Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

The multiple origins of proteins present in tupanvirus particles

Paulo Victor de Miranda Boratto[1], Ana Cláudia dos Santos Pereira Andrade[1], Rodrigo Araújo Lima Rodrigues[1], Bernard La Scola[2] and Jônatas Santos Abrahão[1]

In the last few decades, the isolation of amoebae-infecting giant viruses has challenged established principles related to the definition of virus, their evolution, and their particle structures represented by a variety of shapes and sizes. Tupanviruses are one of the most recently described amoebae-infecting viruses and exhibit a peculiar morphology with a cylindrical tail attached to the capsid. Proteomic analysis of purified viral particles revealed that virions are composed of over one hundred proteins with different functions. The putative origin of these proteins had not yet been investigated. Here, we provide evidences for multiple origins of the proteins present in tupanvirus particles, wherein 20% originate from members of the archaea, bacteria and eukarya.

Addresses
[1] Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 31270-901, Brazil
[2] URMITE, Aix Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, IHU - Méditerranée Infection, AP-HM, 19-21 Boulevard Jean Moulin, Marseille, 13005, France

Corresponding author:
Santos Abrahão, Jônatas (jonatas.abrahao@gmail.com)

Current Opinion in Virology 2019, **36**:25–31

This review comes from a themed issue on **Virus structure and expression**

Edited by **Juliana Reis Cortines** and **Peter Prevelige Jr**

For a complete overview see the Issue and the Editorial

Available online 16th March 2019

https://doi.org/10.1016/j.coviro.2019.02.007

Introduction

Viral particles have a variety of shapes, symmetries, and sizes. The large majority of known viruses have extremely small sizes, with dimensions up to 200 nm in length and relatively simple structures, composed by one or few proteins [1]. This characteristic reflects the genomes of these viruses, which have a reduced number of genes that encode only a few proteins. One group that stands out in this scenario is the giant viruses. These viruses are classified as nucleo-cytoplasmic large DNA viruses (NCLDVs – proposed order Megavirales). They have dimensions larger than 200 nm and extensive genomes reaching up to 2.5 Mb that can encode thousands proteins [2–5].

Most giant viruses, such as mimivirus, pandoravirus, and pithovirus, are associated with free-living amoebae of the genus *Acanthamoeba* [3–5]. Some giant viruses though, have been described infecting flagellate microorganisms such as *Cafeteria roenbergensis* virus and Bodo saltans virus, and both groups are phylogenetically related to the family *Mimiviridae* [6,7]. The giant viruses have extremely complex structures and different shapes or symmetries. The mimiviruses exhibit pseudo-icosahedral particles covered with long glycoproteic fibers reaching ~750 nm in diameter [8,9••], while pandoravirus and pithovirus exhibit an ovoid-to-ellipsoid shape reaching ≥1000 nm in length and contain apical pores [4,5,11•]. Other giant viruses have been described with ovoid particles, such as cedratviruses and orpheovirus [12–14], which exhibit genomic similarities with pithoviruses and together constitute a putative new viral family. Icosahedral viruses are also present among the giant viruses, such as marseilleviruses, faustoviruses, pacmanvirus, and kaumoebavirus, all of which have particles of 220–270 nm in diameter [15–18]. Considering the size and complexity of the particles of these viruses, studies to better characterize their three-dimensional structure through high resolution techniques, such as X-ray crystallography and cryo-electron microscopy, are still limited [9••,10,11•,17,18,19•,20,21].

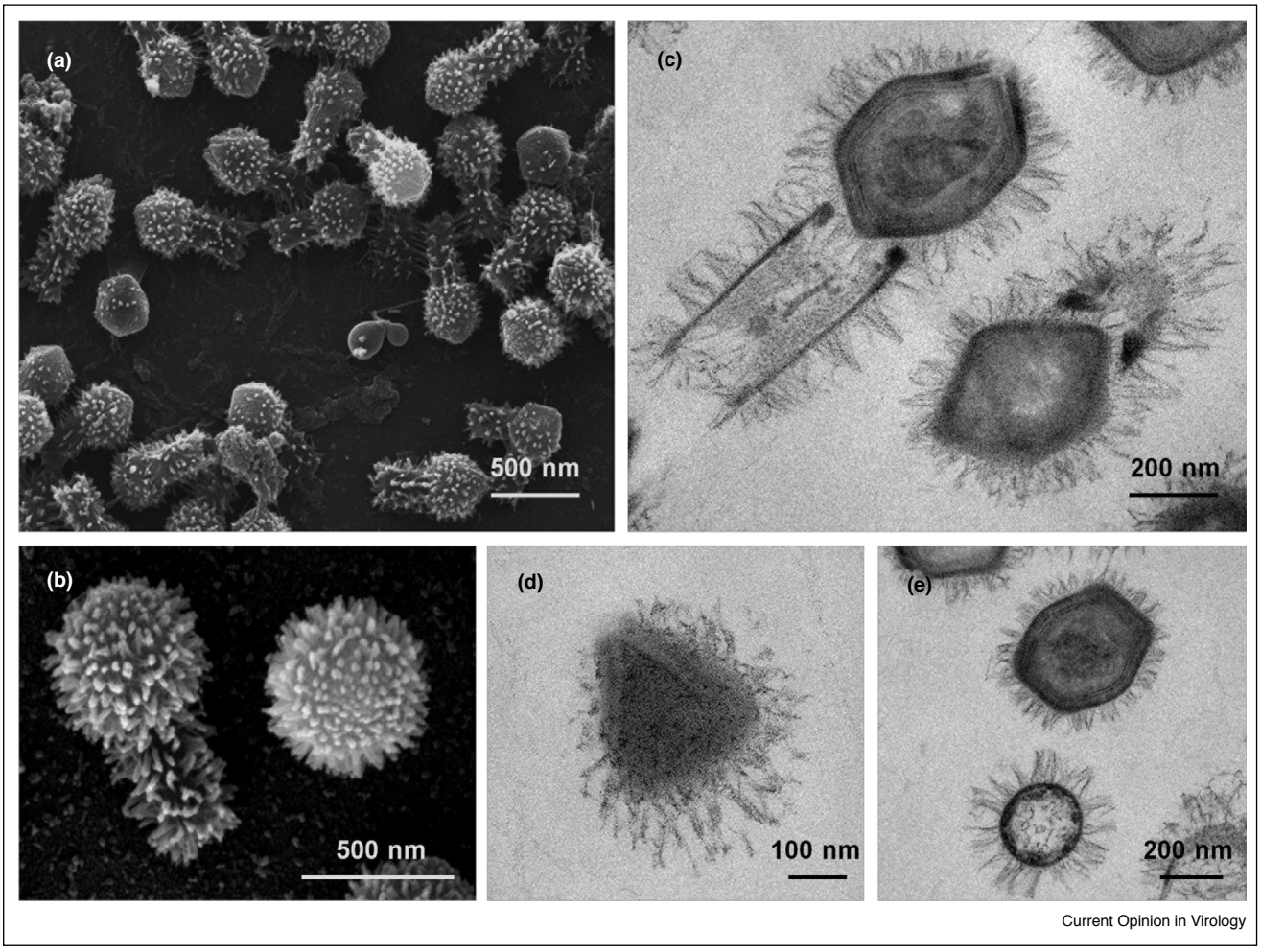
Even more striking is the structure observed for tupanviruses, a new group of viruses within the family *Mimiviridae*. These viruses were recently isolated from extreme environments in Brazil and are capable of infecting a wide variety of amoebae species [22••]. Tupanvirus has more than 1200 genes, and a vast gene arsenal related to the process of protein synthesis, for example, 20 aminoacyl-tRNA synthetases, ~70 tRNAs, and 11 factors related to all translation steps. In addition, it has a cytotoxic profile and causes the host's ribosomal rRNA

to shut down by a still unknown mechanism that is possibly related to the viral particle [22••]. In this review, we explore the peculiar characteristics of the Tupanvirus' particles compared to the other mimiviruses and compile the available proteomics data for the giant viruses. Finally, we perform phylogenetic analyses of the protein coding sequences found in the Tupanvirus' particles to determine the contribution of various taxonomic groups to their genomes.

The complex structure of giant tupanviruses

Tupanviruses are represented by two isolates of giant viruses (tupanvirus soda lake and tupanvirus deep ocean) that were described in 2018 and are putative members of the family *Mimiviridae* [22••]. At the time of their discovery, the impressive features exhibited by the particles were surprising. Tupanviruses not only have characteristically large dimensions for viruses of the *Mimiviridae* family but also a very complex structure marked by the presence of a semi-icosahedral capsid attached to a long cylindrical tail (Figure 1) [22••]. These viruses have a capsid similar to other mimiviruses with a diameter of about 450 nm that is covered by a layer of long fibrils everywhere except in a region named stargate [22••]. The stargate region is a special pentameric vertex that serves as a portal for release of the viral genome [23]. Associated with this capsid there is a tail (also covered by fibrils) that is about 550 nm in length and 450 nm in diameter (fibrils included). Electron microscopy analyses initially suggested a weak form of interaction between these two structures. However, further experiments involving sonication and enzymatic treatment of purified viral particles

Figure 1



Electron microscopy of tupanvirus particles.
(a) Scanning electron microscopy (SEM) image of tupanvirus soda lake (TPV-SL) with viral particles in different positions; **(b)** SEM image visualizing the full structure of the virion (capsid and tail) and the pseudo-icosahedral capsid of about 500 nm; **(c)** Transmission electron microscopy (TEM) image of the internal structures of the particle. Notice the internal membrane in the multi-layered capsid, fibrils covering the whole particle, and the tail attached to it; **(d)** TEM image of the stargate portal of tupanvirus; **(e)** TEM image visualizing a transverse slice of the capsid and tail.

demonstrated that both the capsid and the tail remained tightly attached, hampering complete determination of the nature of interaction between these structures. The average size of these particles was around 1.2 μm, though the tails vary in size and facilitate a substantial plasticity in some of the particles which reach up to 2.3 μm [22••].

In other members of the family Mimiviridae the structure of the virus is known at a somewhat higher level of detail due to the longer period of time over which these viruses were described. In 2009, the structure of the acanthamoeba polyphaga mimivirus (APMV) particle was analyzed in detail by cryo-electron microscopy (cryo-EM) [9••]. It was observed in this study that the particles had a diameter of about 7500 Å. The pseudo-icosahedral capsid is about 5000 Å in diameter and is composed of multiple layers of proteins and lipid membranes surrounding the nucleocapsid. The major capsid protein (MCP) of APMV is formed by two consecutive jelly-roll domains forming capsomers with quasi-sixfold symmetry. As observed for tupanvirus, on the surface of the APMV capsid there is a layer of 1250 Å -long fibers everywhere with the exception of the stargate region. In another study with an APMV-related giant virus, the virus known as Samba virus (SMBV), the authors have performed an in-depth analysis of the structure of the virion by a series of methodologies, including cryo-EM. Apparently, the virion structure seems to be less rigid than the one observed for the particles of APMV. The particles of SMBV are composed by a capsid with a slightly larger diameter (~27 nm) and longer fibers (~30 nm) than the observed for APMV. Furthermore, the structure of SMBV virions appeared to be different from the quasi-icosahedral symmetry of the prototype of the family *Mimiviridae*, evidencing a high level of structural heterogeneity and with unique characteristics, even for individuals belonging to the same viral family [10]. The structure of Cafeteria roenbergensis virus (CroV) was also reconstructed by cryo-EM [19•]. Although CroV infects marine zooplankton and not amoebae, it is phylogenetically related to APMV and belongs to a new genus in the family *Mimiviridae*. Cafeteria roenbergensis virus has an icosahedral capsid with a diameter of 3000 Å, and 30 Å-long surface protrusions which appear to form from loops of its double jelly-roll MCP [19•].

Other giant viruses have had their particles thoroughly analyzed. A member of the faustovirus clade, a group of large viruses that infect amoebae of the genus *Vermamoeba*, have had their particles described using cryo-EM. These viruses are about 2400 Å in diameter and have icosahedral symmetry [21]. It was proposed that the faustovirus capsid is composed of two concentric protein shells. The outer shell is formed by double jelly-roll protein, like those of mimivirus, while the inner shell is formed of different capsid proteins [21]. The internal capsid is flexible, having sizes ranging from 1600 Å to 1900 Å, and contacts the outer shell with protrusions present on its surface [21]. The structure of the largest viral particle known thus far, Pithovirus sibericum, was studied using high-voltage electron cryotomography and energy-filtered cryo-EM [11•]. *Pithovirus* particles are ovoid and can measure up to 2.5 μm in length and 0.9 μm in diameter. At one end, or less often at both extremities, the particles harbor a striated cork-like structure that is characteristic for those isolates [11•]. The *Pithovirus* particles also present a low-density layer that is about 40 nm in thickness on the outermost surface of particles. The density within the particles is higher than expected when considering the 'reduced' size of its genome (600 kbp) and the large volume it occupies indicates a substantial macromolecule component in addition to the genome [11•].

The high degree of detail obtained from cryo-EM studies of giant viruses' particles enabled us to move one step forward in our comprehension of the biology in these complex members of the virosphere. Considering the high level of complexity of tupanvirus particles, this kind of analyses remain to be done. Ultrastructural study of tupanviruses will allow a better characterization of their virions, and possibly generate insights about the nature of the interaction between the capsid and tail. Together with thorough proteomic analyses, the structure of tupanviruses will yield exciting discoveries in the near future.

Proteome of giant viruses

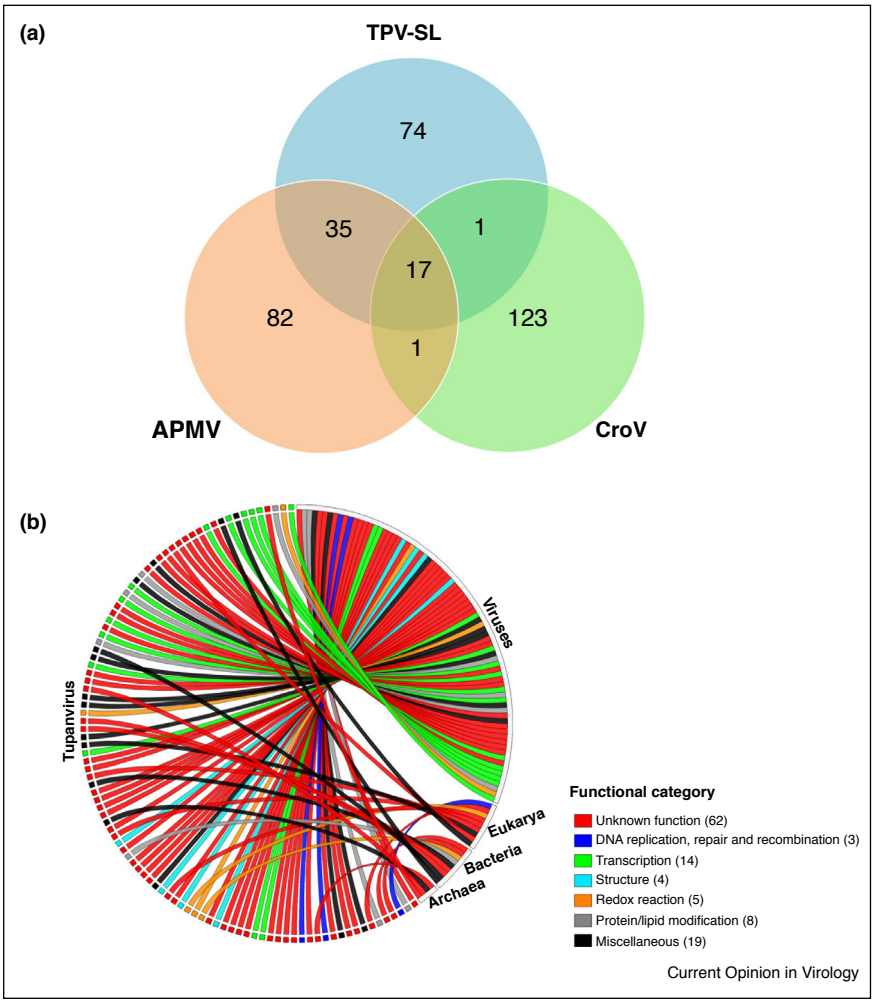
According to many proteomic studies within the known virosphere, the mature particles of amoebal giant viruses are composed of a great number of proteins [4,24••,25,26•]. This has generated hypotheses that seek to explain why only a small fraction of the nearly one thousand proteins encoded by mimiviruses are incorporated into mature virions [24••]. There are larger numbers of proteins present in the mimivirus viral factories (VF) when compared to the protein diversity detected in the mature virion. This observation suggests that these VF are highly elaborate and dynamic structures, with many of these components being specifically required to produce and to propagate the VF structure [24••]. This may partly explain the high number of genes in the genomes of these giant viruses [24••].

Among the amoebal-infecting giant viruses, there are studies into the proteomics of purified viral particles, including APMV, CroV, different pandoravirus isolates, faustovirus E12, pithovirus sibericum, mollivirus sibericum, and also for tupanvirus soda lake [5,16,25,26•,27–30]. In initial comparisons of these viruses' proteomic profiles, it was observed that the proteins predicted to be ORFs primarily belonged to functional categories typical among the members of this group, such as those represented by 'DNA replication, repair and recombination', 'transcription', 'oxidative pathways', 'protein and lipid modification', 'particle structure', and 'nucleotide synthesis'

[5,16,25,26•,27–30]. Generally, the functional category 'transcription factors' is the largest class coding for non-structural proteins in these viruses [4,5,26•]. The number of proteins required to make the mature particle of these viruses was obtained from proteomic works and appears to be broadly similar. Among the giant viruses with available proteomic data, their virions are made with about 130 proteins. This rough value holds even after considering the different techniques used in different analyses and the main problem in viral proteomics: the contamination of analyzed samples with host proteins [29,31]. Interestingly, the size of particles repertoire does not seem to be correlated with the size of the viral genome or the number of proteins that makes up a specific viral particle, for example, faustovirus virion has 164 proteins and pithovirus sibericum virion has 159 proteins.

Proteomic analysis was performed on tupanvirus soda lake (TPV-SL) particles and revealed the presence of 127 proteins constituting the mature virions [22••]. As observed for other giant viruses, an important fraction of these proteins corresponds to sequences of unknown functions and almost 10% of them are related to ORFans (sequences with no match in databases). For the proteins predicted to belong to a functional category, the purified particles of tupanvirus are split into the same groups as the other giant viruses described above [22••]. Comparative analyses of proteomic data from *Mimiviridae* family viruses revealed a set of conserved proteins that compose the mature virions. These conserved proteins were especially related to DNA replication and transcription (e.g. DNA polymerase X family and DNA-directed RNA polymerase) and the major capsid protein, a pivotal

Figure 2



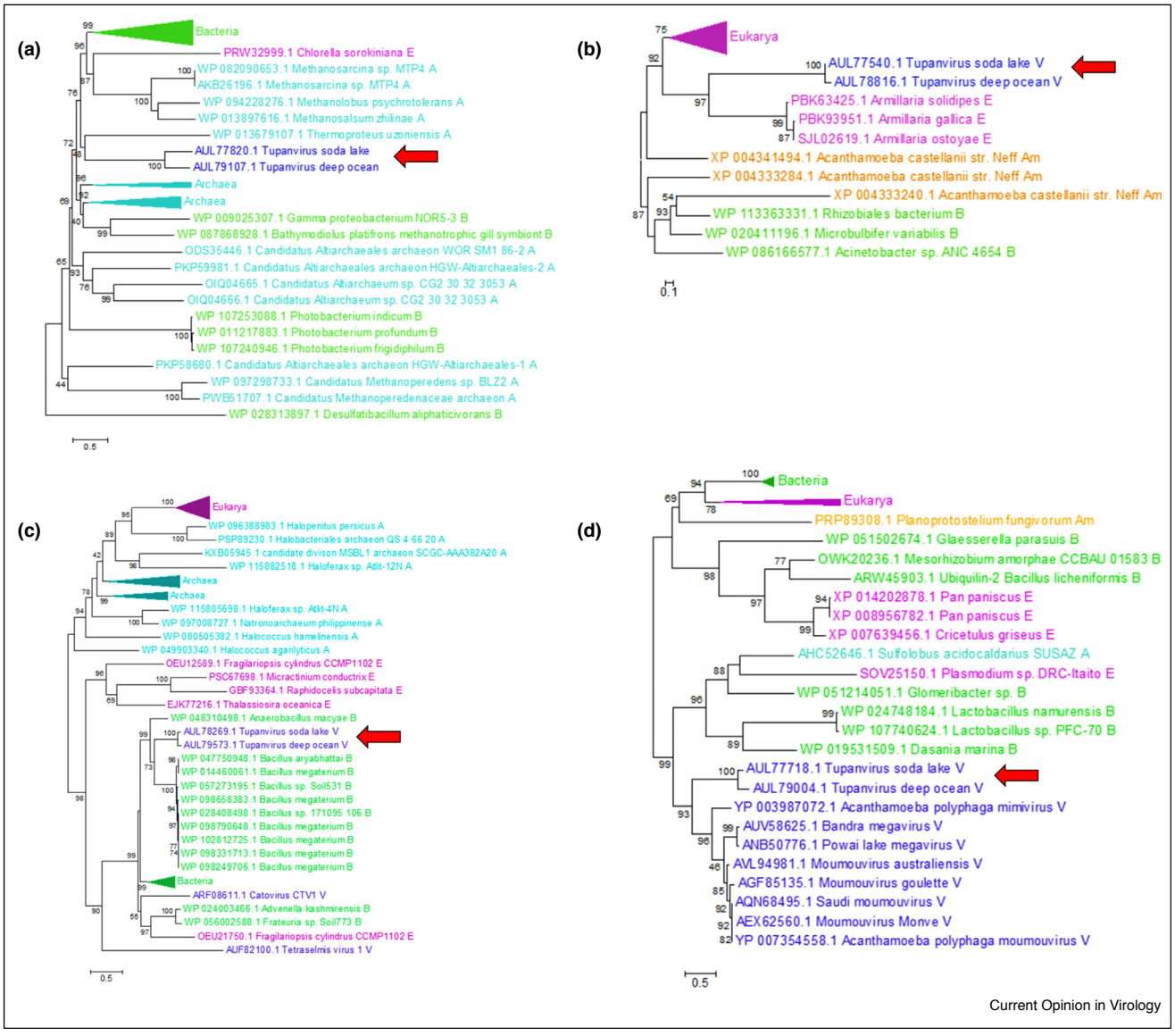
Proteome analysis of tupanvirus soda lake.
(a) Venn diagram of a comparative analysis of the viral particle proteome of TPV-SL, APMV, and CroV containing 127, 136, and 141 proteins, respectively. The analysis was performed using the proteomic data of each virus obtained from the literature [21,25,31] and the software ProteinOrtho with the following parameters: cov = 50%, e-value = 10^{-5}; **(b)** Circos plot representing the putative origin of the proteins constituting the tupanvirus particle. Genes were grouped into functional categories and are depicted in different colors. The number of proteins from each group are specified in the figure.

component in the structure of viruses (Figure 2a). This supports the presence of these components in the last common ancestor of the *Mimiviridae* family [22••,24••,26•]. It is noteworthy that TPV-SL has more proteins in common with APMV than with CroV, reinforcing the close relationship observed in a phylogenetic analysis [22••]. Further, a total of 74 proteins were exclusive to TPV-SL and the majority of these have no known function (Figure 2a). This observation, in addition to the presence of many of these sequences in metagenomic studies indicates participation of these genes in highly coordinated multimolecular processes that are established under tight evolutionary constraints [27]. These data also support the hypothesis that environmental ORFans do indeed correspond to bona fide proteins, however, the role of these still require elucidation in future studies [27].

Contribution of multiple taxonomic groups to the tupanvirus particle structure

One of the most striking features of the known virosphere concerns the genomic and structural characteristics of the

Figure 3



Phylogenetic analyses of the tupanvirus proteome.
Maximum-likelihood trees created with data gathered from the best BLAST-matches within each group of the following organismal (when available) groups: eukarya (30 best-hits), archaea (30 best-hits), amoebozoa (10 best-hits), viruses (10 best-hits), proteobacteria (15 best-hits), and firmicutes (15 best-hits). Inferences were performed for the 127 proteins composing the mature particle of tupanvirus. In these analyses, tupanvirus proteins (red arrows) were grouped with other sequences of **(a)** archaea - gene L352, **(b)** eukarya - gene L996, **(c)** bacteria - gene L1162, and **(d)** other viruses - L250. All of the trees were created with the FastTree 2.1 software and visualized with MEGA 7.

tupanviruses. Their approximately 1.5 Mb double-stranded DNA genomes code for over 1200 proteins, 28% of which have never been identified in other organisms [22••]. Moreover, tupanviruses have many genes related to protein synthesis, including 20 aminoacyl-tRNA synthetases and several translation factors, which appear to have originated from other taxonomic groups [22••]. Similar to tupanviruses, other giant viruses, such as mimiviruses and Marseilleviruses, have genes originating from across the three domains of life, which led to their mosaic genomes [25,32]. It is possible that such genomic mosaicism is reflected in the virion structure.

To avoid artifacts caused by the observation of taxonomic relationships performed with BLAST-searches alone, we have created phylogenetic trees using a maximum-likelihood analysis for all of the proteins predicted present in the purified particles of TPV-SL [22]. These analyses reinforce that multiple groups of organisms contribute to the formation of the virion structure (Figures 2b and 3). The maximum likelihood trees grouped about 20% of the TPV-SL proteome with members of eukarya (9% of total; one third of this 9% originates from amoebae), archaea (3% of total), bacteria (8% of total) (Figures 2 and 3a–c). This result supports data demonstrating the relevance of other groups from the Tree of Life in the evolution of NCLDVs genomes and indicates that parts of these proteins may be incorporated during the formation of the viral particle as well. The high contribution of genes/proteins with a probable bacterial origin has been postulated as a distinctive feature for the NCLDV that infect unicellular eukaryotic hosts, especially for the mimiviruses, Marseilleviruses, and phycodnaviruses [25]. The other 80% of the TPV-SL proteome was related to other groups of viruses, specifically (but not exclusively) to other NCLDVs (Figures 2 and 3d). Finally, since this portion of genes is shared with members of other cellular domains, it may indicate that tupanviruses are not constrained solely to the extreme environments where they have thus far been isolated.

Future perspectives

From more than 300 isolates of giant viruses, only in a dozen the proteins that make up the structure of the particle itself have been analyzed [5,16,25,26•,27–29]. By understanding the components involved in the formation of a virion, it is possible to answer important questions related to the evolution of viruses, their origins, and even what exactly characterizes them. Phylogenetic analyses of all 127 tupanvirus proteins indicate that a substantial portion of the particle structure is influenced by members from across the three Domains of Life. However, it is still an open question whether this portion is relevant when compared to other members of the NCLDVs. To extend this observation to other giant viruses, we must amplify our knowledge on the structure of their virions. This will help so that the currently available viral proteomes can be analyzed and inferred phylogenetically to other cellular groups.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

We thank our colleagues from Grupo de Estudo e Prospecção de Vírus Gigantes (Gepvig) and the Laboratório de Vírus for their excellent technical support. We also thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for scholarship and the Center of Microscopy of Universidade Federal de Minas Gerais. JSA is CNPq researcher and member of a CAPES-COFECUB project.

References

1. Harrison SC: **Principles of virus structure**. In *Fields Virology*. Edited by Knipe DM, Howley PM. Lippincott Williams & Wilkins; 2013:52-86.

2. Colson P, De Lamballerie X, Yutin N, Asgari S, Bigot Y, Bideshi DK, Cheng XW, Federici BA, Van Etten JL, Koonin EV *et al.*: **"Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses**. *Arch Virol* 2013, **158**:2517-2521.

3. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie J-M, Raoult D: **A giant virus in amoebae**. *Science* 2003, **299**:2033.

4. Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, Arslan D, Seltzer V, Bertaux L, Bruley C *et al.*: **Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes**. *Science* 2013, **341**:281-286.

5. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, Lescot M, Poirot O, Bertaux L, Bruley C *et al.*: **Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology**. *Proc Natl Acad Sci U S A* 2014, **111**:1-6.

6. Fischer MG, Allen MJ, Wilson WH, Suttle CA: **Giant virus with a remarkable complement of genes infects marine zooplankton**. *Proc Natl Acad Sci U S A* 2010, **107**:19508-19513.

7. Deeg CM, Chow C-ET, Suttle CA: **The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea**. *eLife* 2018, **7**:1-22.

8. Kuznetsov YG, Xiao C, Sun S, Raoult D, Rossmann M, McPherson A: **Atomic force microscopy investigation of the giant mimivirus**. *Virology* 2010, **404**:127-137.

9. •• Xiao C, Kuznetso YG, Sun S, Hafenstein SL, Kostyuchenko VA, Chipman PR, Suzan-Monti M, Raoult D, McPherson A, Rossmann MG: **Structural studies of the giant mimivirus**. *PLoS Biol* 2009, **7**:958-966.
This is the first study thoroughly describing the structure of a giant virus by using different microscopy techniques, revealing important characteristics of mimivirus particles, including symmetry, capsid organization and T number.

10. Schrad JR, Young EJ, Abrahão JS, Cortines JR, Parent KN: **Microscopic characterization of the Brazilian giant Samba virus**. *Viruses* 2017, **9**:1-16.

11. • Okamoto K, Miyazaki N, Song C, Maia FRNC, Reddy HKN, Abergel C, Claverie JM, Hajdu J, Svenda M, Murata K: **Structural variability and complexity of the giant Pithovirus sibericum particle revealed by high-voltage electron cryo-Tomography and energy-filtered electron cryo-microscopy**. *Sci Rep* 2017, **7**:1-12.

This study reveals the ultrastructural characteristics of pithovirus, revealing viral particles with huge volume and relative small genome, providing insights about the origin and evolution of this group of viruses.

12. Andreani J, Aherfi S, Khalil JYB, Di Pinto F, Bitam I, Raoult D, Colson P, La Scola B: **Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses**. *Viruses* 2016, **8**:1-11.

13. Rodrigues RAL, Andreani J, Andrade ACSP, Machado TB, Abdi S, Levasseur A, Abrahão JS, La Scola B: **Morphologic and genomic analyses of new isolates reveal a second lineage of cedratviruses**. *J Virol* 2018, **92**:1-13.

14. Andreani J, Khalil JYB, Baptiste E, Hasni I, Michelle C, Raoult D, Levasseur A, La Scola B: **Orpheovirus IHUMI-LCC2: a new virus among the giant viruses**. *Front Microbiol* 2018, **8**:1-11.

15. Aherfi S, La Scola B, Pagnier I, Raoult D, Colson P: **The expanding family Marseilleviridae**. *Virology* 2014, **466–467**:27-37.

16. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, Rossmann M, Colson P, Raoult D, La Scola B: **Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae**. *J Virol* 2015, **89**:6585-6594.

17. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, Raoult D, La Scola B: **Kaumoebavirus, a new virus that clusters with Faustoviruses and Asfarviridae**. *Viruses* 2016, **8**:1-10.

18. Andreani J, Khalil JYB, Sevvana M, Benamar S, Di Pinto F, Bitam I, Colson P, Klose T, Rossmann MG, Raoult D *et al.*: **Pacmanvirus, a new giant icosahedral virus at the crossroads between Asfarviridae and Faustoviruses**. *J Virol* 2017, **91**:1-11.

19. • Xiao C, Fischer MG, Bolotaulo DM, Ulloa-Rondeau N, Avila GA, Suttle CA: **Cryo-EM reconstruction of the Cafeteria roenbergensis virus capsid suggests novel assembly pathway for giant viruses**. *Sci Rep* 2017, **7**:1-7.
The study describes the ultrastructure of CroV particle and suggests a new pathway for viral capsids assembly.

20. Ekeberg T, Svenda M, Abergel C, Maia FRNC, Seltzer V, Claverie JM, Hantke M, Jonsson O, Nettelblad C, Van Der Schot G *et al.*: **Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser**. *Phys Rev Lett* 2015, **114**:1-6.

21. Klose T, Reteno DG, Benamar S, Hollerbach A, Colson P, La Scola B, Rossmann MG: **Structure of faustovirus, a large dsDNA virus**. *Proc Natl Acad Sci U S A* 2016, **113**:6206-6211.

22. •• Abrahão J, Silva L, Silva LS, Khalil JYB, Rodrigues R, Arantes T, Assis F, Boratto P, Andrade M, Kroon EG *et al.*: **Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere**. *Nat Commun* 2018, **9**:1-12.
Descripton of isolation and thorough characterization of tupanviruses, a new group of giant viruses exhibiting unprecedent structural and genomic features, including unusual phenotypes with different hosts cells.

23. Zauberman N, Mutsafi Y, Halevy D, Ben, Shimoni E, Klein E, Xiao C, Sun S, Minsky A: **Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus**. *PLoS Biol* 2008, **6**:1104-1114.

24. •• Fridmann-Sirkis Y, Milrot E, Mutsafi Y, Ben-Dor S, Levin Y, Savidor A, Kartvelishvily E, Minsky A: **Efficiency in complexity: composition and dynamic nature of Mimivirus replication factories**. *J Virol* 2016, **90**:10039-10047.
An update on the mimivirus particle proteome and the first study to evaluate the proteome of a giant virus viral factory, suggesting a dynamic nature of these structures.

25. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG *et al.*: **Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms**. *Proc Natl Acad Sci U S A* 2009, **106**:21848-21853.

26. • Fischer MG, Kelly I, Foster LJ, Suttle CA: **The virion of Cafeteria roenbergensis virus (CroV) contains a complex suite of proteins for transcription and DNA repair**. *Virology* 2014, **466–467**:82-94.
A proteomic study revealing over 140 proteins with different functions composing the structure of the CroV particle.

27. Renesto P, Abergel C, Decloquement P, Moinier D, Azza S, Ogata H, Fourquet P, Gorvel J-P, Claverie J-M: **Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products**. *J Virol* 2006, **80**:11678-11685.

28. Aherfi S, Andreani J, Baptiste E, Oumessoum A, Dornas FP, Andrade ACSP, Chabriere E, Abrahao J, Levasseur A, Raoult D *et al.*: **A large open pangenome and a small core genome for giant Pandoraviruses**. *Front Microbiol* 2018, **9**:1-13.

29. Legendre M, Lartigue A, Bertaux L, Jeudy S, Bartoli J, Lescot M, Alempic J-M, Ramus C, Bruley C, Labadie K *et al.*: **In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba**. *Proc Natl Acad Sci U S A* 2015, **112**:1-9.

30. Legendre M, Fabre E, Poirot O, Jeudy S, Lartigue A, Alempic J-M, Beucher L, Philippe N, Bertaux L, Labadie K *et al.*: **Diversity and evolution of the emerging Pandoraviridae family**. *Nat Commun* 2018, **9**:1-2.

31. Moreira D, Brochier-Armanet C: **Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes**. *BMC Evol Biol* 2008, **8**:1-10.

32. Claverie JM, Abergel C, Ogata H: **Mimivirus**. *Curr Top Microbiol Immunol* 2009, **328**:89-121.