique predicted proteins that are not shared with other known viruses, which may indicate a significant level of genomic differentiation. The observation of small clusters between Naiavirus and various MAGs, and a large exclusive cluster between Naiavirus and MAG-Hydrivirus, suggests a possible evolutionary relationship or exchange of genetic material between these viruses and other MAGs, which may indicate processes of coevolution or interaction in specific environments. Interestingly, the analysis of the contribution of paralogs in the genomes of pimascoviruses revealed that Naiavirus and Hydrivirus have the lowest percentage of gene duplications within this group (Supplementary Table 2).

### Naiavirus particles' proteome

Considering that the genome and virion organization of Naiavirus are unprecedented in the virosphere, we sought to examine the viral proteins that make up the particle. To address this, viral particles were produced by infecting 20 T175 flasks containing  $10^7$  *Acanthamoeba castellanii* and then purified twice, sequentially in a sucrose gradient. The ultrapure particles were subjected to proteomics and 254 viral proteins associated with the particle (Supplementary Table 3) were identified, a number that surpasses those reported for all giant viruses, except pandoraviruses. Overall, 17 proteins involved in transcription and RNA processing were detected, containing a variety of elements that could explain the occurrence of transcription of viral genes entirely in the cytoplasm, consistent with electron microscopy findings, where no phase of the cycle was observed in the cell's nucleus. Additionally, 15 proteins involved in DNA replication, recombination, and repair were found in Naiavirus particles, suggesting that replication of its genome likely occurs in the early stages of the cycle. A homolog to a H2B histone was also found in the genome and proteomic results (ID\_277), suggesting that this protein is potentially used for packaging the genome in the capsid. Finally, proteins involved in host interaction (e.g., kinases), lipid metabolism, protein metabolism, and others were also identified. Still, a total of 174 proteins were identified as ORFans or hypothetical proteins. To complement the analysis, we performed Foldseek searches on AlphaFold-predicted structures in the Naiavirus proteome (Supplementary data 1). As a result, we were able to successfully annotate only 229 proteins considering all of the approaches. A total of 19 proteins remained classified as hypothetical proteins, including the most abundant protein (ID\_538), and six proteins were entirely novel, with no matches in any database (Supplementary Table 3), highlighting the vast untapped reservoir of novel proteins in giant viruses. Foldseek analysis also revealed that the viral particle packaging proteins potentially involved in host takeover, such as remote homologs to a interleukin-1 receptor-associated kinase 4 (ID\_542), and a Ubx-domain-containing protein (ID\_192). Notably, we identified some proteins with remote homology to phage structural proteins, including the baseplate tripod (ID\_214), tail fibers (ID\_206 and ID\_561), and tailspike protein (ID\_48). Interestingly, although the major capsid protein (MCP) is typically the most abundant protein in the capsids of various nucleocytopiruses, in Naiavirus particles, it ranks only as the 169th most abundant protein (Supplementary Table 3 and 4). Considering the completely unprecedented structure of Naiavirus particles, it is possible that the MCP's primary structural function has been modified throughout the virus's evolution, perhaps acting as a scaffolding protein. It is worth mentioning that such a phenomenon appears to have occurred in other pimascoviruses as well.

### Discussion

The discovery of Naiavirus emphasizes that viral isolation studies remain crucial and complement the growing exploration of the virosphere through metagenomic methods. Unique structural elements, such as the flexible tail, the envelope, and the ostioles, demonstrate that even after more than 20 years since the description of the first



**Fig. 4 | Genomics and evolution of Naiavirus.** **a** Concatenated phylogenetic tree of nucleocytoplasmic viruses, showing that Naiavirus is positioned among various metagenomic sequences related to pimascoviruses, distant from any other previously isolated virus. The colored squares next to the sequences show the presence/absence of some of the core genes usually identified in giant viruses. **b** Global analysis of the sharing of predicted proteins among pimascoviruses, including

isolated viruses (cedrat-, pitho-, and Orpheovirus) and metagenomic assembled genomes (MAGs). The largest cluster observed for Naiavirus consists of unique predicted proteins not shared with any other virus. **c** Global distribution of Nai-like viruses detected after analysis of over 15,000 metagenomic databases. Although widely distributed around the world, Naiavirus represents the only isolated virus in this group.

amoeba viruses, giant viruses continue to surprise the scientific community with their unique structural adaptations. The fact that Naiavirus is capable of infecting different types of amoebae, a trait previously only described for Tupanviruses, suggests that the host spectrum of giant viruses is likely defined by much more complex factors than previously hypothesized, such as the presence of an almost complete set of genes associated with translation, as seen in Tupanviruses. The understanding of the evolution of giant virus genomes remains a challenge, as do the factors that determine which proteins are incorporated into their complex virions. The discovery of Naiavirus, the largest, enveloped, tailed virus ever described, symbolizes contemporary virology, where concepts and definitions regarding the boundaries of the virosphere must be frequently revisited.

## Methods

### Virological methods

For viral isolation, 96-well plates containing  $4 \times 10^4$  cells per well of *Vermamoeba vermiformis* (ATCC 50237) in Peptone Yeast Extract Glucose (PYG) medium supplemented with ciprofloxacin (4 µg/mL), vancomycin (4 µg/mL), doxycycline (20 µg/mL), penicillin (500 U/mL), streptomycin (100 mg/mL), and amphotericin B (25 mg/mL) were inoculated with 100 µL of pure or 1:10 diluted water samples collected from 44 different points along the Paraguay River in the state of Mato Grosso do Sul, Brazil. The collections were carried out within the context of the Navio project (Expanded Navigation for Intensive and Optimized Surveillance), with support from the Brazilian Navy. For the negative control, 100 µL of PBS was used instead. The plates were

incubated at 29 °C and monitored daily under an optical microscope for cytopathic effects (CPE).

Upon CPE observation, the sample was collected, and viral stocks were produced using twenty T175 cm<sup>2</sup> cell culture flasks containing  $1 \times 10^7$  *V. vermiciformis* cells cultivated in PYG medium supplemented with penicillin (500 mg/mL), streptomycin (100 mg/mL), and amphotericin B (25 mg/mL). The cells were infected with Naiavirus at a multiplicity of infection (MOI) of 0.01 and incubated at 29 °C until typical CPE was observed.

For viral purification, the flask contents were collected and subjected to freezing and thawing to release viral particles that might still be trapped inside intact cells. The lysate was then centrifuged at 1200 × g for 10 min to remove cell debris. The supernatant was collected, layered over a 24% sucrose gradient, and centrifuged twice at 36,000 × g for 1 h. The resulting pellet was resuspended in phosphate-buffered saline (PBS), and the purified viruses were stored at -20 °C. The viral titer was determined using the endpoint method, calculated according to Reed-Muench, and expressed as the number of 50% tissue culture infective doses (TCID<sub>50</sub>) per milliliter<sup>14</sup>. Host spectrum experiments were conducted by infecting the following amoebas with the Naiavirus seed pool: *Acanthamoeba castellanii* (ATCC30010), *Acanthamoeba polyphaga* (UFMG isolate), *Vermamoeba vermiciformis* (ATCC50237), and *Naegleria gruberi* (USP isolate). Considering that the virus did not replicate only in Naegleria, the remaining experiments were primarily conducted in *Acanthamoeba castellanii* and *Vermamoeba vermiciformis*, as described in each section.

For the analysis of Naiavirus susceptibility to Tween 20, the purified viral particles were treated with this detergent at concentrations of 0.1%, 1%, and 10%. After 30 min of incubation, the particles were washed twice with PBS and then subjected to titration in *Acanthamoeba castellanii*. The detergent toxicity to the amoeboid cells was previously calculated at different concentrations and dilutions, and these factors were considered in the interpretation of the results.

## Microscopy

For the TEM assay,  $2 \times 10^7$  cells were infected at an MOI of 0.01. Once CPE was observed, the flask contents were centrifuged for 10 min at 800 × g. The pellet was washed with 0.1 M sodium phosphate buffer (pH 7.4) and fixed with 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer for at least 1 h at room temperature. The pellet was then washed again with 0.1 M sodium phosphate buffer and resuspended in the same solution. Subsequently, the samples were fixed for 2 h in 2% osmium tetroxide, incubated overnight in 2% uranyl acetate at 2–8 °C, dehydrated in a graded ethanol series, incubated in acetone, and embedded in EPON or Spurr resin. Ultrathin sections were analyzed using a Tecnai G2-12 FEI Spirit Biotwin 120 kV electron microscope at the Microscopy Center of UFMG, Brazil.

For the SEM assay, the lysate of  $4 \times 10^6$  cells or purified viral particles were added to round glass coverslips coated with poly-L-lysine and fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer for at least 1 h at room temperature. The samples were then washed three times with 0.1 M cacodylate buffer and post-fixed with 1.0% osmium tetroxide for 1 h at room temperature. After this second fixation, the samples were washed three more times with 0.1 M cacodylate buffer and immersed in 0.1% tannic acid for 20 min. Next, they were washed in cacodylate buffer and dehydrated through a graded ethanol series (35–100%). Critical point drying using CO<sub>2</sub> was then performed, followed by mounting on stubs and coating with a 5-nm gold layer. Analyses were carried out using the ThermoFisher APREO 2 C electron microscope at the Microscopy Center of UFMG. Morphometric analyses were conducted using ImageJ software (version 1.54p).

For IF analyses, purified viral particles were stained with polyclonal anti-Naiavirus whole particle antibodies produced in mice (diluted 1:400 in PBS Tween-20 with bovine serum albumin [BSA] 3%) for 1 h at 37 °C. The samples were then incubated with anti-mouse

secondary antibodies (diluted 1:400 in PBS Tween-20 with BSA 3%) and DAPI Sigma (diluted 1:1000 in PBS) for 1 h at 37 °C. Fluorescently labeled particles were visualized (objective 100x) using an ApoTome Zeiss Axio Imager Z2 Apotome 2 microscope. Image processing was performed using Zen Lite software 2025 version (Zeiss Microscopy) at the Microscopy Center of UFMG, Brazil.

## Electron tomography sample preparation and data capture

For the electron tomography preparation, *Vermamoeba vermiciformis* cells were fixed in culture with 2.5% (v/v) glutaraldehyde for 2 h. Cells were then washed and resuspended in 200 mM phosphate buffer (pH 7.0). The pellet was washed, post-fixed with 1% osmium tetroxide for 2 h and stained en bloc with 2% uranyl acetate for 2 h, then dehydrated in an ethanol series and embedded in Epon resin. Sections with a nominal thickness of 250 nm were cut using an ultramicrotome and a 35 DiATOME diamond knife and collected via a water bath onto slot electron microscopy grids Slot.

Transmission electron micrographs of tilt series for single axis tomographic reconstruction were captured on a Tecnai G2-12 at 120 kV electron acceleration voltage using a 4096 × 4096 Eagle CCD (FEI). Images were captured at 1° steps from -60° to +60° using Serial EM (University of Colorado)<sup>15</sup> and the fiducial less tomographic reconstruction of the tilt series was performed using IMOD 5.1 suite (University of Colorado)<sup>16</sup>. From the complete stack, 107 image slices were selected and cropped to 2183 × 2423 pixels containing the region of interest. This dataset was used for manual segmentation and modeling of each structure presented in this work.

## Genome sequencing and annotation

For DNA extraction, 200 µL of purified Naiavirus was processed using the MagMAX™ CORE Nucleic Acid Purification Kit, following the manufacturer's instructions. The viral DNA was then eluted in 70 µL, and its quality and concentration were assessed using a Nanodrop spectrophotometer (Thermo Scientific, Waltham, MA, USA). The genome was sequenced using an Illumina MiSeq instrument with a single-end library, prepared with the Illumina DNA Prep kit (Illumina Inc., San Diego, CA, USA). A total of 19.7 million reads were generated. The genome was assembled de novo using SPAdes software 3.12 with the parameter “–isolate”<sup>17</sup>. Additionally, tRNA coding sequences were predicted using ARAGORN<sup>18</sup>.

Gene prediction was performed using Prodigal v2.6.3<sup>19</sup>. The codon usage table was generated using Jamie McGowan's Online Bioinformatics Tools ([https://jamiemcgowan.ie/bioinf/codon\\_usage.html](https://jamiemcgowan.ie/bioinf/codon_usage.html)). All of the predicted ORFs were annotated using Diamond (Evalue of < 10<sup>-5</sup>) against the NCBI nonredundant protein sequence database<sup>20</sup>, HHpred server<sup>9</sup>, considering as valid hits those that presented probabilities greater than 80% or E values equal to or smaller than 1; eggNOG-mapper<sup>12</sup> and InterProScan<sup>10,11</sup>. The genomic data were deposited in the Genbank - accession: PV450503. Although the data are not yet publicly available, reviewers and editors can access them using this FigShare link: (<https://figshare.com/s/ef4eddbe9cc10b4789c4>).

## Proteomic analysis

For the proteomics analysis, Naiavirus particles were produced in *Acanthamoeba castellanii* and then subjected to a two-step purification process. For protein extraction,  $10^8$  viral particles were resuspended in Laemmli buffer (90 mM DTT, 2% SDS, 80 mM Tris-HCl pH 6.8, 10% glycerol, 0.1% bromophenol blue) for protein extraction and solubilization. Proteins were then separated using SDS-PAGE, and electrophoresis was halted as soon as the samples entered the resolving gel. The gel was stained with Coomassie Brilliant Blue, and the single band containing all proteins was excised for in-gel digestion. The excised gel band was destained using a solution of 50% methanol and 2.5% acetic acid, followed by protein reduction with 10 mM dithiothreitol (DTT) and alkylation with 50 mM iodoacetamide in the

dark for 30 min. The reaction was quenched by adding 10 mM DTT and incubating in the dark for 15 min. Protein digestion was performed using 1 µg of trypsin (Promega, Madison, WI, USA) at 37 °C for 16 h. The reaction was stopped by adding 5% formic acid, and the peptides extracted with 5% formic acid in 50% acetonitrile. Peptides were desalting using the StageTips method with C18 Empore disks (3 M, St Paul, MN, USA). After desalting, samples were dried in a vacuum concentrator and reconstituted in 10 µL of 0.1% formic acid for further analysis.

The proteomic analysis was performed using LC-MS/MS on an Orbitrap Exploris 240 mass spectrometer (Thermo Fisher Scientific, USA) coupled to an EASY-nLC 1200 system (Proxeon Biosystem, USA) through a Proxeon nanoelectrospray ion source. Peptides were separated by a 2-40% acetonitrile gradient during 80 min, using 80% acetonitrile in 0.1% formic acid, utilizing a trap Acclaim PepMap 100 nanoViper 2PK C18 (2 cm × 100 µm, 3 µm particle size, Thermo Scientific) in line with an analytical PepMap RSLC C18 ES 902 column (50 cm × 100 µm, 2 µm particle size) at a flow rate of 250 nL/min. The nanoelectrospray voltage used was 1.7 kV with a source temperature of 275 °C. Full scan MS spectra (m/z 375–1500) were acquired in the Orbitrap analyzer after accumulation target value of 3e6, operating at a resolution of  $r = 60,000$ . The 20 most intense peptide ions, with charge states of 2 or higher, were sequentially isolated to a target value of 3e5 and fragmented using high-energy collisional dissociation (HCD) with a normalized collision energy of 27%. The signal threshold for triggering an MS/MS event was set at 1e4. Dynamic exclusion was enabled for 20 seconds, with a repeat count of 1. The maximum injection time was set to 60 ms.

For data analysis, protein identification was performed using MaxQuant version 2.4.7.0 against concatenated databases from *Acanthamoeba castellani* and genome sequences (37,160 protein sequences, 15296385 residues) using the Andromeda search engine. Six-frame translation (6FT) was conducted using the built-in six-frame translation tool in MaxQuant. Carbamidomethylation was set as a fixed modification, while N-terminal acetylation and methionine oxidation were considered variable modifications. Peptide identification parameters included a maximum of two missed cleavages for trypsin/P, a precursor mass tolerance of 10 ppm, and a fragment ion tolerance of 0.02 Da. Protein groups were inferred automatically by the Andromeda engine following the parsimony principle. An FDR of 1% was applied to both protein and peptide identifications, calculated using reverse sequences. Protein quantification was carried out using the LFQ algorithm implemented in MaxQuant, normalizing protein abundance based on razor and unique peptide intensity values, with a minimum ratio count of one. Identifications labeled as 'Reverse,' identified only by 'site,' contaminants or present in a single sample were excluded from further analysis. LFQ intensity values were log2-transformed in Perseus version 2.0.11 for subsequent analyses.

For the additional annotation of the proteins predicted in the proteome, we ran AlphaFold v2.2 using default parameters<sup>21</sup>. The predicted structures were searched against PDB and Uniprot50-minimal v4 structural databases using Foldseek easy-search<sup>22–24</sup>. After acquiring the raw proteomics data, a total of 274 viral proteins were detected, but 20 of them were excluded due to having only one unique peptide, low signal, and/or low coverage. The proteomics data from the particles were deposited in the ProteomeXchange platform - Project accession: PXD063493. Although the data are not yet publicly available, reviewers and editors can access them using this FigShare link - <https://figshare.com/s/39d6cf079fbaf30c14ce>.

## Metagenome mining

To search for Naiavirus relatives across a range of different habitats, we surveyed 16,801 metagenomes from the SPIRE database<sup>25</sup>. We chose

the metagenomes in this set such that they represented a wide range of environments, including freshwater ponds and lakes, marine habitats, soil and sediment samples, engineered systems such as wastewater treatment plants, and others. For this analysis, we downloaded the assemblies that were already available on the SPIRE database and considered only contigs > 5 kb in length. We predicted proteins using Prodigal v. 2.6.3 (-p meta option), and then searched all predicted proteins against the Naiavirus and Hydrivivirus PolB proteins using lastal v. 959 (parameter -m 100). We curated all of the best matches, in general retaining all hits with bit scores > 100. We then used these proteins for phylogenetic analysis together with select reference PolB sequences from viruses in the *Pimascovirales* order (Figure S8). Some of these proteins were placed within a clade together with Naiavirus and Hydrivirus, while others were more closely related to other pimascoviruses. For metagenome-derived PolB sequences that placed within a clade with Naiavirus and Hydrivirus, proceeded to bin the entire metagenomes to retrieve giant virus metagenome-assembled genomes (GVMAGs) using methods similar to those previously described<sup>5</sup>. We first downloaded 10 million reads associated with each metagenome and mapped them against the assembled contigs using coverm v0.4.0 (parameters "-m metabat --min-covered-fraction 0") and then used the coverage file to bin the contigs using MetaBat2 v2.12.1 (parameter "-s 200000")<sup>26</sup>. We then identified bins that contained the contigs on which we found Naiavirus-like PolBs and curated these GVMAGs with ViralRecall v. 2.0 (parameter "-c")<sup>27</sup> removing contigs with scores < 0.

## Orthologous groups analysis

For the analysis of orthologous groups, the genomes of 17 pimascovirus representatives available at the National Center for Biotechnology Information (NCBI) database, along with 11 curated GVMAGs, were used to predict proteins using prodigal v2.6.3<sup>19</sup>. Orthologous groups were generated from the protein predictions from a set of 29 genomes, including Naiavirus, using OrthoFinder 2.5.5 (parameter "-og")<sup>28</sup>. A bipartite network was built from the orthofinder output using Gephi 10.1, followed by manual customization of node sizes<sup>29</sup>.

## Phylogeny analysis

For phylogenetic analysis, proteins of different nucleocytoviruses were predicted using prodigal v2.6.3 with default parameters<sup>19</sup>. The ncldv\_markersearch tool was used to predict marker genes, create the dataset and perform the alignments ([https://github.com/fayward/ncldv\\_markersearch.git](https://github.com/fayward/ncldv_markersearch.git)). A set of seven marker genes that were previously benchmarked was used<sup>4</sup>. Alignments of individual or concatenated datasets were trimmed with trimAl v1.5. (parameters "-gt 0.1" or "-gt 0.5", respectively)<sup>30</sup>. Phylogenetic analysis was performed using IQtree2 2.3.6 (parameters "-m LG + F + R10 -bb 1000 -safe")<sup>31</sup>. The resulting phylogenetic trees were visualized and edited using iTOL<sup>32</sup>.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Ethics

We confirm that our research complies with all relevant ethical regulations. The production of anti-Naiavirus antibodies in mice was conducted with approval from the Animal Experimentation Ethics Committee of Universidade Federal de Minas Gerais (235/2023). This research was registered in the National System for the Management of Genetic Heritage and Associated Traditional Knowledge - SISGEN (number A2291C9).

## Data availability

The raw and analyzed proteome data generated in this study have been deposited in the FigShare (<https://figshare.com/s/>

39d6cf079fbaf30c14ce). The Naiavirus proteome generated in this study have been deposited in the ProteomeXchange database under accession code PXD063493. The raw and analyzed genome data generated in this study have been deposited in the FigShare (<https://figshare.com/s/ef4eddbe9cc10b4789c4>). The Naiavirus genome generated in this study have been deposited in the Genbank database under accession code PV450503. Sequences and protein predictions from the metagenome-assembled genomes generated in this study can be found on Zenodo (<https://zenodo.org/records/16950814>).

## References

1. Sun, T. W. et al. Host range and coding potential of eukaryotic giant viruses. *Viruses* **12**, 1337 (2020).
2. La Scola, B. et al. A giant virus in amoebae. *Science* **299**, 2033 (2003).
3. Queiroz, V. F. et al. Amoebae: hiding in plain sight: unappreciated hosts for the very large viruses. *Annu. Rev. Virol.* **9**, 79–98 (2022).
4. Aylward, F. O., Moniruzzaman, M., Ha, A. D. & Koonin, E. V. A phylogenomic framework for charting the diversity and evolution of giant viruses. *PLoS Biol.* **19**, e3001430 (2021).
5. Moniruzzaman, M., Martinez-Gutierrez, C. A., Weinheimer, A. R. & Aylward, F. O. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat. Commun.* **11**, 1–11 (2020).
6. Schulz, F. et al. Giant virus diversity and host interactions through global metagenomics. *Nat* **2020** 5787795 **578**, 432–436 (2020).
7. Abrahão, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **9**, 749 (Nature Publishing Group, 2018).
8. Souza, F. et al. In-depth analysis of the replication cycle of Orpheovirus. *Virol. J.* **16**, 158 (2019).
9. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
10. Quevillon, E. et al. InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, W116–W120 (2005).
11. Blum, M. et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* **49**, 344–354 (2020).
12. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
13. Rigou, S., Santini, S., Abergel, C., Claverie, J. M. & Legendre, M. Past and present giant viruses diversity explored through permafrost metagenomics. *Nat. Commun.* **2022** 131 **13**, 1–13 (2022).
14. Reed, L. J. & Muench, H. A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* **27**, 493–497 (1938).
15. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
16. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
17. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes de novo assembler. *Curr. Protoc. Bioinforma.* **70**, e102 (2020).
18. Laslett, D. & Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **32**, 11–16 (2004).
19. Hyatt, D. et al. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma.* **11**, 1–11 (2010).
20. Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **2021** **18**, 366–368 (2021).
21. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nat* **2021** 5967873 **596**, 583–589 (2021).
22. Varadi, M. et al. AlphaFold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* **52**, D368–D375 (2024).
23. Varadi, M. et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
24. van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **2023** **42**, 243–246 (2023).
25. Schmidt, T. S. B. et al. SPIRE: a searchable, planetary-scale microbiome REsource. *Nucleic Acids Res.* **52**, D777–D783 (2024).
26. Kang, D. D. et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **2019**, e7359 (2019).
27. Aylward, F. O. & Moniruzzaman, M. ViralRecall-a flexible command-line tool for the detection of giant virus signatures in'omic data. *Viruses* **13**, (2021).
28. Emms, D. M. & Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 1–14 (2019).
29. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. *Third Int. AAAI Conf. Weblogs Soc. Media* 361–362. <https://doi.org/10.1136/qshc.2004.010033> (2009).
30. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
31. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
32. Letunic, I., Bork, P. & Gmbh, B. S. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 1–4. <https://doi.org/10.1093/nar/gkab301> (2021).

## Acknowledgements

We thank the Universidade Federal de Minas Gerais, Departamento de Microbiologia (ICB), Programa de Pós Graduação em Microbiologia-UFMG, Laboratório de Vírus-UFMG and Profs. Betânia Drumond, Erna Kroon and Giliane Trindade for their support. J.S.A., M.R., V.Q., B.N., N.A., T.M., M.B., T.A., H.L., O.H.T. and R.A.L.R. received support from the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES) - including projeto PROEX: processo 88887.893670/2023-00 -, São Paulo Research Foundation (FAPESP), and/or the Minas Gerais State Agency for Research and Development (FAPEMIG). R.E.M. and T.D.M.H. received financial support from the Ministry of Science, Technology, and Innovation of Brazil. T.D.M.H. is supported by Serrapilheira (R-2401-47201). This research utilized facilities at the Brazilian Biosciences National Laboratory (LNBio), part of the Brazilian Center for Research in Energy and Materials (CNPEM), a private non-profit organization under the supervision of the Brazilian Ministry of Science, Technology, and Innovations (MCTI). We acknowledge Adriana Franco Paes Leme and Bianca Alver Pauletti for their assistance during experiments at the Mass Spectrometry Laboratory Facility (Proposal number 20242819). L.C.J.A and this study was supported by the NAVIO project, funded by the National Institutes of Health USA grant U01AI151698, through the United World Arbovirus Research Network (UWARN). We thank the Brazilian Navy and the Health Secretaries of the Mato Grosso do Sul and Mato Grosso states. F.O.A. thanks BioIMG Net supported by the Chan Zuckerberg Initiative (CZI). We also acknowledge the use of the Virginia Tech Advanced Research Computing Center for some of the bioinformatic analyses presented in this study. This work was partially funded by a National Institutes of Health R35 grant (no. 1R35GM147290-01) to F.O.A.. R.E.M., J.P.A.J., L.C.J.A., R.A.L.R., and J.S.A. are CNPq researchers. J.S.A. dedicates this work to Ernani Pereira Abrahão Júnior, his beloved

brother, who passed away this year and is deeply missed, and to all those committed to individual and collective mental health.

## Author contributions

Sample collections: T.M., C.B.C.S. and L.C.J.A. Virus isolation: M.R. Viral experiments: M.R., B.N., N.A., E.R. and J.S.A. Microscopical assays and analyses: M.R., T.A., J.S.A. Electron tomography analyses: T.A., H.L., O.H.T., M.R., J.S.A. Genomic analyses: V.Q., M.B., E.L.H.C., J.P.A.J., A.J., R.A.L.R., J.S.A., F.O.A. Proteomic analyses: T.D.M.H., R.E.M., J.R.C., V.Q., A.J., J.S.A., F.O.A. Study design and coordination: J.S.A. and F.O.A.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-63463-6>.

**Correspondence** and requests for materials should be addressed to Frank O. Aylward or Jônatas Santos Abrahão.

**Peer review information** *Nature Communications* thanks Kiran Kondabagil, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025